

UNIVERSITA' DEGLI STUDI DI NAPOLI FEDERICO II

**DOTTORATO DI RICERCA IN BIOCHIMICA E
BIOLOGIA CELLULARE E MOLECOLARE**

TESI DI DOTTORATO, 2006

GENOMA E AMBIENTE: TEMPERATURA E METILAZIONE

DOTTORANDA:

**ANNALISA VARRIALE
LABORATORIO DI EVOLUZIONE
MOLECOLARE
STAZIONE ZOOLOGICA
ANTON DOHRN**

RELATORE:

Prof. GIUSEPPE GERACI

CORRELATORE:

Prof. GIORGIO BERNARDI

COORDINATORE:

Prof. GIUSEPPE D'ALESSIO

INDICE

	pagina
SOMMARIO	i
ABSTRACT	ii
1-INTRODUZIONE	1
<i>1.1 -I pesci</i>	7
<i>1.2 -I rettili</i>	8
<i>1.6 -Ambiente e termostabilità dell'RNA ribosomale (rRNA) 18S</i>	9
2-MATERIALI E METODI	12
<i>2.1 DNA e campioni di tessuto biologico</i>	12
<i>2.2 Estrazione del DNA genomico e analisi per ultracentrifugazione</i>	12
<i>2.3 Idrolisi enzimatica del DNA</i>	12
<i>2.4 Analisi con RP-HPLC</i>	13
<i>2.5 Analisi dei geni ortologi</i>	14
<i>2.6 Analisi delle sequenze di RNA ribosomale 18S in vertebrati</i>	15
3-RISULTATI	16
<i>3.1.1 Analisi del livello di metilazione in pesci</i>	16
<i>3.1.2 Analisi del livello di metilazione nei rettili</i>	22

<i>3.1.3 Analisi del livello di metilazione nei mammiferi</i>	25
<i>3.2 Correlazione del livello di metilazione con la lunghezza del genoma</i>	28
<i>3.3 Livello di CpG in geni ortologhi</i>	29
<i>3.4 Correlazione del livello di metilazione con tassonomia e filogenia</i>	32
<i>3.5 Correlazione tra livello di metilazione e temperatura corporea</i>	32
<i>3.6 Correlazione tra livello di GC derivante dall'analisi dei nucleosidi e dalla centrifugazione in gradiente di densità di cloruro di cesio (CsCl)</i>	33
<i>3.7 Analisi dell'RNA ribosomale 18S endotermi ed endotermi</i>	35
4.DISCUSSIONE	37
5.CONCLUSIONE	41
3-RINGRAZIAMENTI	42
3-BIBLIOGRAFIA	43

Sommario

La temperatura è un fattore ambientale che influenza la vita degli organismi. Nei vertebrati essa ha determinato lo sviluppo di due diverse strategie adattative: l'endotermia e l'ectotermia, ciascuna con i propri vantaggi. Nel nostro lavoro lo scopo principale era mostrare che l'ambiente ha un'influenza su tutto il genoma, inteso come la globalità del DNA nucleare e non solo come contenitore di geni. Eravamo quindi interessati a trovare anche nel DNA delle caratteristiche che discriminassero questi diversi gruppi di vertebrati a temperatura corporea diversa. In particolare ci siamo occupati di studiare l'effetto della temperatura sul livello di metilazione globale e sulla composizione delle regioni di geni per l'RNA ribosomale 18S che sono preposte alla formazione di strutture secondarie. Abbiamo trovato una correlazione negativa tra la temperatura corporea ed i livelli di metilazione nei vertebrati. Inoltre abbiamo visto che la temperatura ha anche un effetto sulla composizione delle parti strutturate del ribosomale 18S. In base ai risultati ottenuti possiamo affermare che l'ambiente ha un effetto sull'adattamento e l'evoluzione dei genomi.

Abstract

Temperature is an environmental factor that affects the life of the organisms. In vertebrates, it has determined the development of two different ways of adaptation: endothermy and ectothermy, each having some advantages.

In our project, the main aim was to show that the environment influences the whole genome that is the whole nuclear DNA and not only the group of genes. Then, we were interested in finding features distinguishing also the DNA of the two types of vertebrates having different body temperatures. More in particular, we have studied the effect of body temperature on the global methylation level and on the nucleotide composition of regions of the genes for ribosomal RNA 18S, which, after the transcription, fold in secondary structures. We found a negative correlation between temperature and methylation levels in vertebrates. Moreover, we found also an effect of temperature on the composition of the stems in ribosomal RNA 18S. From these results, we can claim that environment has an effect on genome adaptation and evolution.

1.Introduzione

La temperatura è un fattore ambientale che riveste una grande importanza nella vita di procarioti ed eucarioti. A livello macroscopico, la temperatura dell'ambiente esterno influenza la temperatura corporea degli organismi, ne accelera il metabolismo e di conseguenza il successo ecologico e riproduttivo. La temperatura corporea è talmente discriminante per i vertebrati da essere stati suddivisi dai fisiologi in due gruppi: endotermi ed ectotermi. I primi (mammiferi e uccelli) sono capaci di ricavare il calore necessario dal proprio metabolismo e sono detti anche "a sangue caldo" o omeotermi in quanto possono mantenere la propria temperatura elevata e costante. Gli ectotermi (pesci, anfibi, rettili), invece, ottengono il calore dall'ambiente esterno ed hanno quindi una temperatura variabile, per cui vengono detti "a sangue freddo" o eterotermi.

Anche a livello microscopico, la temperatura, in questo caso quella dell'ambiente cellulare, costituisce una pressione selettiva in grado di influenzare la termostabilità di DNA, RNA e proteine (Bernardi e Bernardi, 1986). Queste macromolecole infatti sono sensibili alla temperatura elevata, e se vi sono sottoposte per tempi molto lunghi, necessitano di adattarsi incrementando la propria stabilità. Nel corso dell'evoluzione degli organismi che vivono in ambienti molto caldi (es. batteri termofili) o che hanno una temperatura corporea elevata (vertebrati endotermi) sono stati selezionati genomi che hanno subito un incremento di GC (guanina + citosina) che hanno un effetto termostabilizzante sul DNA e di conseguenza su RNA e proteine da

esso codificate. Per questo, il livello di GC è una caratteristica dei genomi correlata alla temperatura sia nei batteri che nei vertebrati.

Nel caso dei batteri si è dimostrato che il livello genomico di GC è maggiore in quelli che vivono a temperature elevate rispetto a quelli che vivono in ambienti meno caldi (Musto et al., 2004). Nei vertebrati, invece, l'aumento in GC riguarda una piccola parte del genoma, quella più ricca in geni (genome core). Questo aumento, chiamato transizione compositazionale intragenomica (major shift), nel corso dell'evoluzione si è verificato sia negli antenati di mammiferi e di uccelli che negli antenati dei vertebrati a sangue freddo (vedi Bernardi, 2005 per una rivista). Nei vertebrati a sangue caldo, tuttavia, l'incremento è stato maggiore rispetto ai vertebrati a sangue freddo perché hanno una temperatura corporea più alta che induce una pressione selettiva più forte sul DNA. Oltre che la composizione nucleotidica, la temperatura influenza anche un'altra proprietà del DNA dei vertebrati: il livello di metilazione. La metilazione del DNA nei vertebrati è coinvolta in numerosi processi quali regolazione dell'espressione genica, imprinting, inattivazione del cromosoma X (Costello e Plass, 2001). Biochimicamente, la metilazione è un processo mediato dall'enzima DNA-metiltransferasi (DNMT). Nelle cellule eucariotiche superiori sono state identificate la DNMT1, DNMT3a e DNMT3b. La prima è coinvolta nella metilazione delle eliche appena sintetizzate del DNA dopo la replicazione, ed è per questo definita come metiltransferasi di "mantenimento". Quest'enzima ha un segnale di localizzazione nucleare e un altro di legame al complesso di replicazione, provvedendo durante la sintesi replicativa a mantenere lo stato metilato. Essa quindi presiede alla metilazione globale del genoma. Le DNMT3a e DNMT3b

hanno invece la capacità di metilare il DNA non metilato, e quindi presiedono alla metilazione *de novo* (Bestor, 2000). Quest'ultima è tipica dell'imprinting e delle regolazioni dello sviluppo. Tutte le DNMT catalizzano il trasferimento di un gruppo metilico dalla S-adenosilmetionina (SAM) al carbonio-5 delle citosine (Fig. 1) seguite da una guanina (siti CpG). Queste vengono metilate divenendo 5-metilcitosine. Esse vengono indicate come 5mC e i siti CpG metilati come mCpG.

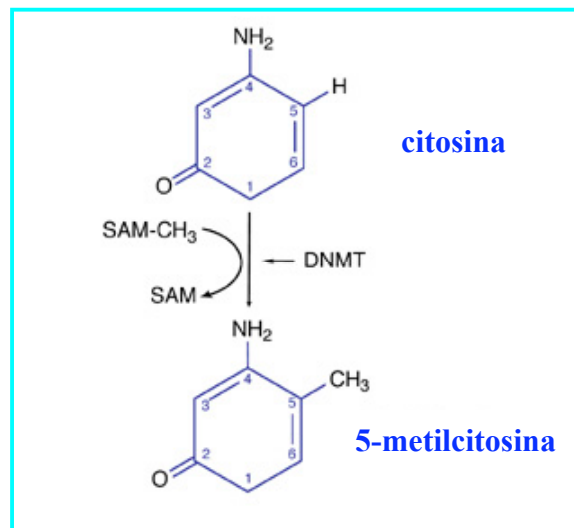


Fig. 1- Formazione della 5-metilcitosina.

Da molti anni è noto che nel genoma dei vertebrati la frequenza di CpG è molto inferiore rispetto all'atteso quasi ovunque (Josse et al., 1961; Swartz et al., 1962) se si escludono le isole CpG (sequenze regolatrici associate a geni). Fu proposto che tale diminuzione fosse dovuta alla metilazione (Doskocil e Sorm, 1962; Grippo et al., 1968). I vertebrati sono infatti organismi il cui DNA è metilato e mC è un punto soggetto a mutazioni (Salser, 1977; Bird, 1980) perché può deaminare spontaneamente a T (Fig. 2).

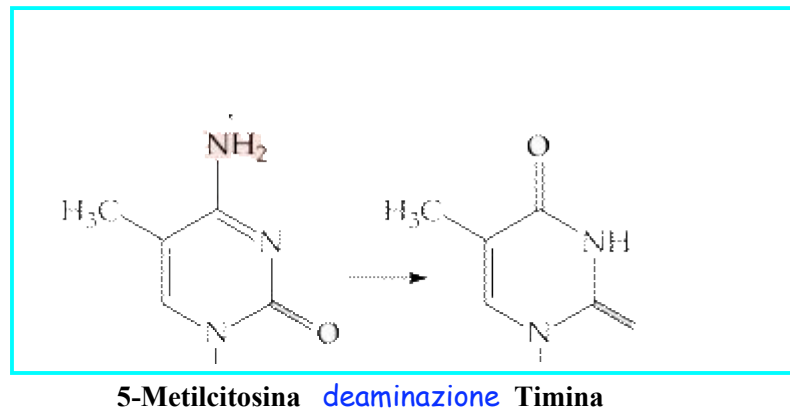


Fig. 2- Deaminazione della 5-metilcitosina.

La velocità della deaminazione spontanea fu studiata *in vitro* a varie temperature superiori a 50°C (Ehrlich et al., 1986) e si vide che essa è tanto più alta quanto più alta è la temperatura. Questa correlazione fu dimostrata anche in batteri in cui si determinò che a 37°C la velocità di deaminazione spontanea era $5.8 \times 10^{-13} \text{s}^{-1}$ (Shen et al., 1994). A una velocità così bassa, per stabilire se la correlazione esiste anche nei vertebrati si dovevano confrontare i livelli di 5mC in vertebrati che da tempi molto lunghi hanno temperature corporee diverse. Per questo motivo, nel nostro laboratorio (Jabbari et al., 1997; Jabbari, 2004) sono stati confrontati i livelli di metilazione e di CpG di vertebrati a sangue freddo ed a sangue caldo. Da questo studio si è trovato che i genomi di pesci e anfibi sono metilati in media il doppio rispetto a quelli di mammiferi e uccelli (Fig. 3) ed inoltre possiedono il doppio del livello di CpG (Fig. 4). Ciò è in accordo con l'ipotesi che a una temperatura più bassa corrisponde una velocità di deaminazione più bassa e quindi una minore perdita di 5mC e di CpG. La deaminazione spontanea rappresenta il maggior fattore di perdita della metilazione in quanto enzimi deaminasi abbiano una funzione strettamente confinata a determinati geni (Morgan et al., 2004), e con una velocità 6 ordini di grandezza inferiore rispetto alla deaminazione spontanea.

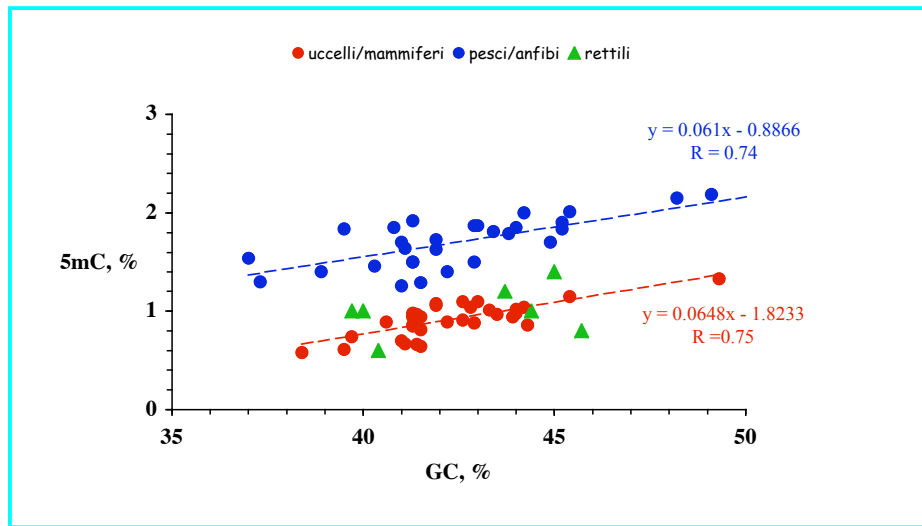


Fig. 3- Andamento della metilazione in funzione del livello di GC nei genomi totali di vertebrati a sangue freddo ed a sangue caldo (Jabbari et al., 1997). Vengono date le equazioni della retta e coefficienti di correlazione (R).

I dati di mammiferi e uccelli non sono sovrapposti a quelli di anfibi e pesci, suggerendo l'esistenza di due equilibri. Questo indicava che la perdita di metilazione è un processo che va all'equilibrio.

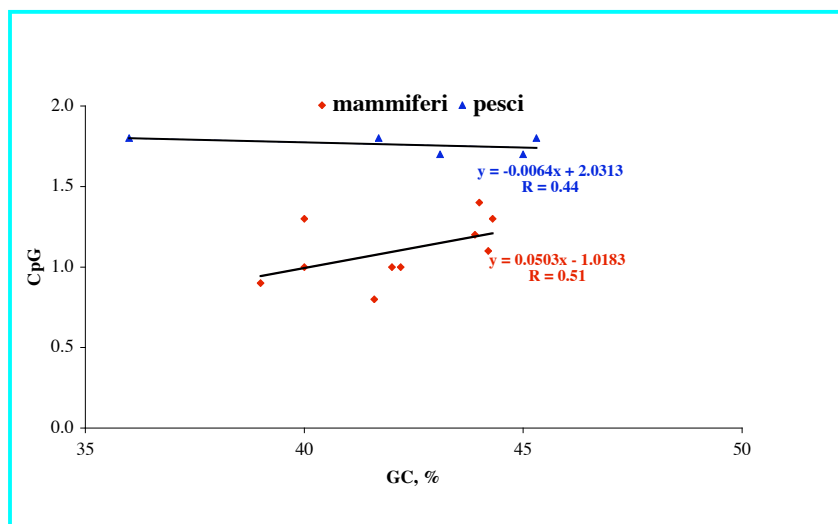


Fig. 4- Diversi livelli di CpG in genomi poco metilati (mammiferi; punti rossi) e molto metilati (pesci; punti blu) in funzione del livello di GC (modificato da Jabbari e Bernardi, 2004).

Nei DNA dei mammiferi, ad esempio, nonostante la deaminazione, circa il 2-10% dei residui di citosina sono metilati (Ehrlich e Wang, 1981). Questa percentuale relativamente alta può essere considerata conseguenza o dell'equilibrio tra perdita e creazione di siti CpG (Sved e Bird, 1990) o dell'azione di uno specifico sistema di riparazione (Kramer et al., 1984; Brown e Jiricny, 1987; Rusmintratip e Sowers, 2000). L'importanza della metilazione del DNA richiede quindi che essa venga protetta tramite un sistema di riparazione che ne eviti la perdita totale (Coulondre et al., 1978; Wang et al., 1982; Ehrlich et al., 1986; Shen et al., 1994) anche nelle regioni non codificanti. Attualmente non è ben noto il motivo per cui dev'essere mantenuta la metilazione anche in queste regioni, ma ciò potrebbe essere dovuto a fattori strutturali o di silenziamento di sequenze trasponibili. Infatti, il maggiore rapporto di mCpG/CpG non si trova nel nucleo del genoma (genome core), che contiene la maggior densità di geni ed è la parte più ricca in GC, ma nel deserto del genoma (genome desert; Jabbari e Bernardi, 2004), che costituisce più dell'80% del genoma (Bernardi, 2005). Questo vuol dire che quando parliamo di metilazione globale ci riferiamo essenzialmente proprio alle regioni che contengono meno geni. Inoltre questo spiega perché la temperatura da un lato è positivamente correlata con il livello di GC e dall'altro lato è negativamente correlata con il livello di 5mC e di CpG. Infatti, il GC negli endotermi è aumentato nel genome core dove i CpG non essendo metilati non vengono persi (correlazione positiva tra Temperatura, GC e CpG). Nel genome desert, invece, c'è più metilazione e più probabilità che mCpG vengano mutati a TpG (correlazione negativa tra Temperatura, 5mC e CpG). Il discorso per gli ectotermi è diverso in quanto la temperatura, e di conseguenza la velocità di

deaminazione, sono più basse, e quindi pochi mCpG vengono persi anche nel genome desert.

Nel presente studio abbiamo voluto confermare l'influenza dell'ambiente sul genoma e più in particolare della correlazione tra temperatura e metilazione. Inoltre abbiamo voluto affinare la ricerca estendendola a pesci che vivono a temperature molto diverse tra loro, ed ai rettili, che notoriamente hanno temperature molto variabili. Inoltre abbiamo mostrato che è possibile trovare differenze anche indipendentemente dalla vicinanza filogenetica.

1.1- I pesci

I pesci rappresentano un soggetto ideale per lo studio dell'effetto della temperatura corporea sulla struttura del genoma poiché la loro temperatura segue l'andamento di quella ambientale caratteristica delle acque in cui sono immersi (termoconformer). Nel presente lavoro, abbiamo analizzato pesci a temperature che vanno da circa 0°C per i pesci artici ed antartici ai 20°-30°C per i temperati e tropicali.

Tra i pesci antartici abbiamo analizzato prevalentemente rappresentanti dei Nototenioidi (ordine *Perciformes*) che rappresentano la componente faunistica più massiccia del Polo Sud. Essi iniziarono a diversificarsi durante l'isolamento della piattaforma continentale Antartica che avvenne nel medio Terziario. In quel periodo infatti si formò il Fronte Polare Antartico (APF) una corrente circolare che creò una barriera naturale risultante nell'isolamento termale ed idrografico dell'Oceano Australe (Kennett, 1982). Attualmente il Polo Sud offre un ambiente unico quanto a stabilità termica con una temperatura annuale compresa tra -1.5°C e +1.5°C. Questa condizione ha

ridotto quasi del tutto le migrazioni dei pesci che adattarono il proprio metabolismo al progressivo raffreddamento (Eastman e McCune, 2000) sviluppando caratteristiche fisiologiche peculiari quali la produzione di glicopeptidi anticongelamento (AFGPs; De Vries e Wohlschlag, 1969), ed un sistema speciale di trasporto dell'ossigeno nel sangue (di Prisco, 2000).

Sebbene molti Nototenioidei vivano nell'Antartico, almeno 28 specie sono sub-Antartiche in quanto vivono lungo le coste Sudamericane e Neozelandesi (Eastman, 2005) dove si trovano a contatto con temperature variabili e più alte (4°-10°C) rispetto alle loro controparti antartiche. Di questo gruppo fanno parte sia specie che si sono diversificate dopo la formazione dell'APF (Bargelloni et al., 2000) che specie di origine antartica che sono riuscite a migrare all'esterno (Stankovic et al., 2002) e conservano i geni per le AFGPs. Nell'artico l'isolamento è più blando e la fluttuazione della temperatura è maggiore, facilitando la migrazione dell'ittiofauna. Studi condotti sui pesci artici hanno mostrato l'esistenza di numerose differenze tra gli ambienti marini Artico ed Antartico quanto a clima, migrazioni, e storia evolutiva. Nonostante ciò, esistono casi di evoluzione convergente, come la presenza di AFGPs (Chen et al., 1997).

Nel caso dei pesci tropicali analizzati, la temperatura dell'acqua è circa 15-25°C, mentre i temperati vivono a 8-15°C.

1.2 -I rettili

I rettili sono vertebrati la cui termofisiologia è particolarmente interessante. Essi, respirando aria, hanno bisogno di compensare più ampie fluttuazioni di temperatura rispetto agli ectotermi acquatici, e quindi tollerano

un più ampio raggio di temperature. Questo fatto, unito alle loro peculiari strategie comportamentali e fisiologiche di termoregolazione, fa sì che la loro temperatura corporea sia in molti casi particolarmente variabile, e costante in altri casi. I tre ordini di rettili studiati (coccodrilli, tartarughe e squamati) presentano caratteristiche fisiologiche diverse.

I coccodrilli possiedono un alto grado di termoregolazione (Seebacher et al., 1999) e sono capaci di produrre calore metabolico sufficiente ad elevare la loro temperatura corporea al di sopra della temperatura dell'acqua. Un articolo recente, inoltre, evidenzia la similarità tra il cuore dei coccodrilli e degli endotermi, mostrando la possibilità di un antenato endotermo dei moderni coccodrilli ed alligatori (Summers, 2005).

Le tartarughe, sia terrestri che acquatiche possono raggiungere temperature superiori a quelle dell'acqua e dell'aria grazie all'esposizione al sole al sole o come risultato dell'accumulo di calore di provenienza metabolica (Mrosovsky e Pritchard, 1971).

La maggior parte delle lucertole si espone al sole durante il giorno cosicché esse possano mantenere una temperatura relativamente alta e costante nell'intervallo 30-35°C (Avery, 1982). Nei serpenti, i maggiori rapporti superficie-volume e la forma allungata fanno sì che la relazione termica con l'ambiente tenda ad essere meno precisa rispetto a quella delle lucertole.

1.3 -Ambiente e termostabilità dell'RNA ribosomale (rRNA) 18S

Gli RNA ribosomali sono i maggiori costituenti dei ribosomi e formano complesse strutture secondarie: le basi che compongono le regioni più importanti si appaiano formando strutture a stelo (stem) mentre le altre restano

in singolo filamento formando anse (loop). Uno schema per un RNA ribosomiale 18S eucariotico è rappresentato in Fig. 5.

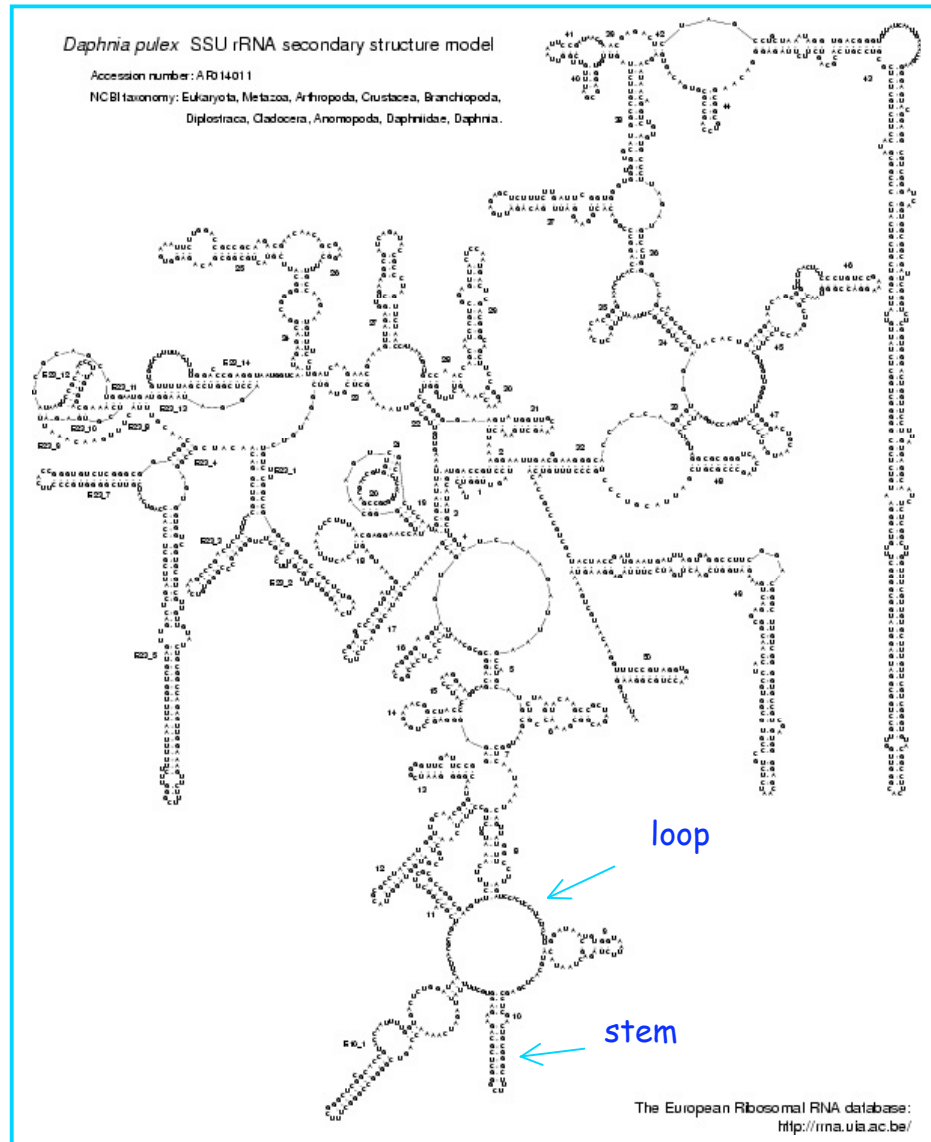


Fig. 5- Modello di un RNA ribosomiale 18S eucariotico (*Daphnia pulex*). Tratto dal sito web dell'European Ribosomal RNA Database.

Negli stem il livello di GC è circa 80%, una percentuale ben più alta di quella che si trova in sequenze a composizione casuale e in media negli RNA non codificanti (50%). Questa è un'indicazione che negli stem le coppie GC sono importanti in relazione alla stabilità. In batteri, ad esempio, Galtier e

Lobry (1997) mostrarono che il livello di coppie GC negli stem dell'RNA ribosomale e transfer è correlato positivamente alla temperatura ottimale di crescita. Essi affermarono perciò che l'incremento di coppie GC con la temperatura fosse necessario per la stabilizzazione della struttura secondaria. Tuttavia, non trovando relazione tra la temperatura e il GC genomico, rifiutarono l'ipotesi della selezione naturale come fattore guidante. Armbruster (2001) ha mostrato come virtualmente in alcuni eucarioti una stessa molecola di RNA ribosomale 18S possa assumere una diversa struttura secondaria a seconda della temperatura alla quale viene generata. In seguito vennero effettuati degli studi anche sui vertebrati. Inizialmente, Xia et al. (2003) avevano osservato che il livello di GC delle sequenze dell'RNA 18S era più basso nei vertebrati a sangue freddo che in quelli a sangue caldo. Successivamente, Wang et al. (2006) hanno confermato l'esistenza della correlazione temperatura-GC nei batteri e l'hanno mostrata anche negli *Archaea*. Tuttavia, estendendo l'analisi ai vertebrati a sangue caldo ed a sangue freddo, essi hanno riscontrato differenze piccole e non confinate alle regioni appaiate. Da questi risultati, gli autori hanno escluso che l'aumento del GC nell'RNA ribosomale fosse dovuto ad un adattamento evolutivo.

2. Materiali e Metodi

2.1 DNA e campioni di tessuto biologico

DNA o campioni di tessuto sono stati ottenuti da numerose specie di pesci, rettili e mammiferi. Per ogni specie abbiamo analizzato un solo tessuto (soprattutto fegato o muscolo) poiché in una stessa specie le differenze nei livelli di 5mC tra i tessuti sono poco rilevanti.

2.2 Estrazione del DNA genomico e analisi per ultracentrifugazione

Abbiamo estratto il DNA dai tessuti utilizzando il metodo del fenolo-cloroformio come descritto da Sambrook et al., (1989) o usando il kit Genomix (Talent). Abbiamo poi quantificato il DNA misurando l'assorbanza a 260 nm e ne abbiamo controllato la purezza mediante lettura dell'assorbanza a 230 e 280 nm e calcolo dei rapporti A_{260}/A_{230} e A_{260}/A_{280} . Al fine di individuare l'eventuale presenza di satelliti nel DNA genomico, abbiamo condotto un'ultracentrifugazione analitica in Cloruro di Cesio (CsCl) sui campioni mediante un apparecchio Beckman modello Optima XL-A. Abbiamo infine calcolato la densità modale di galleggiamento (ρ_0) e il rapporto molare di deossiguanosina + deossicitidina (GC) come descritto nel 1976 da Thiery et al. I profili analitici in CsCl dei pesci e dei rettili sono presentati come Materiale Supplementare negli articoli da noi pubblicati (Varriale e Bernardi, 2006a,b).

2.3 Idrolisi enzimatica del DNA

La procedura che abbiamo seguito è una modifica del metodo descritto da Gehrke et al. (1984). Abbiamo disciolto da 3 a 10 μg di DNA in acqua, e li

abbiamo fatti denaturare per riscaldamento a 100°C per due minuti e subito dopo raffreddati in ghiaccio. Ai campioni abbiamo aggiunto 50 µl di sodio acetato (NaAc) 30 mM a pH 5.3, 2.5 µl di zinco solfato 20mM, e 10 unità di nucleasi P1 (1U/µl in 30 mM NaAc pH 5.3). L'idrolisi del DNA è stata condotta per la notte a 37°C. Al termine abbiamo aggiunto 10 unità di fosfatasi alcalina di timo di vitello (CIAP) e 20 µl di tampone (Roche) e abbiamo incubato il tutto per ulteriori tre ore a 37°C. Abbiamo quindi filtrato i 2'-deossinucleosidi risultanti con dei microfiltri con MWCO 3 kDa (Millipore) e infine li abbiamo iniettati sulla colonna HPLC.

2.4 Analisi con RP-HPLC

Come metodo di separazione dei deossinucleosidi abbiamo scelto la cromatografia liquida ad alta prestazione a fase inversa (RP-HPLC) perché è più sensibile e più riproducibile rispetto alle altre tecniche (Gehrke et al., 1984), e perché essa permette la separazione dei deossiribonucleosidi dagli eventuali ribonucleosidi originatisi da RNA contaminante. Per potere riconoscere i singoli nucleosidi abbiamo registrato i tempi di ritenzione degli standard. Una modifica del metodo descritto da Gehrke et al. (1984) è stata sviluppata usando una colonna C18 a fase inversa di 25 cm (Beckman-Coulter). Il tampone A è KH₂PO₄ 50 mM filtrato su Millipore GS-22 (0.22 µm). Il solvente B è 95% (v:v) metanolo (HPLC grade; J.T. Baker).

Abbiamo adottato un sistema di gradiente a scalino con due tamponi seguito da un gradiente lineare a 100% di solvente B in 10 minuti, dal lavaggio della colonna, e dal riequilibrio. Il flusso è stato di 1.2 ml/min e la

temperatura della colonna è stata tenuta costantemente a 35°C mediante un apparecchio Gecko.

Abbiamo identificato i deossiribonucleosidi (Fig. 6) a λ_{MAX} e poi li abbiamo quantificati determinando l'Area $_{MAX}$ per ciascuno mediante uno spettrofotometro Diode Array (Detector 168, Beckman-Coulter).

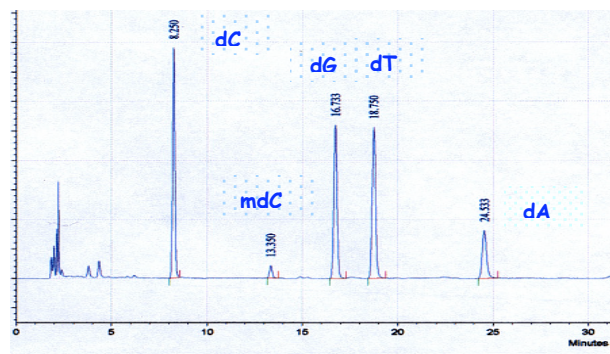


Fig. 6- Cromatogramma in cui sono mostrati i nucleosidi analizzati

Abbiamo verificato la qualità dell'idrolisi confrontando con DNA standard di timo di vitello le cui percentuali per singoli deossiribonucleosidi sono note. Abbiamo controllato la precisione del metodo iniettando aliquote di replicati per ogni singolo campione ed abbiamo determinato le percentuali dei singoli deossiribonucleosidi. Infine, abbiamo calcolato una media per la percentuale molare di ciascuno e così da poter eliminare i campioni con una deviazione standard superiore a 0.15% 5mC.

2.5 Analisi dei geni ortologi

Per determinare la frequenza di CpG nelle sequenze codificanti (cds) da pesci polari e temperati/tropicali, abbiamo preso le cds da pesci polari antartici

(sottordine *Notothenioidei* e da artici (ordine *Gadiformes*) dalla banca dati Genbank usando lo strumento WWW-QUERY dal sito internet <http://pbil.univ-lyon.fr>. Il set finale di dati conteneva 49 geni. Abbiamo quindi cercato gli orologi per i pesci tropicali o temperati usando lo strumento BLASTN (<http://www.ncbi.nlm.nih.gov/BLAST>). Abbiamo poi allineato le sequenze recuperate ciascuna con il corrispettivo orologio polare mediante l'applicazione Se-Align versione 2.0a11 Carbon (Rambaut, A., 1996), disponibile sul web alla URL <http://evolve.zoo.ox.ac.uk>. Abbiamo calcolato la frequenza totale di CpG su ciascuna sequenza mediante CODONW (J. Peden), nella versione online (<http://bioweb.pasteur.fr/seqanal/interfaces/codonw.html>).

2.7 Analisi delle sequenze di RNA ribosomale 18S in vertebrati

Per il confronto della frequenza di appaiamenti GC ed AU negli stem delle sequenze di RNA ribosomale 18S di vertebrati abbiamo utilizzato gli allineamenti in formato DCSE dal database dell'European Ribosomal RNA Database (<http://www.psb.ugent.be/rRNA/index.html>). Sulle sequenze allineate vengono dati i punti di inizio e fine degli stem e la loro numerazione, in modo tale da rendere possibile un conteggio manuale di tutti gli appaiamenti degli stem. Abbiamo quindi considerato le percentuali di coppie GC e AU sul totale degli appaiamenti nei singoli stem e poi le abbiamo confrontate con quelle degli stem nella sequenza dell'uomo.

3. Risultati

3.1.1 Analisi del livello di metilazione in pesci

Il nostro dataset per i pesci comprende 65 specie appartenenti a 12 su 38 ordini di *Teleostei* (classe *Actinopterygii*; Nelson, 1994). Nel nostro lavoro, quindi, rispetto a quello di Jabbari et al. (1997), l'insieme dei pesci è molto più numeroso e tutti i campioni sono stati analizzati con la stessa metodica. Dato che il nostro scopo era scoprire possibili differenze nel livello di metilazione in relazione alle differenze in temperatura corporea, abbiamo studiato le caratteristiche ecologiche e fisiologiche dei pesci analizzati, per suddividerli in base alla loro temperatura. Inoltre abbiamo scelto di analizzare un numero di specie grande abbastanza da ottenere un campione rappresentativo sia da un punto di vista tassonomico che ecologico. Le indicazioni climatiche, geografiche e tassonomiche insieme alla fonte di provenienza dei campioni analizzati sono presentate in Tabella 1a. Qui vi sono anche elencati i dati di GC e 5mC che abbiamo ottenuto nel presente lavoro con l'analisi HPLC sulle singole specie, i valori medi calcolati per generi o per famiglie (in un certo numero di casi), i rapporti 5mC/GC, e la lunghezza del genoma, laddove disponibile (fonte: Animal Genome Size Database).

La procedura seguita per mettere in grafico i dati è quella utilizzata anche nel confronto tra vertebrati ectotermi ed endotermi (Jabbari et al., 1997). I valori di 5mC per le specie sono stati messi in grafico contro i livelli di GC nella Fig. 7a (specie) e 7b (generi/famiglie). I nostri risultati per i pesci temperati/tropicali e quelli per i polari si allineano su due linee di regressione

Tabella 1a. Proprietà dei pesci analizzati (1)

Ordine	Famiglia	Specie	Clima ⁽²⁾	Valore-c ⁽³⁾	Specie		Famiglie/ Generi ⁽⁴⁾		
					GC	5mC	GC	5mC	
Osteoglossiformes	Pantodontidae	<i>Pantodon buchholzi</i>	TR	0.77	45.66	1.77			
	Notopteridae	<i>Notopterus notopterus</i>	TR		44.96	1.22			
Clupeiformes	Clupeidae	<i>Sardina pilchardus</i>	ST		47.12	1.40			
Cypriniformes	Cyprinidae	<i>Carassius auratus</i>	ST	1.60	39.53	1.44	39.36	1.40	
		<i>Brachydanio rerio</i>	TR	1.68	39.19	1.35			
Salmoniformes	Salmonidae	<i>Oncorhynchus keta</i>	TE	2.49	45.75	1.49			
Gadiformes	Merlucciidae	<i>Merluccius merluccius</i>	TE		48.69	2.18			
		<i>Merluccius bilinearis</i>		0.84					
	Gadidae	<i>Boreogadus saida</i>	PO	0.88	48.40	2.22			
		<i>Gadus morhua</i>	PO	0.40	48.61	2.37			
		<i>Arctogadus glacialis</i>	PO		48.13	2.74			
Cyprinodontiformes	Aplocheilidae	<i>Nothobranchius flammicomantis</i>	TR		47.26	1.59			
		<i>Aphyosemion elegans</i>	TR		43.49	1.52			
		<i>Aphyosemion coeruleum</i>		1.20					
		Rivulidae	<i>Aphyolebias peruensis</i>	TR		45.70	1.50		
		Fundulidae	<i>Fundulus heteroclitus</i>	TE	1.36	42.80	1.76		
	Poeciliidae	<i>Xiphophorus maculatus</i>	TR	0.76	41.28	1.39			
		Cyprinodontidae	<i>Jordanella floridae</i>	ST		41.43	1.10		
Zeiformes	Caproidae	<i>Aphanius fasciatus</i>	ST		43.17	1.62			
		<i>Capros aper</i>	ST		46.69	1.87			
Scorpaeniformes	Scorpaenidae	<i>Scorpaena guttata</i>	ST		41.27	1.48			
		<i>Scorpaena brasiliensis</i>		1.40					
		Hexagrammidae	<i>Ophiodon elongatus</i>	ST		44.31	1.78		
		Cottidae	<i>Leptocottus armatus</i>	TE		46.72	2.11		
Perciformes	Liparidae	<i>Liparis tunicatus</i>	PO	0.88	46.30	2.79			
		Serranidae	<i>Serranus cabrilla</i>	TE		44.39	1.66		
		<i>Alphesites immaculatus</i>	TR		43.40	1.40			
		<i>Paralabrax maculatofasciatus</i>	ST		43.63	1.36			
	Apogonidae	<i>Apogon imberbis</i>	ST		43.25	1.51			
		<i>Apogon bandanensis</i>		0.94					
	Carangidae	<i>Trachurus mediterraneus</i>	ST		44.58	1.65			
		<i>Trachurus declivis</i>		0.90					
	Sparidae	<i>Boops boops</i>	ST	0.52	44.65	1.83			
		Mullidae	<i>Mullus barbatus</i>	ST		48.86	2.19		
	Pomacanthidae	<i>Holacanthus passer</i>	TR		44.10	1.43			
		Kyphosidae	<i>Hermosilla azurea</i>	ST		44.26	1.41		
	Pomacentridae	<i>Microspathodon dorsalis</i>	TR		43.15	1.54			
		<i>Chromis</i> sp.		1.10	43.80	1.50	43.48	1.52	
	Labridae	<i>Coris julis</i>	TE	1.21	44.90	1.28			
		<i>Thalassoma grammaticum</i>	TR		42.59	1.29			
		<i>Halichoeres poeyi</i>	TR		44.43	1.41			
		<i>Halichoeres bathyphilus</i>		0.80					
	Scaridae	<i>Symphodus tinca</i>	ST	0.73	45.22	1.81			
		<i>Scarus ghobban</i>	TR		41.95	1.32			
		<i>Scarus guacamaia</i>		1.25					
		Zoarcidae	<i>Iluoetes fimbriatus</i>	TE	1.07	46.62	1.92		
	Bovichtidae	<i>Cottoperca gobio</i>	ST		43.65	1.98			
		<i>Bovichtus diacanthus</i>	ST		41.95	1.83			
	Eleginopidae	<i>Eleginops maclovinus</i>	TE		44.04	1.66			
		Nototheniidae	<i>Notothenia coriiceps</i>	PO		44.40	1.87		
		<i>Notothenia rossii</i>	PO		44.52	1.78	44.46	1.83	
		<i>Gobionotothen gibberifrons</i>	PO		42.62	2.15			
		<i>Gobionotothen marionensis</i>	PO		44.32	1.92	43.47	2.04	
		<i>Trematomus newnesi</i>	PO		44.57	1.82			
		<i>Trematomus bernacchii</i>	PO	1.19	43.59	1.82	44.08	1.82	
		<i>Dissostichus mawsoni</i>	PO	1.02	44.09	1.85			
		<i>Patagonotothen guntheri</i>	TE		44.08	1.69			
		<i>Lepidonotothen squamifrons</i>	TE		43.79	1.87			
		<i>Lepidonotothen kempfi</i>	PO		43.31	1.64			
		<i>Lepidonotothen nudifrons</i>	PO		43.74	1.79	43.61	1.77	
	Bathdraconidae	<i>Gymnodraco acuticeps</i>	PO		42.60	1.80			
		<i>Parachaenichthys charcoti</i>	PO		43.24	1.99	42.92	1.89	
		<i>Cygnodraco mawsoni</i>		1.41					
		Channichthyidae	<i>Chionodraco rastrispinosus</i>	PO		43.65	2.28		
		<i>Chionodraco hamatus</i>	PO	1.83	43.90	2.40	43.78	2.34	
		<i>Champscephalus esox</i>	TE		45.53	2.22			
		<i>Pseudochaenichthys georgianus</i>	PO		44.90	2.22			
		<i>Chaenocephalus aceratus</i>	PO		44.27	2.10			
		<i>Neopagetopsis ionah</i>	PO		43.80	2.08			
		Tripterygiidae	<i>Lepidonectes corallicola</i>	TR		44.52	1.42		
	Labrisomidae	<i>Dialommus fuscus</i>	TR		45.61	1.85			
		Pleuronectiformes	<i>Paralichthys californicus</i>	ST	0.80	46.50	1.58		
Tetraodontiformes	Tetraodontidae	<i>Tetraodon fluviatilis</i>	TR		48.39	1.93			
		<i>Tetraodon nigroviridis</i>		0.35					

Tab. 1a (1), (2) I dati di tassonomia e distribuzione climatica provengono da www.fishbase.org. (TR=tropicale, PO= polare, ST=subtropicale, TE=temperato). (3) valore-c da www.genomesize.com. (4) Le medie per GC e 5mC per famiglie/generi sono state calcolate solo laddove non ci fossero grosse differenze (linee verticali spesse), mentre i casi con le differenze sono indicati con linee verticali sottili.

parallele, la più alta delle quali riguarda i polari. Questi pesci hanno la temperatura più bassa e il livello medio di metilazione più alto tra tutti i pesci e tra tutti i vertebrati finora analizzati. La differenza tra i livelli di metilazione tra pesci polari e temperati/tropicali è evidente ed altamente significativa (t-test due campioni assumendo varianze diseguali; $p < 10^{-7}$).

Fig.7a

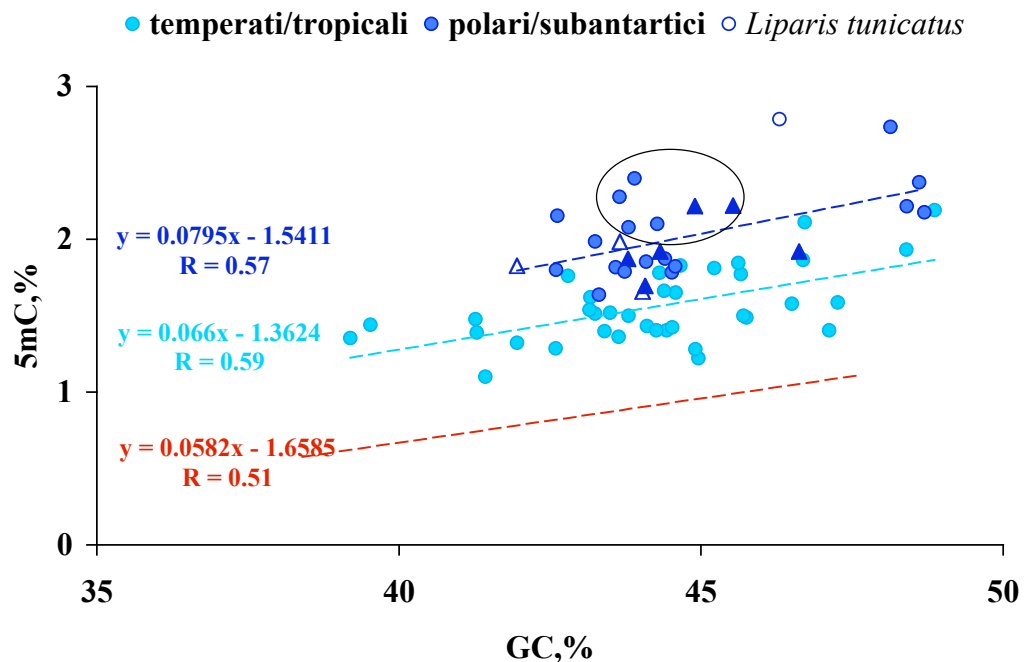


Fig. 7a- Andamento della metilazione in funzione del livello di GC in pesci polari (cerchi blu scuro: antartici e artici; triangoli pieni e triangoli vuoti: sub-Antartici di origine polare, e sub-Antartici nativi, rispettivamente) e temperati/tropicali (punti blu chiaro). I *Channichthyidae* sono racchiusi da un ovale. Il cerchio vuoto è *Liparis tunicatus*, che non è stato considerato nel calcolo

della retta di regressione. La retta dei mammiferi (rossa) è data come riferimento, inoltre vengono date le equazioni delle rette e i coefficienti di correlazione.

La retta dei pesci tropicali e temperati è comunque più alta rispetto alla retta dei mammiferi (che hanno una temperatura corporea più alta), mostrata come riferimento, ed è statisticamente simile (t-test, $p < 10^{-5}$) a quella riportata per il set di pesci da Jabbari et al. (1997). La differenza tra pesci di diverse temperature corrisponde a quella vista tra vertebrati ectotermi ed endotermi, ma è una situazione ancor più interessante perché è stata vista all'interno di un singolo gruppo filogenetico, i pesci.

Oltre al risultato principale ottenuto dal confronto tra pesci di temperatura diversa, uno sguardo più approfondito ai dati di Fig.7a rivela delle caratteristiche interessanti che possiamo esaminare singolarmente. *Liparis tunicatus*, il punto anomalo (outlier) col più alto livello di metilazione è un pesce artico che vive a una temperatura pari a quella degli altri artici, ma che si può trovare in acque di profondità fino a 3000 mt. Questo suscita l'ipotesi che anche fattori diversi dalla temperatura possano influenzare il livello di metilazione. Tra i pesci antartici, la famiglia dei Channictidi (pesci privi di emoglobina), pur avendo una temperatura uguale alle altre famiglie dei nototenioidi mostra i livelli di metilazione più alti (senza considerare *L.tunicatus*). Anche in questo caso c'è la possibilità che altri fattori influenzino i livelli di metilazione. In particolare il metabolismo, che è collegato alle mutazioni, allo stress ossidativo ed alla deaminazione, è particolarmente basso nei Channictidi, inferiore anche a quello degli altri nototenioidi (Hemmingsen, 1991). Le pendenze delle correlazioni di 5mC vs. GC erano

state già riportate da Jabbari et al. (1997) per l'insieme di dati allora analizzato. I dati quindi sono consistenti e indicano che abbiamo trovato una stessa dipendenza dal GC in tutte le classi di vertebrati analizzati. Questo suggerisce innanzitutto che la metilazione dipende dal livello di GC perché esso è correlato positivamente con la frequenza di CpG (Bernardi, 1985; Aïssani e Bernardi, 1991a,b). Inoltre suggerisce che i dati che spingono le pendenze oltre questo andamento “parallelo” possono essere considerati outliers.

Nella Fig. 7b sono rappresentati i risultati ottenuti dai dati delle famiglie rappresentate da una singola specie e dai livelli medi di generi e famiglie. Quest'ultima analisi evita, almeno in parte, errori associati alla presenza di un diverso numero di specie nelle varie famiglie o generi. Per le famiglie *Gadidae* e *Labridae*, le differenze in GC e/o 5mC tra le specie erano piuttosto rilevanti e non abbiamo calcolato una media né incluso i dati nella Figura.

Fig.7b

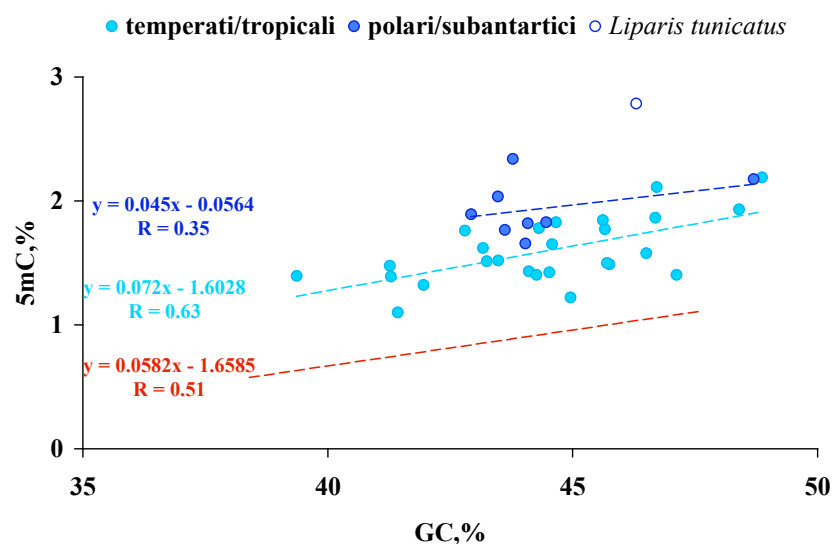


Fig. 7b- Andamento dei livelli medi di 5mC vs. GC per generi/famiglie di pesci polari (punti blu scuro) e temperati/tropicali (punti azzurri). Sono state incluse anche le famiglie di cui abbiamo analizzato una sola specie, mentre non

abbiamo considerato quelle famiglie le cui specie avevano 5mC o GC molto diversi tra loro. Il cerchio vuoto è *L. tunicatus*, che è stato considerato outlier. Vengono date la retta dei mammiferi (rossa) come riferimento, le equazioni delle rette e i coefficienti di correlazione.

Se i pesci tropicali vengono separati dai temperati (o subtropicali, come da denominazione utilizzata nella banca dati FishBase), si vede che due diverse linee possono essere tracciate per i due gruppi e la più bassa delle quali appartiene ai tropicali (Fig. 7c) che vivono a temperature più alte. La differenza tra le due linee, tuttavia, è statisticamente poco significativa ($p=0.06$).

Fig. 7c

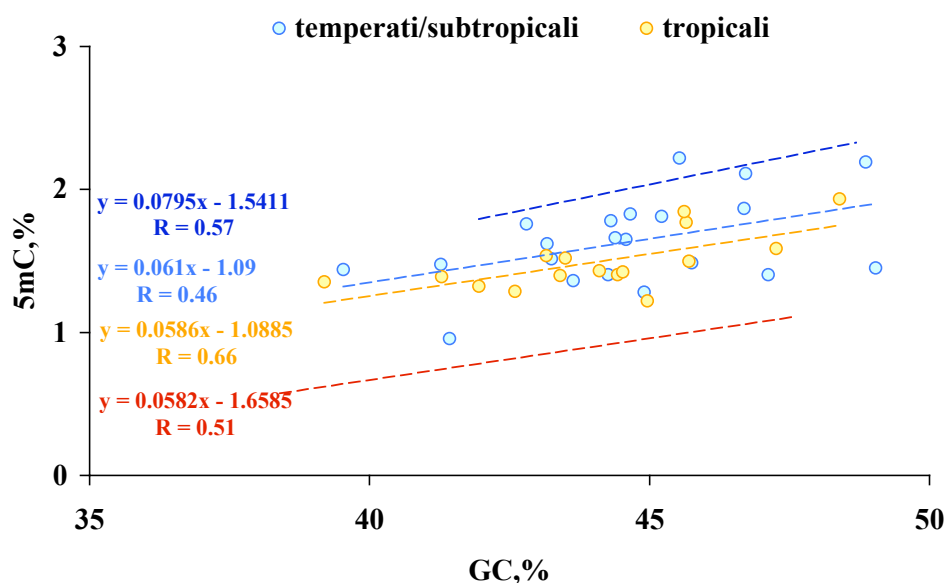


Fig. 7c- Andamento di 5mC vs. GC per i genomi di pesci temperati (punti in celeste) e tropicali (punti gialli). Sono date le equazioni ed i coefficienti di correlazione per i due gruppi, così come per i pesci polari (retta tratteggiata blu scura) e per i mammiferi (retta tratteggiata rossa) indicati qui come riferimento.

3.1.2 Analisi del livello di metilazione nei rettili

Il nostro insieme di dati comprende per i rettili 43 specie che fanno parte di tre dei quattro ordini esistenti: *Testudines* (tartarughe), *Squamata* (lucertole e serpenti) e *Crocodylia* (coccodrilli). Il quarto ordine, *Rhynchocephalia*, non studiato in questo lavoro, comprende solo la famiglia *Sphenodontidae* con due specie tuttora esistenti. La Tabella 1b presenta tassonomia, indicazione geografica, fonte di provenienza dei genomi analizzati, e i dati percentuali di GC e 5mC ottenuti dalle nostre analisi RP-HPLC per ciascuna specie. Ad eccezione di *Testudo graeca* non abbiamo utilizzato altri dati dalla letteratura perché questi erano stati ottenuti con tecniche diverse dall'HPLC o perché quelle specie sono state analizzate *de novo* nel corso del presente lavoro. I valori di 5mC e GC sono stati messi in grafico in Fig. 8a. Si nota subito che i rettili rappresentano il gruppo più eterogeneo tra tutti i vertebrati analizzati. Infatti, i livelli di 5mC nei rettili mostrano una dispersione che copre l'intero intervallo di dati dai mammiferi fino ai pesci temperati/tropicali. All'estremo inferiore troviamo coccodrilli e tartarughe che hanno una temperatura piuttosto alta (vedi Introduzione) e che pur essendo i più ricchi in GC sono anche i meno metilati. Gli squamati rappresentano invece l'ordine con la termofisiologia più eterogenea (vedi Introduzione) e difatti producono gran parte dell'eterogeneità riscontrata sul grafico. Il diverso comportamento della metilazione del DNA nei rettili confrontato alle altre classi di vertebrati è rimarcato dal fatto che in questo caso non viene ottenuta una correlazione significativa ($R=0.04$).

Tabella 1b. Proprietà dei rettili analizzati (1)

Ordine	Famiglia	Specie	Distribuzione ⁽²⁾	Valore-c ⁽³⁾	Specie				Famiglie/ Generi ⁽⁴⁾	
					GC	5mC	GC	5mC	GC	5mC
Testudines	Chelydridae	<i>Chelydra serpentina</i>	Canada, USA, C America	2.63	47.68	1.33	48.33	1.22		
		<i>Macrochelys temminckii</i>	USA		48.98	1.11				
	Emydidae	<i>Trachemys scripta elegans</i>	USA, C America, Thailandia, Singapore	2.62	46.98	1.14	47.21	1.05		
		<i>Chrysemis picta</i>	S Canada, USA		47.44	0.96				
	Testudinidae	<i>Testudo graeca</i>	Mediterraneo	3.52	45.70	0.77				
		<i>Testudo hermanni</i>	Oceani Atlantico, Indiano e Pacifico							
Cheloniidae	<i>Chelonia mydas</i>	Oceani Indiano e Atlantico, Mediterraneo e Mar Nero	2.64	47.38	1.00	47.06	0.98			
	<i>Caretta caretta</i>			46.74	0.96					
Squamata	Agamidae	<i>Chlamydosaurus kingii</i>	Australia, Nuova Guinea	2.89	44.67	0.93				
	Chamaeleonidae	<i>Furcifer oustaleti</i>	Madagascar		44.49	1.14				
	Iguanidae	<i>Iguana iguana</i>	USA, C/S America	2.65	44.33	1.36				
	Phrynosomatidae	<i>Sceloporus magister</i>	USA		45.84	0.85				
	Gekkonidae	<i>Tarentola mauritanica</i>	Mediterraneo, Uruguay, USA	2.84	46.59	0.94	46.32	1.01		
		<i>Gekko gecko</i>	Bangladesh, Indonesia, India, Thailand, USA		46.05	1.09				
	Gerrhosauridae	<i>Zonosaurus madagascariensis</i>	Madagascar, Seychelles	2.70	47.11	1.19				
	Teiidae	<i>Tupinambis teguixin</i>	S America		45.60	0.86				
	Lacertidae	<i>Podarcis sicula</i>	Mediterraneo	2.36	47.02	1.47	47.60	1.41		
		<i>Podarcis muralis</i>	C/S Europa, USA		48.18	1.34				
	Scincidae	<i>Mabuya gravenhorstii</i>	Madagascar	1.72	45.94	0.83				
		<i>Mabuya</i> sp.								
	Anguidae	<i>Anguis fragilis</i>	Europa, Asia Medio-orientale, N Africa	2.23	47.60	1.01				
	Boidae	<i>Acrantophis dumerili</i>	S Madagascar		44.41	0.94				
		<i>Boa constrictor</i>	C/S America	1.90	41.95	0.68				
		<i>Python molurus</i>	India, Malaysia, Sri Lanka		43.18	0.81				
		<i>Python reticulatus</i>	Indonesia, Malaysia, Filippine, Thailandia	1.60	42.38	0.65				
		Colubridae	<i>Hierophis viridiflavus</i>		C/S Europa	44.16				
	<i>Elaphe lineata</i>		S Italia	2.14	43.90	1.40				
	<i>Elaphe obsoleta</i>		C/E USA, S Canada		43.31	1.19				
<i>Elaphe guttata</i>	USA		41.83	1.29						
<i>Elaphe mandarina</i>	India, Laos, Vietnam, S China		43.05	1.23						
<i>Sibon longifrenis</i>	Panama, Costa Rica, Honduras		1.83	47.82	1.36					
<i>Natrix tessellata</i>	C/E Europa			44.12	1.03					
<i>Liopholidophis lateralis</i>	Madagascar		44.36	1.41						
<i>Geodipsas</i> sp	Madagascar		43.28	1.38						
<i>Madagascarophis</i> sp	Madagascar		42.71	1.10						
<i>Clelia rustica</i>	Brazil, Uruguay, Argentina		42.84	1.31						
Elapidae	<i>Walterinnesia aegyptia</i>		Asia Medio-orientale	2.25	42.77	0.80	42.61	0.92		
	<i>Notechis scutatus</i>	Australia	42.44		1.05					
Viperidae	<i>Bothrops jararaca</i>	Brasile, Argentina, N/E Paraguay	2.91	42.99	1.16	43.16	1.17			
	<i>Bothrops neuwiedi</i>	Brasile, Paraguay, Bolivia, Peru, Uruguay		43.33	1.17					
	<i>Viper aspis francisredi</i>	S Svizzera, N/C Italia, Slovenia, NO Croazia	2.07	43.40	1.20					
	<i>Daboia palestinae</i>	Pakistan, India, Sri Lanka, Bangladesh, Indonesia		43.94	1.17					
	<i>Daboia russelii</i>		1.30	42.53	0.92					
	<i>Echis coloratus</i>	N Africa, Asia Medio-orientale								
<i>Echis carinatus</i>										
Crocodylia	Crocodylidae	<i>Crocodylus niloticus</i>	N/C Africa	3.40	48.44	0.85	48.50	0.90		
	Alligatoridae	<i>Alligator mississippiensis</i>	S/E USA, Messico		48.56	0.96				

Tab. 1b. (1), (2) Tassonomia e distribuzione provengono da EMBL Reptile Database (www.embl-heidelberg.de/~uetz/LivingReptiles.html). (3) Il valore c è stato ricavato da Animal Genome Size Database (<http://www.genomesize.com>). (4) E' stata calcolata una media dei dati di GC e 5mC solo per le famiglie e i generi con differenze poco rilevanti.

Fig. 8a

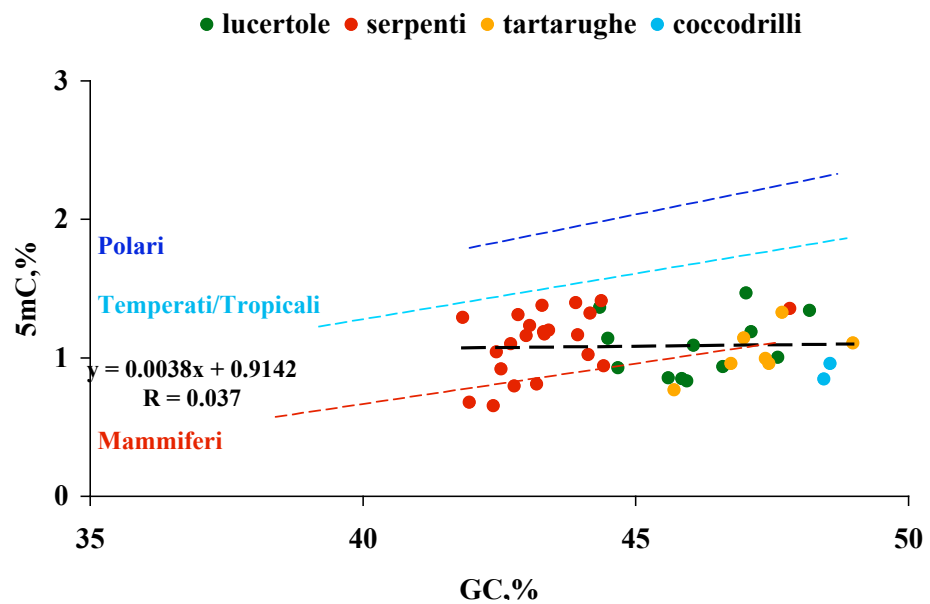


Fig. 8a- Andamento di 5mC in funzione di GC per i rettili. La retta in nero è stata tracciata considerando tutto il gruppo dei rettili e ne vengono riportate l'equazione ed il coefficiente di correlazione. Le rette dei pesci e dei mammiferi sono date come riferimento.

Abbiamo anche elaborato un grafico utilizzando le medie calcolate per generi e famiglie, laddove non ci fossero grandi differenze in 5mC o in GC (Fig. 8b). Anche in questo caso, così come nel caso del grafico per generi/famiglie dei pesci, il risultato resta lo stesso.

Fig.8b

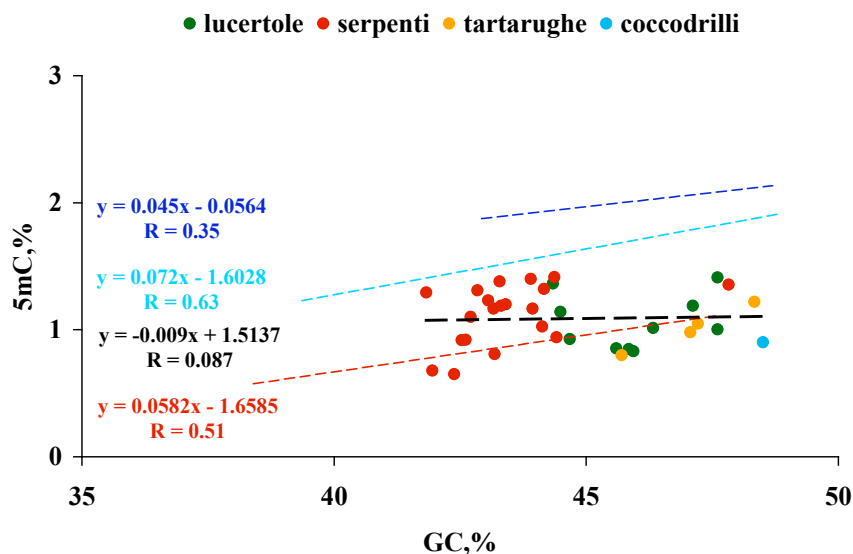


Fig. 8b- Andamento di 5mC in funzione di GC per generi/famiglie di rettili. Sono state incluse le famiglie di cui abbiamo analizzato una sola specie, mentre non abbiamo considerato quelle famiglie le cui specie avevano dati di 5mC o GC molto diversi tra loro. Le rette dei pesci e dei mammiferi sono date come riferimento.

3.1.3 Analisi del livello di metilazione nei mammiferi

Nella Tabella 1c abbiamo riportato la tassonomia, il GC, 5mC e i rapporti 5mC/GC per i genomi dei mammiferi analizzati. Nella Fig. 9 mostriamo i risultati dall'analisi RP-HPLC nel grafico 5mC vs. GC. Abbiamo suddiviso i dati dei mammiferi secondo l'ordine tassonomico in *Marsupialia* (marsupiali), *Eutheria* (euteri) e *Monotremata* (monotremi). Ciò ha permesso di evidenziare i punti fondamentali nella correlazione in quanto il primo gruppo ha GC e 5mC più bassi, mentre il terzo è quello con i livelli più alti. Il risultato importante è che quasi tutti i punti per i mammiferi analizzati sono sulla linea di regressione, e questo è chiaro se si pensa che i tre gruppi di mammiferi hanno una differenza di temperatura di soli 3°- 4° C circa.

Tabella 1c. Proprietà dei mammiferi analizzati

Ordine	Famiglia	Specie		GC	5mC
Monotemata	Tachyglossidae	Tachyglossus aculeatus	*	48.9	1.0
	Ornithorhynchidae	Ornithorhynchus anatinus	*	48.5	1.2
Didelphimorphia	Didelphidae	Monodelphis domestica	*	39.1	0.3
		Didelphis virginiana	*	39.2	0.3
Diprotodontia	Macropodidae	Macropus rufus	*	41.7	0.4
		Macropus robustus	*	41.2	0.4
	Vombatidae	Vombatus ursinus	*	40.9	0.3
Primates	Potoroidae	Potorous tridactyla	*	39.8	0.5
	Hominidae	Homo sapiens	*	42.8	0.7
	Tupaïidae	Tupaia montana		41.9	1.1
	Lemuridae	Hapalemur griseus		41.4	0.9
Insectivora	Erinaceidae	Erinaceus europeus		45.5	0.5
	Soriceidae	Crocidura russula		41.4	0.7
Dermoptera	Cynocephalidae	Cynocephalus variegatus		40.6	0.9
Megachiroptera	Pteropodidae	Pteropus hypomelanus	*	44.0	1.5
		Pteropus poliocephalus		41.5	1.4
		Pteropus sp	*	40.5	1.0
Microchiroptera	Noctilionidae	Noctilio albiventris	*	43.3	0.6
	Phyllostomidae	Ariteus planirostris	*	47.6	0.7
		Hipposideros galeritus		41.4	0.9
		Rhinolophus creaghi		41.5	0.8
		Myotis lucifugus		43.5	1.0
		Chiroderma salvinii		41.4	0.9
		Nycteris thebaica		42.9	0.9
Chaerephon pumila		41.4	0.9		
Pholidota	Manidae	Manis sp		42.4	0.6
Lagomorpha	Leporidae	Oryctolagus cuniculus		44.3	0.9
Rodentia	Sciuridae	Sciurus vulgaris		39.5	0.6
	Cricetidae	Cricetus norvegicus	*	40.7	0.4
	Spalacidae	Spalax sp.		38.4	0.6
	Caviidae	Cavia porcellus		39.7	0.7
	Muridae	Rattus norvegicus		43.9	0.9
	Procaviidae	Procavia capensis		41.0	0.7
	Balenopteridae	Balenoptera physalus		41.3	0.9
	Physiteridae	Physeter macrocephalus		41.9	1.1
	Phocoenidae	Phocena phocena		41.4	1.0
	Canidae	Canis familiaris		41.1	0.7
	Felidae	Panthera uncia		41.5	0.9
	Equidae	Equus caballus		42.8	1.0
	Cetartiodactyla	Suidae	Sus scrofa		44.6
Bovidae		Bos taurus	*	46.9	1.4

Tab. 1c. I dati di GC e 5mC contrassegnati da un asterisco sono ricavati dall'analisi con HPLC nel presente lavoro, mentre gli altri provengono da Jabbari et al. (1997).

Ciò vuol dire che, fatta eccezione per i punti anomali, l'unica causa delle differenze in livello di 5mC che vediamo in grafico è dato dal livello di GC. Un risultato simile al nostro era stato ottenuto da Jabbari et al. (1997), ma in più nel nostro lavoro abbiamo analizzato diverse specie di marsupiali rendendone note le caratteristiche di metilazione e GC.

Fig. 9

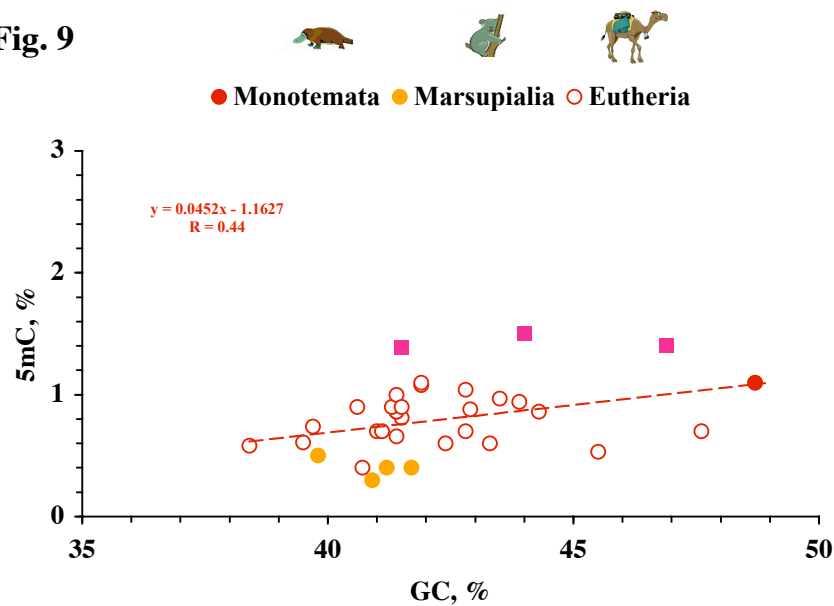


Fig. 9- Andamento di 5mC in funzione di GC per i mammiferi. I dati (vedi Tab. 1c) sono stati presi in parte dal presente lavoro ed in parte dalla letteratura (Jabbari et al., 1997). I quadrati rappresentano i tre outliers. Sono date l'equazione della retta ed il coefficiente di correlazione

Nel nostro gruppo di dati solo tre punti deviano dalla correlazione mostrata da tutte le altre specie analizzate. Questi sono *Bos taurus*, ed i pipistrelli (*Megachiroptera*) *Pteropus hypomelanus* e *Pteropus* sp. Per quanto riguarda *B.taurus*, è noto che il suo genoma contiene 25% di DNA satellite ricco in GC che è altamente metilato nei tessuti somatici. E'interessante notare che Jabbari et al. (1997) nel loro studio sui genomi di mammiferi avevano

trovato gli stessi outliers, cioè *B.taurus* e un altro pipistrello del genere *Pteropus*, *P. poliocephalus*. I pipistrelli potrebbero avere una situazione simile a *B.taurus*, ma attualmente non è ben chiaro. Un'indicazione viene data da Sabeur et al (1993) che analizzando la densità del DNA di *Pteropus* la trovarono particolarmente bassa nonostante la vasta distribuzione e la presenza di un satellite ricco in GC. Essi quindi ipotizzarono che l'abbassamento della densità potesse causato dalla metilazione (vedi par. 3.6).

3.2- Correlazione negativa di 5mC con la lunghezza del genoma

E' noto che 5mC è correlato positivamente con il livello di GC e che quest'ultimo mostra una correlazione negativa con la lunghezza del genoma (v. sopra; Bernardi e Bernardi, 1990) dato che accrescerla e quest'ultimo equivale ad aumentare il numero di regioni povere in GC. La correlazione negativa tra 5mC e lunghezza del genoma era stata vista per gli ectotermi e gli endotermi (Jabbari et al., 1997) ed è stata confermata nel presente studio per i pesci polari e temperati/tropicali. Qui sono state trovate due correlazioni, una per ciascun gruppo (Fig. 10a). Un grafico analogo è stato elaborato per i rettili. La correlazione trovata per i rettili quando si mette in grafico 5mC vs. la lunghezza del genoma (Fig. 10b) è evidentemente più bassa ($R=0.12$) di quella trovata per i pesci. Inoltre la pendenza è leggermente positiva mentre per i pesci è negativa. Questi risultati indicano che nei rettili la metilazione è così variabile che non è possibile individuare degli effetti dati dalla lunghezza del genoma.

Fig. 10a

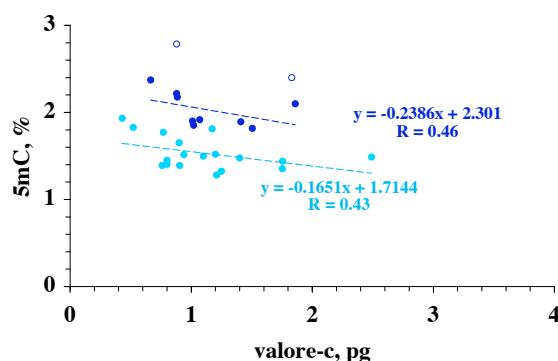


Fig. 10b

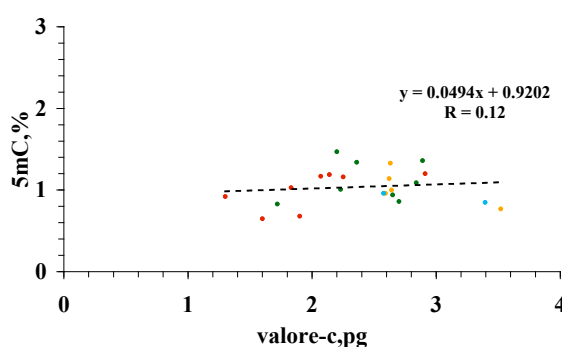


Fig. 10a- Livelli di 5mC vs. lunghezza del genoma in pesci polari (punti blu scuri pieni; i punti vuoti sono outliers) e temperati/tropicali (punti azzurri). **10b.** Livelli di 5mC vs. lunghezza del genoma in rettili. Per il riferimento ai colori vedi legende a Fig. 7. I dati per valore-c sono stati ricavati dal sito web Animal Genome Size Database (<http://www.genomesize.com/>).

3.3 Livello di CpG in geni ortologhi

Sappiamo che i livelli di CpG sono più alti nei genomi di pesci che in quelli di mammiferi (Jabbari et al., 1997) e che il CpG è correlato linearmente con i livelli di metilazione (Cacciò et al., 1997). Tuttavia non si sapeva se ci fossero differenze tra i livelli di CpG dei geni di pesci polari e temperati/tropicali. In questo lavoro, quindi, abbiamo confrontato il livello di CpG dei geni ortologhi dei due gruppi di pesci (Tab. 2). I risultati (Fig. 11)

mostrano che i livelli di CpG sono più elevati in geni dai pesci polari che negli ortologhi da pesci di regioni più calde. Questo risultato ben si relaziona a quanto trovato per i livelli di 5mC nei genomi di pesci e induce a ipotizzare che nei pesci polari anche i genomi, come i geni, possano essere più ricchi in CpG per il fatto di non essere andati incontro a perdita di 5mC.

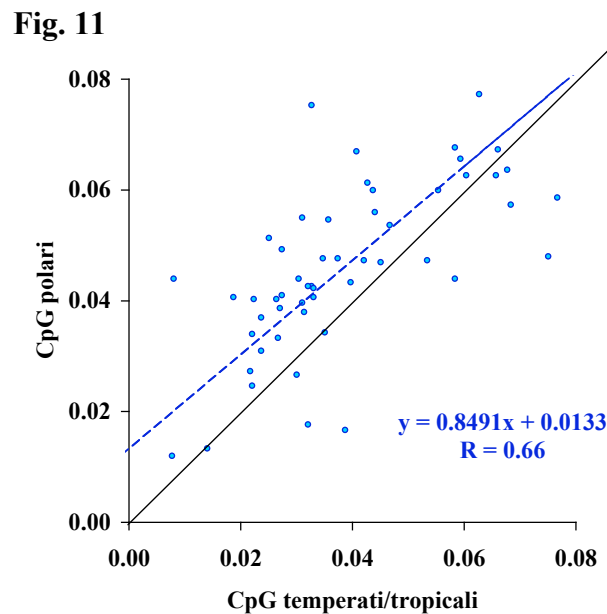


Fig. 11- Livelli di CpG dei geni di pesci polari vs. i rispettivi livelli in ortologhi di pesci temperati/tropicali.

Un altro confronto riguarda (Fig. 12) i livelli di GC, GC1, GC2, GC3 (il livello in GC alla prima, seconda e terza posizione) dei geni ortologhi dai due gruppi di pesci: mentre le differenze sono minime per i confronti tra GC e GC nelle prime due posizioni, essi sono notevoli per il GC3.

Tabella 2. Geni di pesci utilizzati per la determinazione della frequenza di CpG

Codice	Specie	CpG	GC1	GC2	GC3	GC	GC	GC3	GC2	GC1	CpG	Specie	Codice
DQ109810	<i>Gadus morhua</i>	6.7	57.9	42.7	87.7	62.8	57.3	74.6	42.7	54.5	4.1	<i>Salmo trutta</i>	AF247395
AF073925	<i>Gadus morhua</i>	6.0	50.3	39.6	94.9	61.6	49.7	62.4	41.9	44.8	5.5	<i>Danio rerio</i>	NM_131436
AB162779	<i>Chionodraco hamatus</i>	3.7	61.5	49.0	74.1	61.5	55.8	60.4	48.6	58.3	2.4	<i>Cottus kazika</i>	AB162777
AY291435	<i>Gadus morhua</i>	5.5	56.6	43.4	83.2	61.1	55.1	71.8	40.4	53.0	3.6	<i>Oncorhynchus mykiss</i>	AB047032
AF043269	<i>Coryphaenoides</i>	6.4	60.9	46.8	71.8	59.8	62.6	48.3	61.1	78.5	6.8	<i>Carassius auratus</i>	AF432155
X82223	<i>P.magellanica</i>	4.1	53.2	45.9	78.8	59.3	58.0	78.4	46.4	49.1	3.3	<i>Oncorhynchus keta</i>	AB091439
DQ270487	<i>Gadus morhua</i>	6.8	84.3	51.6	41.8	59.3	52.0	62.4	40.8	52.9	5.8	<i>Danio rerio</i>	BC051626
AJ609233	<i>Gadus morhua</i>	4.8	55.6	37.5	84.4	59.2	63.3	91.3	38.3	60.2	7.5	<i>Fundulus heteroclitus</i>	L23785
U71054	<i>Champscephalus gunnari</i>	4.0	61.2	51.5	64.1	58.9	53.5	61.9	38.1	60.5	2.6	<i>Thunnus thynnus</i>	AF291836
AY823272	<i>Notothenia coriiceps</i>	7.7	59.1	38.6	77.8	58.5	58.6	75.7	39.1	61.1	6.3	<i>Danio rerio</i>	BC059558
AB073379	<i>Theragra chalcogramma</i>	6.3	55.7	40.8	78.0	58.2	58.9	80.4	40.8	55.4	6.0	<i>Siniperca chuatsi</i>	AY395872
AY126448	<i>Gadus morhua</i>	6.1	47.5	41.3	85.5	58.1	52.3	76.0	41.9	39.1	4.3	<i>Danio rerio</i>	AY078161
DQ054840	<i>Gadus morhua</i>	4.7	57.5	36.5	80.3	58.1	54.3	71.9	35.3	55.7	4.5	<i>Sparus aurata</i>	AY055749
AJ532569	<i>Chionodraco hamatus</i>	4.2	57.5	38.7	77.3	57.8	56.2	72.4	38.7	57.5	3.3	<i>Danio rerio</i>	BC067566
AY161313	<i>Chaenocephalus aceratus</i>	5.4	57.4	36.8	78.9	57.7	57.9	80.5	35.5	57.6	4.7	<i>Oreochromis mossambicus</i>	AY034098
AF255553	<i>Notothenia coriiceps</i>	4.8	56.9	39.6	76.4	57.6	55.4	69.0	39.6	57.5	3.5	<i>Pleuronectes americanus</i>	X74492
AY046595	<i>Gadus morhua</i>	5.6	55.3	39.0	78.4	57.6	56.2	72.4	40.2	55.9	4.4	<i>Dicentrarchus labrax</i>	AY863148
CO542457	<i>Gadus morhua</i>	6.7	52.4	62.1	58.3	57.6	56.3	57.3	58.3	53.4	6.6	<i>Tetraodon nigroviridis</i>	CR1713588
AJ011585	<i>Trematomus bernacchii</i>	1.7	25.0	75.0	71.7	57.2	60.0	78.3	76.7	25.0	3.9	<i>Sparus aurata</i>	U58774
U58945	<i>Dissostichus mawsoni</i>	3.4	57.8	43.0	70.7	57.2	57.2	68.8	43.6	59.2	3.5	<i>Pleuronectes americanus</i>	AF012462
U09187	<i>Notothenia coriiceps</i>	6.6	56.5	34.0	81.0	57.1	59.0	87.8	36.1	53.1	5.9	<i>Oryzias latipes</i>	AB080120
AJ007878	<i>Chionodraco hamatus</i>	5.1	54.8	43.4	72.7	57.0	49.7	53.9	43.3	51.9	2.5	<i>Danio rerio</i>	AJ278268
Y08567	<i>Trematomus bernacchii</i>	6.3	53.7	42.6	74.6	57.0	59.9	80.3	41.5	57.7	6.6	<i>Takifugu rubripes</i>	AB179551
AB042054	<i>Theragra chalcogramma</i>	6.0	59.8	35.1	75.8	56.9	57.4	69.2	38.9	64.1	4.4	<i>Euthynnus pelamis</i>	AB042037
AY635584	<i>Gadus morhua</i>	5.7	58.5	34.5	77.3	56.8	59.0	42.2	56.8	77.9	6.8	<i>Tetraodon nigroviridis</i>	CR724573
AY927651	<i>Dissostichus mawsoni</i>	4.0	47.0	41.1	82.1	56.7	52.3	68.7	41.9	46.4	3.1	<i>Labidochromis chisumulae</i>	AY775098
U09186	<i>Notothenia coriiceps</i>	4.4	55.9	39.2	74.1	56.4	57.9	75.7	42.4	55.6	5.8	<i>Thunnus thynnus</i>	AB093569
CO541550	<i>Gadus morhua</i>	4.7	56.9	46.0	66.4	56.4	56.4	68.2	46.0	55.0	5.3	<i>Danio rerio</i>	BC075977
AJ535730	<i>Gadus morhua</i>	4.3	53.8	40.3	74.7	56.3	53.5	68.8	38.3	53.4	4.0	<i>Sparus aurata</i>	AJ277166
AY035584	<i>Gadus morhua</i>	5.5	61.5	41.3	65.1	56.0	54.3	64.8	38.9	59.3	3.1	<i>Salmo salar</i>	X97825
AF079822	<i>Notothenia coriiceps</i>	3.1	59.2	34.1	74.0	55.8	49.7	58.4	35.5	55.1	2.4	<i>Danio rerio</i>	BC067188
AF329903	<i>Gadus morhua</i>	5.9	49.1	46.1	71.9	55.7	62.8	81.1	52.0	55.2	7.7	<i>Oreochromis aureus</i>	AF270790
D10639	<i>Gadus morhua</i>	4.1	53.6	34.0	79.6	55.7	51.4	68.4	32.0	53.7	2.7	<i>Sparus aurata</i>	L49205
AJ550951	<i>Trematomus bernacchii</i>	4.8	50.8	43.4	71.2	55.1	55.3	72.2	42.1	51.6	3.7	<i>Pleuronectes americanus</i>	AF156787
AF255557	<i>Notothenia coriiceps</i>	3.4	57.2	40.4	67.6	55.1	53.5	62.1	40.1	58.3	2.2	<i>Oncorhynchus nerka</i>	AY026060
AY660881	<i>Gadus morhua</i>	4.3	46.9	34.7	83.7	55.1	49.2	68.4	33.7	45.6	3.3	<i>Sparus aurata</i>	AY660879
DQ011268	<i>Trematomus bernacchii</i>	4.0	54.1	39.8	70.9	54.9	51.0	60.0	39.4	53.6	2.2	<i>Tilapia mossambica</i>	AF109409
CO542193	<i>Gadus morhua</i>	7.5	52.8	32.4	79.6	54.9	51.6	69.0	32.4	53.5	3.3	<i>Oncorhynchus mykiss</i>	AY622862
AJ555405	<i>Gadus morhua</i>	4.4	25.3	70.9	67.1	54.4	48.1	55.7	62.0	26.6	0.8	<i>Dicentrarchus labrax</i>	DLA298290
AB086413	<i>Theragra chalcogramma</i>	3.9	62.7	27.1	73.2	54.3	52.0	67.3	27.1	61.6	2.7	<i>Pennahia argentata</i>	AB045645
AY141302	<i>Notothenia coriiceps</i>	2.7	44.2	63.7	53.0	53.6	54.4	53.7	64.5	44.9	3.0	<i>Girella punctata</i>	AB158262
AF326974	<i>Notothenia coriiceps</i>	4.9	55.5	39.6	64.1	53.1	49.2	53.1	40.0	54.5	2.7	<i>Danio rerio</i>	BC100000
AY103471	<i>Pagothenia borchgrevinkii</i>	3.3	50.6	39.7	68.9	53.1	48.4	56.5	39.5	49.2	2.7	<i>Danio rerio</i>	BC066715
DQ165081	<i>Gadus morhua</i>	4.3	49.3	37.3	72.5	53.1	53.2	70.2	35.9	53.4	3.2	<i>Tetraodon nigroviridis</i>	CR733511
AF094531	<i>Trematomus bernacchii</i>	2.7	48.0	54.4	55.0	52.4	49.9	50.9	50.9	48.0	2.2	<i>Takifugu rubripes</i>	AB217616
CO541708	<i>Gadus morhua</i>	3.8	57.4	30.2	69.2	52.3	53.9	75.3	27.6	58.8	3.1	<i>Salmo salar</i>	S77332
CO541984	<i>Gadus morhua</i>	4.1	62.6	31.8	62.6	52.3	44.9	54.2	29.9	50.5	1.9	<i>Ictalurus punctatus</i>	AF293651
AY736280	<i>Trematomus bernacchii</i>	2.5	60.3	43.7	49.0	51.0	52.5	53.6	43.7	60.3	2.2	<i>Oplegnathus fasciatus</i>	AY613390
CO541620	<i>Gadus morhua</i>	4.4	51.2	32.4	68.6	50.7	45.7	56.2	31.7	49.0	3.0	<i>Danio rerio</i>	BC053309
AY771352	<i>Dissostichus mawsoni</i>	1.8	46.8	41.9	61.8	50.2	52.7	67.9	42.0	48.3	3.2	<i>Pseudopleuronectes americanus</i>	AY631037
AY454308	<i>Notothenia coriiceps</i>	1.2	54.8	37.3	56.5	49.5	46.2	52.6	35.0	50.9	0.8	<i>Danio rerio</i>	AY454307
AJ243767	<i>Notothenia coriiceps</i>	1.3	56.6	30.9	59.6	49.0	49.4	60.6	30.7	57.1	1.4	<i>Pennahia argentata</i>	AB039672
U92447	<i>Pseudochaenichthys charcoti</i>	4.7	47.8	33.6	62.7	48.0	49.8	66.4	36.6	46.3	4.2	<i>Danio rerio</i>	AY628221

Tab.2. I geni dei pesci polari (a sinistra) sono stati elencati per GC decrescente. Sulla destra ci sono i corrispondenti geni ortologhi dai pesci temperati/tropicali.

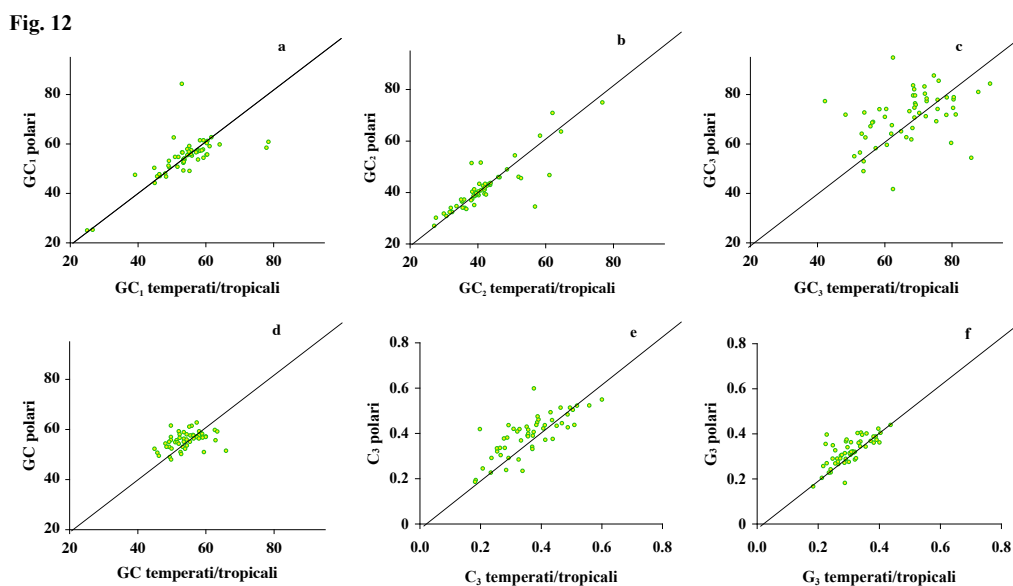


Fig. 12- Livelli di GC1, GC2, GC3, GC, C3 e G3 dai geni di pesci polari contro i rispettivi livelli per geni ortologhi da pesci temperati/tropicali elencati nella Tab. 3.

3.4 Correlazione del livello di metilazione con tassonomia e filogenia

In generale, abbiamo visto che pesci appartenenti allo stesso genere e di simile temperatura corporea, come molti generi della famiglia *Nototheniidae* e le specie del genere *Chionodraco*, mostrano livelli simili di GC e 5mC. Questo accade anche in specie di generi diversi, ma appartenenti a stesse famiglie antartiche e a specie da generi diversi delle famiglie *Cyprinidae*, *Bathydraconidae*, *Pomacentridae*. Questa relazione con la filogenesi a condizioni simili di temperatura viene a cadere quando ci si trova con temperature corporee diverse. Ciò si applica a casi nelle famiglie *Gobiidae*, *Scombridae*, *Gadidae*, *Cichlidae* (risultati non mostrati).

3.5 Correlazione tra livello di metilazione e temperatura corporea

La Fig. 13 mostra un grafico della metilazione contro la temperatura corporea, Tb. I campioni sono stati suddivisi in tre gruppi a temperature

diverse: polari, temperati e tropicali. I punti sono distribuiti attorno a livelli medi decrescenti in correlazione inversa rispetto alla temperatura e l'errore standard è minimo.

Fig. 13

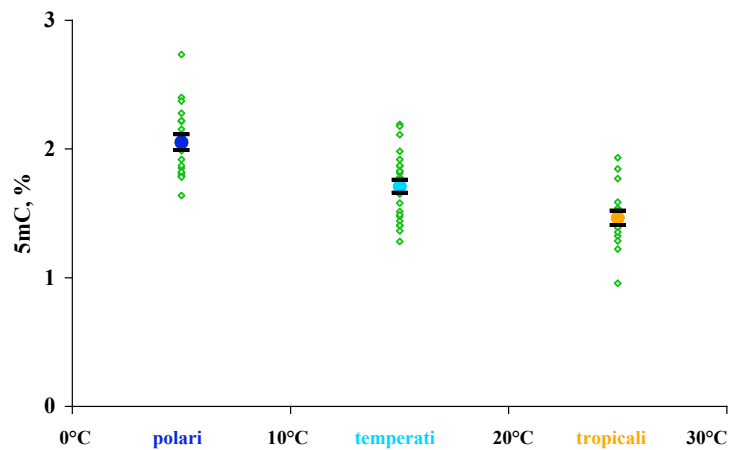


Fig. 13- Correlazione tra livelli di metilazione e temperatura corporea. I due punti estremi corrispondono ai genomi di pesci polari e tropicali, rispettivamente, mentre nel mezzo si trova la temperatura corporea media per i pesci temperati. Le linee nere riportate rappresentano l'errore standard.

3.6 Correlazione tra livello di GC derivante dall'analisi dei nucleosidi e dalla centrifugazione in gradiente di densità di cloruro di cesio (CsCl)

In Fig. 14a (pesci) e 14b (rettili) mostrano che se i livelli di GC sono calcolati a partire dalla densità modale di galleggiamento del DNA, si ricavano valori inferiori di circa 2-3% in GC rispetto a quelli derivanti dall'analisi dei nucleosidi.

Fig. 14a

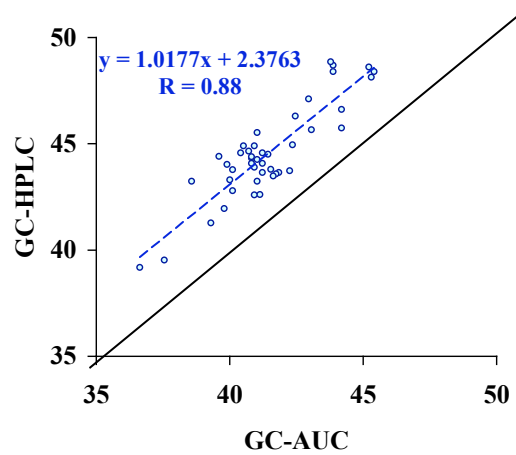


Fig. 14b

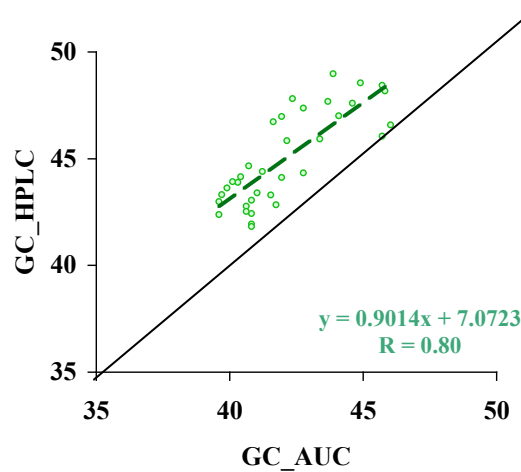


Fig. 14a- Correlazione tra i livelli genomici di GC ottenuti mediante HPLC e mediante ultracentrifugazione analitica in pesci. 14b. La correlazione è qui vista per i rettili.

I dati utilizzati per i grafici sono su Varriale e Bernardi (2006a,b).

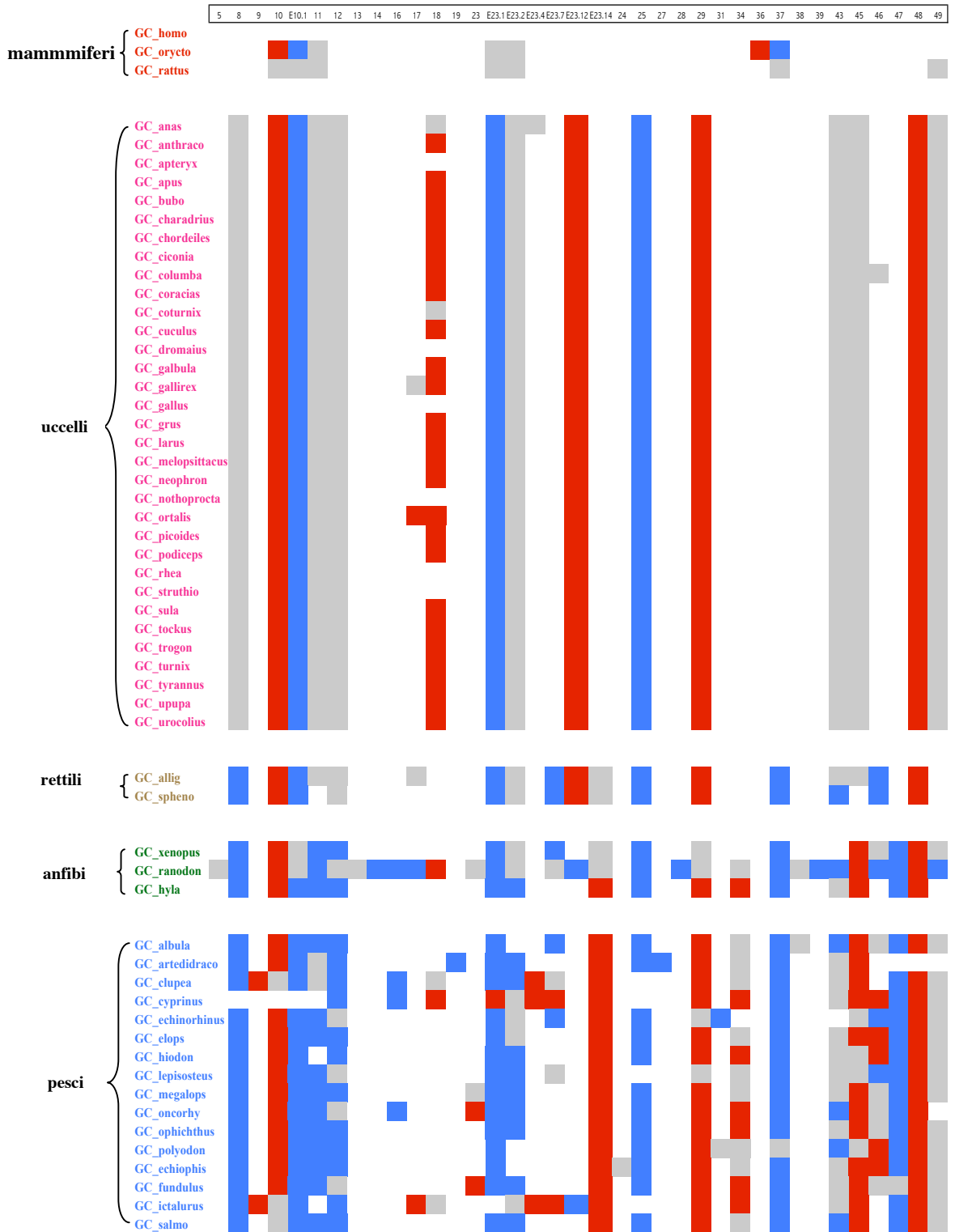
Tale differenza è dovuta, almeno in parte, alla metilazione del DNA. Kirk (1967) mostrò infatti che l'effetto della metilazione del DNA sulla sua densità di galleggiamento è di abbassarla. Ciò comunque non rappresenta una spiegazione completa della differenza in GC perché la differenza stimata di circa il 3% verrebbe ad essere dovuta a circa il 4% di metilazione, mentre

questo livello è circa il 2% nei pesci polari ed è ancor più basso nei pesci non polari e nei mammiferi. Un contributo ulteriore potrebbe provenire dalle frequenze dei di- e tri-nucleotidi, ma la causa potrebbe essere anche un errore sistematico nel calcolo del GC dalla densità di galleggiamento in CsCl. Questo calcolo è basato sulla correlazione tra densità e GC in DNA batterici, che generalmente sono metilati molto poco o per nulla.

3.7 Analisi dell'RNA ribosomale 18S endotermi ed ectotermi

Nel confronto delle percentuali di GC nelle regioni di struttura secondaria degli RNA ribosomali 18S abbiamo considerato solo gli stem la cui composizione variasse tra un vertebrato e l'altro. Tutti i casi in cui la sequenza era invariata o non dava luogo a differenza di composizione non sono stati esaminati. I casi analizzati sono elencati in Tab. 3 dove abbiamo riportato le percentuali di GC per gli stem in un certo numero di vertebrati. Il blu indica un GC più basso nei vertebrati a sangue freddo mentre il rosso indicano un GC più alto. Grigio e bianco indicano rispettivamente una forte similitudine ($\pm 5\%$ GC) e un'identità. Come si può vedere, negli endotermi circa il doppio degli stem hanno un GC più alto rispetto agli ectotermi indicando che l'aumento del livello di GC può avere una funzione nella stabilizzazione della struttura alle temperature elevate.

Tabella 3. Confronto dei livelli di GC degli stem (in colonna) dell'RNA ribosomale 18S tra classi di vertebrati.



Tab. 3. Ciascuna colonna rappresenta uno stem, mentre ciascuna riga rappresenta un vertebrato.

4. Discussione

I primi studi che furono svolti sui livelli globali di metilazione nel DNA da vertebrati (Vanyushin et al., 1970, 1973; Pollock et al., 1978; Ehrlich et al., 1982; Gama-Sosa et al., 1983a,b; Serrano et al., 1993) non portarono ad alcuna conclusione circa le differenze trovate. Un primo passo nella comprensione del significato della metilazione globale era stato fatto da Jabbari et al. (1997). Essi fornirono la prima dimostrazione del fatto che il livello medio di 5mC e di CpG nel DNA dei pesci e degli anfibi è circa il doppio di quello osservato in mammiferi ed uccelli. Varie critiche all'ipotesi della relazione temperatura-livello di metilazione potrebbero essere mosse, in quanto si potrebbe pensare ad altri fattori in grado di influenzare i livelli di metilazione (vedi Jabbari et al., 1997). Uno di questi è rappresentato dalla quantità e livello di metilazione delle sequenze ripetute, sia satelliti che intersperse. Non conosciamo l'entità di queste variabili per tutti i genomi finora analizzati, ma sappiamo che un confronto delle frazioni composizionali da uomo, topo, gallo e *Xenopus* mostra che quest'ultimo possiede un livello intragenomico di 5mC più alto di quello dei tre vertebrati a sangue caldo (Cacciò et al., 1997). Inoltre sappiamo che le sequenze ripetute sono variabili da una famiglia, genere ed a volte anche da una specie all'altra; è più probabile quindi che esse contribuiscano alla dispersione dei punti piuttosto che ad un errore sistematico. Un secondo fattore che potrebbe avere influenza sul livello di 5mC è la lunghezza del genoma, ma come abbiamo mostrato nei Risultati, essa non ha un effetto. Inoltre è stato mostrato che Gadiformi e Tetraodontiformi hanno un livello di metilazione

uguale o superiore ad altri pesci con lunghezza del genoma e numero di sequenze ripetute superiori.

La connessione che c'è tra 5mC e temperatura e l'ipotesi dell'esistenza dei due equilibri di livelli di 5mC, uno per pesci/anfibi e l'altro per mammiferi/uccelli (Jabbari et al., 1997) grazie al presente lavoro sono state ulteriormente dimostrate grazie al confronto tra pesci polari, temperati e tropicali. Per quanto riguarda l'analisi dei pesci, la situazione appare molto chiara in quanto tutti i pesci che hanno una temperatura corporea bassa hanno anche un livello di metilazione alto. Questo è vero per i due gruppi di pesci polari *strictu sensu* ma anche per i pesci sub-Antartici e i temperati freddi, come *Gadus morhua* e *Merluccius merluccius*. Tutti questi infatti vivono in un intervallo di temperatura compreso tra 0° e 10° C, inferiore quindi rispetto alle temperature a cui vivono gli altri pesci qui analizzati. Il caso dei Channictidi si presenta particolarmente interessante e meriterà certamente ulteriori approfondimenti.

Per quanto riguarda i rettili, la loro eterogeneità di temperatura corporea fa sì che la situazione non sia altrettanto chiara, ma spiega comunque la dispersione dei punti che troviamo e che riflette quindi questa variabilità eco-fisiologica. Ciò che noi troviamo, inoltre, è in accordo con altre evidenze sperimentali sulle caratteristiche dei genomi dei rettili. Queste sono la composizione nucleotidica, l'asimmetria delle bande in CsCl, l'eterogeneità compositiva (Aïssani e Bernardi, 1991a,b; Hughes et al., 2002; Bernardi, 2005) e la velocità dell'evoluzione del cariotipo (Olmo, 2005). I risultati qui ottenuti forniscono una prova di tale diversità. I rettili studiati precedentemente (Jabbari et al., 1997) comprendevano un cocodrillo, due tartarughe, un varano

e tre serpenti. Questo campione era piccolo, eterogeneo (in quanto comprendeva tre vecchi dati; Vanyushin, 1970, 1973) e non rappresentativo da un punto di vista filogenetico dato che le due tartarughe e due dei serpenti appartenevano allo stesso genere. Questi rettili mostravano livelli di metilazione del DNA che coprivano un intervallo ampio con la maggior parte dei valori più vicini a quelli dei vertebrati a sangue caldo che a quelli dei vertebrati a sangue freddo. I risultati disponibili a quel tempo suggerivano quindi che i rettili potevano essere caratterizzati da un livello di metilazione più basso rispetto ad anfibi e pesci. Tali scoperte potevano far ipotizzare una transizione nel livello di metilazione già durante la comparsa di un antenato comune a mammiferi e rettili, ossia alla comparsa degli amnioti.

Se le transizioni composizionali avute nei genomi dei vertebrati alla comparsa dei moderni mammiferi ed uccelli fossero avvenute in un antenato comune ai tetrapodi, come proposto da Hughes et al. (1999), o in un antenato comune ai rettili e mammiferi, come proposto da Duret et al. (2002), avremmo dovuto trovare tra tutti i rettili dei livelli di 5mC molto simili, cosa che non si verifica affatto. Noi abbiamo trovato che, mentre la metilazione è bassa in coccodrilli e tartarughe (che hanno anche un genoma molto eterogeneo; Aïssani e Bernardi, 1991a,b), la cui temperatura è relativamente alta, le lucertole e i serpenti hanno livelli di 5mC che coprono tutto l'intervallo dai pesci tropicali ai mammiferi.

Per quanto riguarda l'RNA, i nostri risultati differiscono da quanto trovato da Wang et al. (2006), pur avendo utilizzato i medesimi dati di partenza. La ragione di tale discrepanza, sta molto probabilmente nel diverso trattamento dei dati: nel loro lavoro essi hanno calcolato solo la media di tutti gli stem di tutte le specie in ogni singola classe. In tal modo hanno inserito nel

calcolo anche molti stem la cui sequenza era invariante tra i diversi gruppi di vertebrati o addirittura in cui erano presenti nucleotidi non determinati (indicati come N o S). Questo tipo di computo non fornisce un confronto diretto e produce degli artifici di calcolo. Nel nostro lavoro, invece, effettuando un confronto manuale su tutti gli stem ad uno ad uno e separatamente per ciascuna specie abbiamo potuto scartare a priori tutti gli stem non informativi. Inoltre abbiamo potuto effettuare un ulteriore controllo mediante mezzi bioinformatici. A differenza di quanto affermato da Wang et al. (2006), quindi, abbiamo trovato una correlazione piuttosto netta tra livello di GC e temperatura corporea ed abbiamo riscontrato per gli uccelli una situazione più simile a quella dei mammiferi che non a quella degli anfibi.

Conclusione

Mediante questo lavoro abbiamo fornito evidenze chiare del fatto che la temperatura è correlata con il livello di metilazione nei genomi e con la composizione delle regioni strutturate nell'RNA ribosomale. Possiamo così affermare che i livelli di 5mC e CpG integrano gli effetti di temperatura e tempo di esposizione ad essa. Ciò vuol dire che i cambiamenti più complessi quali quelli nella struttura del genoma, e nella formazione di isocore (lunghe regioni a composizione omogenea) ricche in GC, potrebbero essere correlati meglio con la metilazione del DNA che con la temperatura corporea. Nel caso di confronti tra genomi di endotermi ed ectotermi quali pesci ed anfibi, la temperatura è un buon parametro perché i tempi d'esposizione sono sufficientemente lunghi. Per i rettili, invece, non essendo noti i tempi d'esposizione, possiamo correlare l'eterogeneità composizionale del DNA e la formazione di isocore ricche in GC con la metilazione e i livelli di CpG.

4. RINGRAZIAMENTI

Ringrazio in maniera particolare il direttore della ricerca, professor Giorgio Bernardi, per la sua fondamentale assistenza durante tutto il periodo di svolgimento del programma di dottorato e per aver incoraggiato la comunicazione dei risultati del lavoro presso meeting nazionali (FISV 2003) ed internazionali (6th Anton Dohrn Workshop, Ischia 2005; Marine Genomics, Sorrento 2006).

Ringrazio con pari stima il relatore della tesi, professor Giuseppe Geraci.

Ringrazio tutti i collaboratori elencati in tabella che hanno fornito materiale biologico per la ricerca, e in particolare i dottori Ennio Cocca, Guido Di Prisco e Cinzia Verde per aver fornito anche informazioni bibliografiche riguardanti i pesci polari.

Ringrazio ancora i professori Ettore Olmo e Giacomo Bernardi in quanto editori delle pubblicazioni di questo lavoro sulla rivista Gene, i dottori Kamel Jabbari e Oliver Clay, rispettivamente per la revisione dei manoscritti e per la collaborazione all'analisi statistica dei dati.

Ringrazio infine il dottor Antimo D'Aniello per la collaborazione nella messa a punto della metodica HPLC, la dottoressa Sandra Hochscheid per le indicazioni di letteratura sulla termoregolazione nei rettili, il dottor Gabriele Amore per la lettura della tesi, e tutti gli altri membri del Laboratorio di Evoluzione Molecolare.

5. Bibliografia

- 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.
- Armbruster, G.F.J., 2001. Temperature-based variation of rRNA secondary structure models: a case study in the insect *Drosophila simulans*, the land snail *Isabellaria adriani*, and the crustacean *Daphnia pulex*. *Can. J. Zool.* 79, 334–345.
- Avery, R.A., 1982. Field studies of body temperatures and thermoregulation. In: Gans, C., Pough, F.H. (Eds), *Biology of the reptilia*, vol. 12 C. Academic Press, pp. 93-166.
- Bargelloni, L., Marcato, S., Zane, L., Patarnello, T., 2000. Mitochondrial phylogeny of notothenioids: a molecular approach to antarctic fish evolution and biogeography. *Syst. Biol.* 49:114–129.
- Bernardi, G., 1985. The organization of the vertebrate genome and the problem of the CpG shortage. In: Cantoni, G.L. and Razin, A. (eds.). *Biochemistry and biology of DNA methylation*. Alan Liss, N.Y., U.S.A. Pp. 3-10.
- Bernardi, G., 2005 (reprinted from 2004). *Structural and Evolutionary Genomics. Natural Selection in Genome Evolution*. Elsevier, Amsterdam.
- Bernardi, G., Bernardi G., 1986. Compositional constrains and genome evolution. *J. Mol. Evol.* 24, 1-11.

- Bernardi, G., Bernardi, G., 1990. Compositional patterns in the nuclear genome of cold-blooded vertebrates. *J. Mol. Evol.* 31, 265-281.
- Bestor, T.H., 2000. The DNA methyltransferases of mammals. *Hum Mol Genet.* 9, 2395-2402.
- Bird, A.P., 1980. DNA methylation and the frequency of CpG in animal DNA. *Nucleic Acids Res.* 8, 1499-1503.
- Brown, C.T., Jiricny, J., 1987. A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. *Cell* 50, 945-950.
- 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. USA.* 94, 3817-3822.
- Costello, J.F., Plass C., 2001. Methylation matters. *J. Med. Genet.* 38, 285-303.
- Coulondre, C., Miller, J.H., Farabaugh, P.J., Gilbert, W., 1978. Molecular basis of base substitution hotspots in *Escherichia coli*. *Nature* 274, 775-780.
- 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. Marine Syst.* 27, 253-265.

- Doskocil, J. and Sorm, F. (1962) Distribution of 5-methylcytosine in pyrimidine sequences of deoxyribonucleic acids. *Biochim. Biophys. Acta* 55, 95.
- Duret, L., Semon, M., Piganeau, G., Mouchiroud, D., Galtier, N., 2002. Vanishing GC-rich isochores in mammalian genomes. *Genetics* 162, 1837-1847.
- Eastman, J.T., 2005. The nature of the diversity of Antarctic fishes. *Polar Biol.* 28, 93-107.
- 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., Wang, R.Y., 1981. 5-Methylcytosine in eukaryotic DNA. *Science* 212, 1350-1357.
- Ehrlich, M., Gama-Sosa, M.A., Huang, L.H., Midgett, R.M., Kuo, K.C., McCune, R.A., Gehrke, C., 1982. Amount and distribution of 5-methylcytosine in human DNA from different types of tissues of cells. *Nucleic Acids Res.* 10, 2709-2721.
- Ehrlich, M., Norris, K.F., Wang, R.Y.H., Kuo, K.C., Gehrke, C.W., 1986. DNA cytosine methylation and heat-induced deamination. *Biosci. Rep.* 6, 387-393.
- Galtier, N., Lobry, J.R., 1997. Relationships between genomic G+C content, RNA secondary structures, and optimal growth temperature in prokaryotes. *J Mol Evol.* 44, 632-636.
- Gama-Sosa, M.A., Midgett, R.M., Slagel, V.A., Githens, S., Kuo, K.C., Gehrke, C.W., Ehrlich, M., 1983. 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., 1983. 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. and Kuo, K.C., 1984. Quantitative reversed-phase high-performance liquid chromatography of major and modified nucleosides in DNA. *J. Chromatogr.* 301, 199-219.
- Grippo, P., Iaccarino, M., Parisi, E. Scarano, E., 1968. Methylation of DNA in developing sea urchin embryos. *J. Mol. Biol.* 36, 195-208.
- Hemmingsen, E. A., 1991. Respiratory and cardiovascular adaptation in hemoglobin free fish: resolved and unresolved problems. In: di Prisco, G., Maresca, B. and Tota, B. (eds.). *Biology of Antarctic Fish*. New York: Springer-Verlag. Pp. 191-203.
- Hughes, S., Zelus, D., Mouchiroud, D., 1999. Warm-blooded isochore structure in Nile crocodile and turtle, *Mol. Biol. Evol.* 16, 1521-1527.
- Hughes, S., Clay, O., Bernardi, G., 2002. Compositional patterns in reptilian genomes. *Gene* 295, 323-329.
- 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.
- Jabbari, K., Bernardi G., 2004. Cytosine methylation and CpG, TpG (CpA) and TpA frequencies. *Gene* 333, 143-149.
- Josse, J., Kaiser, A.D., Kornberg, A., 1961. Enzymatic synthesis of deoxyribonucleic acid. VIII. Frequencies of nearest neighbor base sequences in deoxyribonucleic acid. *J Biol Chem.* 236, 864-875.

- Kennett, J.P., 1982. Marine geology. Prentice-Hall, Englewood Cliffs, New Jersey.
 - Kirk, J.T., 1967. Effect of methylation of cytosine residues on the buoyant density of DNA in caesium chloride solution. *J. Mol. Biol.* 28, 171-2.
 - Kramer, B., Kramer, W., Fritz H.J., 1984. Different base/base mismatches are corrected with different efficiencies by the methyl-directed DNA mismatch-repair system of *E. coli*. *Cell* 38, 879-887.
1. Morgan, H.D., Dean, W., Coker, H.A., Reik, W. and Petersen-Mahrt, S.K. (2004) Activation-induced cytidine deaminase deaminates 5-methylcytosine in DNA and is expressed in pluripotent tissues: implications for epigenetic reprogramming. *J. Biol. Chem.* 279, 52353-52360
- Mrosovsky, N., Pritchard, P.C.H., 1971. Body temperatures of *Dermochelys coriacea* and other sea turtles. *Copeia* 4, 624-631.
 - Musto, H., Naya, H., Zavala A., Romero, H., Alvarez-Valín F. and Bernardi G. (2004) Correlations between genomic GC levels and optimal growth temperatures in prokaryotes. *FEBS Lett.* 573, 73-77.
 - Nelson, J.S., 1994. Fishes of the world, 2nd Edition. John Wiley & Sons Inc., New York.
 - Olmo, E., 2005. Rate of chromosome changes and speciation in reptiles. *Genetica* 125, 185-203.
 - Pollock, J.M. Jr, 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-AI: Sequence Alignment Editor. Available at <http://evolve.zoo.ox.ac.uk/>.

- Rusmintratip, V. and Sowers, L.C., 2000. An unexpectedly high excision capacity for mispaired 5-hydroxymethyluracil in human cell extracts. *Proc. Natl. Acad. Sci. USA* 97, 14183-14187.
- Sabeur, G., Macaya, G., Kadi, F., Bernardi, G., 1993. The isochore patterns of mammalian genomes and their phylogenetic implications. *J Mol Evol.* 37, 93-108.
- Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. *Molecular Cloning, A Laboratory Manual*, 2nd ed., Cold Spring Harbor (N.Y.) Laboratory Press.
- Salser, W., 1977. Globin mRNA sequences: analysis of base pairing and evolutionary implications. *Cold Spring Harbor Symp. Quant. Biol.* 40, 985-1002.
- Seebacher, F., Grigg, G.C., Beard, L.A., 1999. Crocodiles as dinosaurs: behavioral thermoregulation in very large ectotherms leads to high and stable body temperatures. *J. Exp. Biol.* 202, 77-86.
- 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.
- Stankovic, A., Spalik, K., Kamler, E., Borsuk, P., Weglenski, P., 2002. Recent origin of sub-Antarctic notothenioids. *Polar Biol.* 25, 203–205.

- Summers, A.P., 2005. Warm-hearted crocs. *Nature* 434, 833-834.
- Sved, J., Bird, A., 1990. The expected equilibrium of the CpG dinucleotide in vertebrate genomes under a mutational model. *Proc. Natl. Acad. Sci. USA* 87, 4692-4696.
- Swartz, M.N., Trautner, T.A, Kornberg, A., 1962. Enzymatic synthesis of deoxyribonucleic acid. XI. Further studies on nearest neighbor base sequences in deoxyribonucleic acids. *J Biol Chem.* 237, 1961-1967.
- Thiery, J.P., Macaya, G., Bernardi, G., 1976. An analysis of eukaryotic genomes by density gradient centrifugation. *J. Mol. Biol.* 108, 219-35.
- 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 and body temperature in fishes. *Gene*. In press.
- Varriale, A., Bernardi, G., 2006. DNA methylation in reptiles. *Gene*. In press.
- Wang, R.Y.H., Kuo, K.C., Gehrke, C.W., Huang, L.H. and Ehrlich, M., 1982. Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. *Biochim. Biophys. Acta* 697, 371-377.
- Wang, H.C., Xia, X., Hickey, D., 2006. Thermal adaptation of the small subunit ribosomal RNA gene: a comparative study. *J Mol Evol.* 63, 120-126.
- Xia, X., Xie, Z. and Kjer, K.M., 2003. 18S ribosomal RNA and tetrapod phylogeny. *Syst. Biol.* 52, 283-295.

Localization of DNA Sequences Tightly Associated with the Synaptonemal Complex in Compositional Fractions of the Golden Hamster Genome*

O. I. Karpova¹, S. Saccone², A. Varriale³, T. V. Sizova¹,
M. V. Penkina¹, and Yu. F. Bogdanov¹

¹Vavilov Institute of General Genetics, Russian Academy of Sciences,
ul. Gubkina 3, Moscow, 119991 Russia

E-mail: karpova@vigg.ru

²Dipartimento di Biologia Animale "M. La Greca,"
via Androne 81, 95124 Catania, Italy

³Stazione Zoologica "Anton Dohrn," Villa Comunale, 80121 Napoli, Italy

Received January 5, 2004

Abstract—Synaptonemal complex (SC) isolated from spermatocyte nuclei after their exhaustive hydrolysis by DNase II contains DNA sequences tightly associated with it (SCAR DNA). Here, the compositional properties of a cloned family of golden hamster SCAR DNA were studied. For this purpose, 27 SCAR DNA clones were hybridized with compositionally fractionated golden hamster genomic DNA. The sequences of the SCAR DNA family were mainly localized in the GC-poor isochore families L1 and L2, which accounted for 63% of hybridization signals. The remaining 37% of signals pertained to the GC-rich isochore families H1 and H2. Thus, SCAR DNA proved to be distributed throughout the genome, irrespective of differences in density and sequence type between isochore families. Moreover, the SCAR DNA sequences containing the regions of homology with LINE/SINE repeats were found in all the isochore families. The compositional localization of SCAR DNA is in agreement with the hypothesis that the SC and SCAR DNA participate in chromatin reorganization during meiosis prophase I, which should result in the attachment of chromatin loops to the lateral elements of SC throughout its length.

Key words: golden hamster, meiosis, DNA regions associated with synaptonemal complex, compositional fractionation, isochores

INTRODUCTION

Meiosis in a cell is provided for by the work of more than a hundred genes that are known today [1–3]. The synaptonemal complex (SC) formation and the process of chromatin spatial packing in prophase I of meiosis are also under genetic control [4–6; for review, see 7]. The SC is formed along two homologous chromosomes and resembles a zipper: its distinctive morphological feature in animals, plants, and fungi is the presence of two parallel lateral elements located at a certain distance from each other. Moving from the opposite ends, transverse filaments extend across this distance and meet in the middle, forming the central SC element [8, 9]. As shown in numerous microscopic investigations, chromatin at prophase I of meiosis is spatially arranged in loops tightly adjoining the lateral SC elements [9–12].

Studies of SCAR DNA have shown that they form a distinct family of genomic DNA consisting mainly of unique sequences that contain no extended open reading frames, are enriched in both microsatellite repeats (GT/AC)_n [13] and LINE/SINE sequences [14], and have a number of other structural peculiarities.

Intracellular functions of SCAR DNA are poorly known. These sequences are supposedly involved in the process of homologous chromosome recognition and synapsis [1, 6, 14]. During crossing-over, the SC keeps homologous chromosomes in pairs, which provides for the chiasma formation and crossing-over interference, and is probably involved in crossing-over proper [2, 15, 16]. In addition, the SC and SCAR DNA are likely to participate in the chromatin rearrangement in prophase I, during which chromatin loops are attached to the lateral elements of SC throughout its length.

The vertebrate genome represents a mosaic of isochores, compositionally fairly homogeneous DNA fragments belonging to several families (two to six, depending on the species). In the golden hamster genome, four isochore families are distinguished: the

* This article was submitted by the authors in English.

Abbreviations: SC, synaptonemal complex; SCAR DNA, synaptonemal complex-associated regions of DNA.

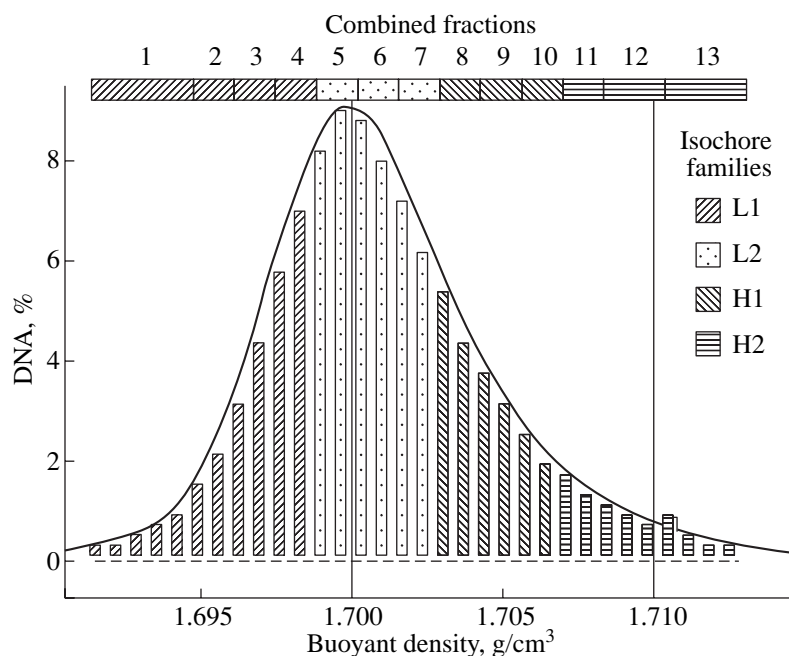


Fig. 1. Compositional properties of the golden hamster DNA fractions. Each bar corresponds to a DNA fraction obtained in a shallow density gradient and characterized with respect to relative DNA content, modal buoyant density, and isochore family. The CsCl profile of the hamster main band DNA is also shown. Figures on the top (1–13) show the numbers of combined fractions used as compositional probes.

GC-poor L1 and L2, and the GC-rich H1 and H2 [17]. In the mammalian genomes, two isochore patterns, general and special, were described, with the former implying a wider compositional distribution of DNA fragments, compared to the latter [18]. The vast majority of mammalian genomes comply with the general model, whereas the special pattern has been described in only a few groups, such as the suborder Myomorpha (Rodentia) to which the golden hamster belongs [18, 19]. Interestingly, DNA sequences in the GC-rich compartments of the genomes of warm-blooded vertebrates have a high degree of homology [19].

The human genome, as well as the genomes of other warm-blooded vertebrates—for example, birds [20]—has a number of structural and functional properties directly related to the GC levels and, specifically, to the compositional features of the isochores [21]. These properties include gene density, replication timing in the S phase of the cell cycle, recombination frequency, chromatin compaction, and the level of transcription activity. Moreover, repeated sequences such as LINES and SINES are distributed in the human genome nonuniformly, concentrating in the GC-rich and GC-poor isochores, respectively [22].

The purpose of this work was to localize the cloned SCAR DNA sequences in the compositional fractions of the golden hamster genome.

EXPERIMENTAL

High-molecular-weight DNA was isolated from the hamster liver and fractionated in a shallow density gradient as described [23]. The resulting DNA fraction was then compositionally characterized by analytical centrifugation in a CsCl density gradient [18, 24].

We used 27 cloned sequences of golden hamster SCAR DNA with the accession numbers Z32797–Z32799, Z32801–Z32803, Z32805–Z32812, Z32837, Z86070–Z86078, Z86080, Z86083–Z86088, and AJ626864–AJ626874 (EMBL Database).

The SCAR DNA clones (5.0, 1.0, and 0.2 μg per spot) were applied to Hybond-N nylon filters (Amersham) using a dot-blot device. The DNA fractions obtained as described above were ^{32}P -labeled and used as probes in DNA–DNA hybridization [25]. The results of hybridization were recorded by exposing filters in the Cyclone Phosphor Imager System and quantitatively assessed using the OptiQuant software.

RESULTS

The hamster genomic DNA was separated into 32 fractions differing in modal buoyant density (as defined by CsCl analytical centrifugation) and DNA content (Fig. 1). They were combined as shown in Fig. 1, and each of the resulting 13 fractions was used as a probe in hybridization with 27 clones of hamster SCAR DNA immobilized on filters. These DNA frac-

Table 1. Compositional distribution of SCAR DNA clones containing unique sequences

Group of SCAR DNA clones	SCAR DNA clones	Isochore families*												
		L1				L2			H1			H2		
		compositional DNA probes**												
		1	2	3	4	5	6	7	8	9	10	11	12	13
1	4	+	+	+	+	+	+	+						
	5				+		+	+						
	6	+	+	+	+		+							
	8		+		+		+							
	59		+		+		+	+						
	60		+		+									
2	221				+									
	53		+	+	+		+	+				+	+	+
	57	+	+		+		+	+		+	+			+
3a	77	+	+	+	+					+				
	1	+	+	+	+			+	+	+	+			+
	84				+				+	+	+	+		+
	101	+		+		+	+	+			+			+
3b	174				+	+		+		+	+			+
	76				+			+			+			
	82	+					+				+			
4	219	+										+		+
	134								+					
	242									+				
	260										+			+

* Isochore families contained in the corresponding DNA probes.

** Each probe corresponds to one of combined fractions shown in Fig. 1.

tions, encompassing the entire compositional range of the hamster genomic DNA (compositional DNA probes; Tables 1, 2), allowed the compositional localization of SCAR DNA clones in the isochore families of the golden hamster.

Figure 2 shows the distribution of SCAR DNA clones on filters. The control spots contained the DNA of plasmid vector pBS(+), in which the SCAR DNA sequences were cloned [26]. The results of hybridization corresponding to the SCAR DNA clones containing only unique sequences and to the clones exhibiting homology with repeated sequences are shown separately in Tables 1 and 2, respectively. Note that the isochore family positions shown in Fig. 1 are approximate, because the boundaries between contiguous isochore families could not be precisely located [21]. In fact, some fractions contained a mixture of the two adjacent isochore families, such as the compositional DNA probe 5 that contained DNA fragments from both the L1 and the L2 isochores (Table 1).

The compositional localization of the SCAR DNAs from the golden hamster genome showed that these sequences were distributed, with different densities, over all the isochore families (Tables 1, 2). Among the clones containing only unique sequences, four groups could be distinguished (Table 1). The first group included the sequences located only in the GC-poor isochore families L1 and L2. Moreover, two of these clones, 60 and 221, were only found in the GC-poorest isochore family L1. The second group consisted of the SCAR DNA sequences localized mainly in the GC-poor isochore families L1 and L2 and, to a lesser extent, in the GC-rich isochore families H1 and H2. The third group included the sequences evenly distributed in the genome, either widespread (clones 1, 84, 101, and 174) or rare (clones 76, 82, and 219). The fourth group consisted of the sequences localized only in the GC-rich isochore families H1 and H2 (clones 134, 242, and 260).

The SCAR DNA clones exhibiting homology to the SINE (short interspersed nuclear element)

Table 2. Compositional distribution of SCAR DNA clones containing repeated sequences

Type of repeats	SCAR DNA clones	Isochore families*												
		L1				L2			H1		H2			
		compositional DNA probes**												
		1	2	3	4	5	6	7	8	9	10	11	12	13
SINEs	16		+	+	+	+	+	+	+		+	+	+	+
	96	+	+		+	+	+	+	+		+	+		+
	103	+	+	+	+	+	+	+	+	+	+	+	+	+
	162		+		+	+	+	+	+	+	+	+	+	+
LINEs	120	+		+		+	+	+			+	+		+
	133	+	+	+	+	+	+	+	+		+	+	+	+
	203	+	+											

* Isochore families contained in the corresponding DNA probes.

** Each probe corresponds to one of combined fractions shown in Fig. 1.

sequences were widely represented in the hamster genome. The sequences homologous to the Alu and B1 repeats and to the MIR (mammalian-wide interspersed repeat) sequences were localized in both GC-poor and GC-rich isochores. In fact, there were no significant differences in the localization of each SINE repeat. It should be noted that the sequence of clone 103, homologous to the B1 repeat, was not only located in all the isochore families, but the corresponding hybridization signal always had the highest intensity.

The SCAR DNA sequences having regions homologous to the LINE (long interspersed nuclear element) repeats were localized in all the isochore families. However, clone 203 was only observed in the GC-poorest isochore family L1 (Table 2).

The presence of microsatellite (GT/AC)_n repeats of different lengths in clones 4, 16, and 219 did not interfere with their compositional localization in the isochore families. In fact, clones 4 and 219 did not differ significantly from other unique SCAR DNA sequences in the type of hybridization. The greatest number of microsatellite repeats ($n = 40$) was found in clone 16, but its compositional localization in the isochore families depended on homology with the MIR sequences and, hence, was similar to the localization of other SCAR DNA clones having homology with the SINE repeats (clones 96, 103, and 162).

On the whole, it was shown that 97 out of 154 hybridization signals (63%) were recorded after hybridization of SCAR DNA clones with the GC-poor isochore families L1 and L2, and 57 signals (37%), after hybridization with the GC-rich isochore families H1 and H2 (Tables 1, 2).

DISCUSSION

Studies on SCAR DNA isolated from the golden hamster [13, 26] and rat [14] genomes have shown that it consists mainly of unique sequences containing no extended open reading frames. The SCAR DNA sequences constitute a special family of nuclear DNA with a number of specific characteristics, such as enrichment with both microsatellite repeats (GT/AC)_n [13] and LINE/SINE sequences [14]. In addition, the SCAR DNA of the golden hamster is rich in short palindromes [13] and differs from the genomic DNA in a higher frequency of some tri- and tetranucleotides [27, 28].

Many properties of chromosomes and chromatin correlate with the isochore organization of the genome. The distribution of the GC-richest (GC > 50%) and the GC-poorest (GC < 37%) isochores in the mammalian chromosomes provides for the existence of two compositionally different chromosomal regions: the former isochores usually concentrate in the telomeric regions, and the latter, in the more inner parts of chromosomes [29–33]. Based on isochore composition estimated by *in situ* hybridization with probes for different isochore families, these regions were named H3+ and L1+, respectively. This nomenclature was confirmed by analyzing the recent draft of the human genome sequence [34, 35]. Moreover, in the nuclei of human cells, two different regions have been described: the region located closer to the center contains the GC-richest and gene-richest isochores at a very high level of chromatin decondensation, and the other, peripheral region contains the GC-poorest and gene-poorest isochores at a very high level of chromatin compaction [36]. However, no information on SCAR DNA localization in the isochores has been obtained to date.

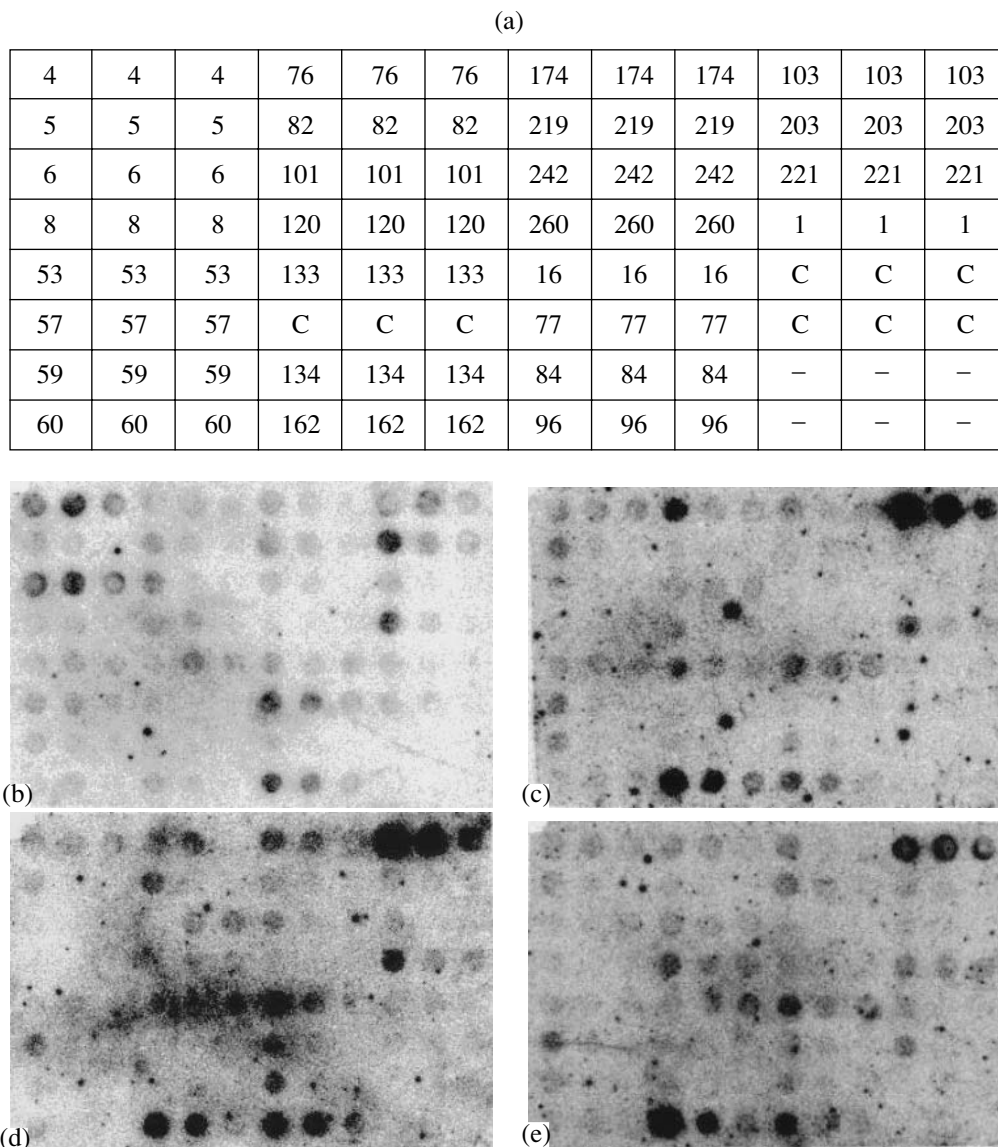


Fig. 2. Results of hybridization of SCAR DNA clones with compositional DNA probes. (a) Scheme of a dot-blot filter with the SCAR DNA clones (numbered), each applied as three spots containing 5.0, 1.0, and 0.2 µg of DNA (from left to right). Control spots (C) contained DNA of pBS(+) vector. Examples of clones hybridization with the compositional DNA probes (see Fig. 1): (b) probe 1 (isochore family L1), (c) probe 7 (L2), (d) probe 10 (H1), and (e) probe 11 (H2).

The results of this work show that the unique sequences of the SCAR DNA family are widespread in the golden hamster genome. These sequences concentrate in the GC-poor isochore families L1 and L2, but some of them also occur in the GC-rich isochore families H1 and H2. Of 87 signals of SCAR DNA hybridization with compositional DNA fractions, 44, 24, 17, and 15% were recorded in the isochore families L1, L2, H1, and H2, respectively.

A different pattern of compositional distribution was observed with the SCAR DNA clones containing regions homologous to repetitive sequences. There were no significant differences in the localization of each SINE repeat. In the human genome, Alu repeats were shown to

be located mainly in GC-rich isochores, especially in the H2 family [37, 38]. In this work, the SCAR DNA sequences containing the regions homologous to the Alu repeats were not found to be concentrated in the GC-rich isochores. On the one hand, this may be indicative of differences in isochore organization between the human and rodent genomes. In this relation, it is noteworthy that the GC-richest isochore H3 characteristic of the human genome was not detected in some rodents [18, 24, 39–41]. On the other hand, the “monotonous” hybridization of SINE-containing SCAR DNA sequences with isochores may suggest that SINE repeats from all genomic isochore fractions are equally represented in the SCAR DNA family.

The LINE (long interspersed nuclear element) repeats in the human and mouse genomes are concentrated in the GC-poor isochores, although they also occur in all other isochores [42]. The SCAR DNA sequences with regions homologous to the LINE repeats are represented in all isochores. However, clone 203 was localized only in the GC-poor isochore L1.

Thus, this study has shown that SCAR DNA sequences are located mainly in GC-poor isochore families L1 and L2 and in GC-rich isochore family H1. This may be due to their base composition (on average, 57% of AT pairs), the absence of coding sequences, and the compositional homogeneity of isochore families. A compositional correlation is observed between the isochores and the introns located therein, the third position of codons, and the intergenic regions [17]. The location of some GC-poor SCAR DNA sequences in GC-rich isochores is quite explicable. The length of SCAR DNA sequences averages 300–500 bp and does not exceed 1000 bp, whereas the compositional homogeneity with respect to GC content is manifested in the genome regions of 300 kbp and longer.

The distribution of SCAR DNA sequences over all isochore families of the golden hamster genome agrees with the concept that an important function of the SC and SCAR DNA is their participation in the structural organization of chromatin in prophase I of meiosis, which entails the appearance of chromatin loops attached to the lateral elements of the SC throughout its length. The basal parts of these loops probably contain SCAR DNA, which has been localized mainly in the isochore families L1 and L2 but is actually distributed throughout the genome.

ACKNOWLEDGMENTS

We are grateful to Prof. G. Bernardi for his constructive criticism and useful comments on the manuscript.

The work was supported by the Russian Foundation for Basic Research (project no. 02-04-49 199a).

REFERENCES

1. Roeder G.S. 1997. Meiotic chromosomes: it takes two to tango. *Genes Devel.* **11**, 2600–2621.
2. Zickler D., Kleckner N. 1999. Meiotic chromosome: integrating structure and function. *Annu. Rev. Genet.* **33**, 603–754.
3. Bogdanov Yu.F. 2003. Variation and evolution of meiosis. *Genetika* (Moscow). **39**, 453–473.
4. Meuwissen R.L.J., Offenbergh H.H., Dietrich A.J., Riesewijk A., van Iersel M., Heyting C. 1992. A coiled-coil related protein specific for synapsed regions of meiotic prophase chromosomes. *EMBO J.* **11**, 5091–5100.
5. Sym M., Engebrecht J., Roeder G.S. 1993. Zip1 is a synaptonemal complex protein required for meiotic chromosome synapsis. *Cell.* **72**, 365–378.
6. Heyting C. 1996. Synaptonemal complex: structure and function. *Curr. Opin. Cell Biol.* **8**, 389–396.
7. Penkina M.V., Karpova O.I., Bogdanov Yu.F. 2002. Synaptonemal complex proteins: specific proteins of meiotic chromosomes. *Mol. Biol.* **36**, 397–407.
8. Moses M.J. 1956. Chromosome structure in crayfish spermatocytes. *J. Biophys. Biochem. Cytol.* **2**, 215–218.
9. von Wettstein D., Rasmussen S.W., Holm P.B. 1984. The synaptonemal complex in genetic segregation. *Annu. Rev. Genet.* **18**, 331–413.
10. Bogdanov Yu.F. 1975. *Ul'trastruktura khromosom v meioze i sinaptonemnyi kompleks* (Ultrastructure of Meiotic Chromosomes and the Synaptonemal Complex). Moscow: Nauka, 33–48.
11. Moens P.B., Pearlman R.E. 1988. Chromatin organization at meiosis. *BioEssays.* **9**, 151–153.
12. Stack S.M., Anderson L.K. 2000. A model for chromosome structure during the mitotic and meiotic cell cycle. *Chromosome Res.* **9**, 175–198.
13. Karpova O.I., Safronov V.V., Zaitseva S.P., Bogdanov Yu.F. Some properties of DNA isolated from the murine synaptonemal complex fraction. *Mol. Biol.* **23**, 571–579.
14. Pearlman R.E., Tsao N., Moens P.B. 1992. Synaptonemal complexes from DNase-treated rat pachytene chromosomes contain (GT)_n and LINE/SINE sequences. *Genetics.* **130**, 865–872.
15. Stack S.M. 1984. Heterochromatin, the synaptonemal complex and crossing over. *J. Cell Sci.* **71**, 159–176.
16. Schwarzacher T. 2003. Meiosis, recombination and chromosomes: a review of gene isolation and fluorescent *in situ* hybridization data in plants. *J. Exp. Botany.* **54**, 11–23.
17. Bernardi G. 2000. Isochores and evolutionary genomics of vertebrates. *Gene.* **224**, 3–17.
18. Sabeur G., Macaya G., Kadi F., Bernardi G. 1993. The isochore patterns of mammalian genomes and their phylogenetic implications. *J. Mol. Evol.* **37**, 93–108.
19. Caccio S., Perani P., Saccone S., Kadi F., Bernardi G. 1994. Single-copy sequence homology among the GC-richest isochores of the genomes from warm-blooded vertebrates. *J. Mol. Evol.* **39**, 331–339.
20. Andreozzi L., Federico C., Motta S., Saccone S., Sazanov A., Sazanov A., Smirnov A., Galkina S.A., Lukina N.A., Rodionov A.V., Carels N., Bernardi G. 2001. Compositional mapping of chicken chromosomes and identification of the gene-richest regions. *Chromosome Res.* **9**, 521–532.
21. Bernardi G. 2004. *Structural and Evolutionary Genomics*. Amsterdam: Elsevier Science.
22. Soriano P., Meunier-Rotival M., Bernardi G. 1983. The distribution of interspersed repeats is non-uniform and conserved in the mouse and human genomes. *Proc. Natl. Acad. Sci. USA.* **80**, 1816–1820.
23. De Sario A., Geigl E.M., Bernardi G. 1995. A rapid procedure for the compositional analysis of yeast artificial chromosomes. *Nucleic Acids Res.* **23**, 4013–4014.

24. Thiery J.-P., Macaya G., Bernardi G. 1976. An analysis of eukaryotic genomes by density gradient centrifugation. *J. Mol. Biol.* **108**, 219–235.
25. Sambrook J., Fritsch E.F., Maniatis T. 1989. *Molecular Cloning. A Laboratory Manual*. Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press.
26. Karpova O.I., Penkina M.V., Dadashev S.Ya., Milshina N.V., Hernandez J., Radchenko I.V., Bogdanov Yu.F. 1995. Primary structure of golden hamster synaptonemal complex DNA. *Mol. Biol.* **29**, 512–521.
27. Penkina M.V., Karpova O.I., Dadashev S.Ya., Milshina N.V., Bogdanov Yu.F. 1997. Periodicity of trinucleotide distribution in golden hamster synaptonemal complex DNA. *Mol. Biol.* **31**, 240–243.
28. Penkina M.V., Karpova O.I., Dadashev S.Ya., Milshina N.V., Bogdanov Yu.F. 1999. Distribution of trinucleotide repeats in the family of golden hamster DNA sequences tightly associated with the synaptonemal complex. *Mol. Biol.* **33**, 586–591.
29. Saccone S., De Sario A., Della Valle G., Bernardi G. 1992. The highest gene concentration in the human genome are in T-bands of metaphase chromosomes. *Proc. Natl. Acad. Sci. USA.* **89**, 4913–4917.
30. Saccone S., Caccio S., Perani P., Andreozzi L., Rapisarda A., Motta S., Bernardi G. 1997. Compositional mapping of chromosomes and identification of the gene-rich regions. *Chromosome Res.* **5**, 293–300.
31. Saccone S., Federico C., Solovei I., Croquette M.F., Della Valle G., Bernardi G. 1999. Identification of the gene-richest bands in human prometaphase chromosomes. *Chromosome Res.* **7**, 379–386.
32. Federico C., Andreozzi L., Saccone S., Bernardi G. 2000. Gene density in the Giemsa bands of human chromosomes. *Chromosome Res.* **8**, 737–746.
33. Federico C., Andreozzi L., Motta S., Saccone S., Russo V., Carels N., Bernardi G. 2004. Identification of the pig chromosomal regions compositionally correlated to the human gene-richest bands. *Gene*. In press.
34. Saccone S., Pavlicek A., Federico C., Paces J., Bernardi G. 2001. Genes, isochores and bands in human chromosomes 21 and 22. *Chromosome Res.* **9**, 533–539.
35. Lander E.S., Linton L.M., Birren B., Nusbaum C., Zody M.C., *et al.* 2001. Initial sequencing and analysis of the human genome. *Nature.* **409**, 860–921.
36. Saccone S., Federico C., Bernardi G. 2002. Localization of the gene-richest and gene-poorest isochores in the interphase nuclei of mammals and birds. *Gene.* **300**, 169–178.
37. Zerial M., Salinas J., Filipski J., Bernardi G. 1986. Gene distribution and nucleotide sequence organization in the human genome. *Eur. J. Biochem.* **160**, 479–485.
38. Jabbari K., Bernardi G. 1998. CpG doublets, CpG islands and Alu repeats in long human DNA sequences from different isochore families. *Gene.* **224**, 123–128.
39. Salinas J., Zerial M., Filipski J., Bernardi G. 1986. Gene distribution and nucleotide sequence organization in the mouse genome. *Eur. J. Biochem.* **160**, 469–478.
40. Mouchiroud D., Gautier C. 1988. High codon-usage changes in mammalian genes. *Mol. Biol. Evol.* **5**, 192–194.
41. Robinson M., Gautier C., Mouchiroud D. 1997. Evolution of isochores in rodents. *Mol. Biol. Evol.* **14**, 823–828.
42. Smit A.F.A. 1999. Interspersed repeats and other mementos of transposable elements in mammalian genomes. *Curr. Opin. Genet. Dev.* **9**, 657–663.



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

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.

© 2006 Elsevier B.V. All rights reserved.

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, in press).

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 Notothe-

nioids 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 ichthyofauna. 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, ρ_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	<i>Pantodon buchholzi</i>		0.77	45.66	1.77	3.88		
	Notopteridae	<i>Notopterus notopterus</i>			44.96	1.22	2.72		
Clupeiformes	Clupeidae	<i>Sardina pilchardus</i>	a		47.12	1.40	2.98		
Cypriniformes	Cyprinidae	<i>Carassius auratus</i>	b	1.60	39.53	1.44	3.64		
		<i>Brachydanio rerio</i>	b	1.68	39.19	1.35	3.45	39.36	1.40
Salmoniformes	Salmonidae	<i>Oncorhynchus keta</i>		2.49	45.75	1.49	3.25		
Gadiformes	Merlucciidae	<i>Merluccius merluccius</i>	c		48.69	2.18	4.47		
		<i>Merluccius bilinearis</i>		0.84					
	Gadidae	<i>Boreogadus saida</i>	d	0.88	48.40	2.22	4.58		
		<i>Gadus morhua</i>	d	0.40	48.61	2.37	4.88		
		<i>Arctogadus glacialis</i>	d		48.13	2.74	5.68		
		<i>Notobranchius flammicomantis</i>	e		47.26	1.59	3.36		
Cyprinodontiformes	Aplocheilidae	<i>Aphyosemion elegans</i>	f		43.49	1.52	3.49		
		<i>Aphyosemion coeruleum</i>		1.20					
		<i>Aphyolebias peruensis</i>	e		45.70	1.50	3.28		
	Rivulidae	<i>Fundulus heteroclitus</i>		1.36	42.80	1.76	4.11		
	Poeciliidae	<i>Xiphophorus maculatus</i>		0.76	41.28	1.39	3.37		
	Cyprinodontidae	<i>Jordanella floridae</i>			41.43	1.10	2.66		
		<i>Aphanius fasciatus</i>	g		43.17	1.62	3.75		
Zeiformes	Caproidae	<i>Capros aper</i>	a		46.69	1.87	4.00		
Scorpaeniformes	Scorpaenidae	<i>Scorpaena guttata</i>	b		41.27	1.48	3.58		
		<i>Scorpaena brasiliensis</i>		1.40					
	Hexagrammidae	<i>Ophiodon elongatus</i>	b		44.31	1.78	4.02		
	Cottidae	<i>Leptocottus armatus</i>	b		46.72	2.11	4.52		
	Liparidae	<i>Liparis tunicatus</i>	d	0.88	46.30	2.79	6.02		
Perciformes	Serranidae	<i>Serranus cabrilla</i>	a		44.39	1.66	3.75		
		<i>Alphestes immaculatus</i>	b		43.40	1.40	3.23		
		<i>Paralabrax maculatofasciatus</i>	b		43.63	1.36	3.13		
	Apogonidae	<i>Apogon imberbis</i>	a		43.25	1.51	3.50		
		<i>Apogon bandanensis</i>		0.94					
	Carangidae	<i>Trachurus mediterraneus</i>	a		44.58	1.65	3.70		
		<i>Trachurus declivis</i>		0.90					
	Sparidae	<i>Boops boops</i>	c	0.52	44.65	1.83	4.10		
	Mullidae	<i>Mullus barbatus</i>	c		48.86	2.19	4.49		
	Pomacanthidae	<i>Holocanthus passer</i>	b		44.10	1.43	3.25		
	Kyphosidae	<i>Hermosilla azurea</i>	b		44.26	1.41	3.18		
		<i>Microspathodon dorsalis</i>	b		43.15	1.54	3.57		
	Pomacentridae	<i>Chromis sp.</i>	b	1.10	43.80	1.50	3.42	43.48	1.52
		<i>Coris julis</i>	c	1.21	44.90	1.28	2.85		
	Labridae	<i>Thalassoma grammaticum</i>	b		42.59	1.29	3.02		
		<i>Halichoeres poeyi</i>	b		44.43	1.41	3.16		
		<i>Halichoeres bathyphilus</i>		0.80					
		<i>Symphodus tinca</i>	a	0.73	45.22	1.81	4.01		
	Scaridae	<i>Scarus ghobban</i>	b		41.95	1.32	3.15		
		<i>Scarus guacamaia</i>		1.25					
Zoarcidae	<i>Iluocoetes fimbriatus</i>	h*	1.07	46.62	1.92	4.11			
Bovichtidae	<i>Cottoperca gobio</i>	h*		43.65	1.98	4.54			
	<i>Bovichtus diacanthus</i>	h*		41.95	1.83	4.35			
Eleginopidae	<i>Eleginops maclovinus</i>	h*		44.04	1.66	3.76			
Nototheniidae	<i>Notothenia coriiceps</i>	h		44.40	1.87	4.22			
	<i>Notothenia rossii</i>	h		44.52	1.78	4.00	44.46	1.83	
	<i>Gobionotothen gibberifrons</i>	h		42.62	2.15	5.05			
	<i>Gobionotothen marionensis</i>	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	<i>Trematomus newnesi</i>	h		44.57	1.82	4.09			
		<i>Trematomus bernacchii</i>	h	1.19	43.59	1.82	4.17	44.08	1.82	
		<i>Dissostichus mawsoni</i>	h	1.02	44.09	1.85	4.20			
		<i>Patagonotothen guntheri</i>	h*		44.08	1.69	3.84			
		<i>Lepidonotothen squamifrons</i>	h*		43.79	1.87	4.27			
		<i>Lepidonotothen kempfi</i>	h		43.31	1.64	3.78			
			<i>Lepidonotothen nudifrons</i>	h		43.74	1.79	4.09	43.61	1.77
	Bathydraconidae	<i>Gymnodraco acuticeps</i>	h		42.60	1.80	4.23			
			h		43.24	1.99	4.59	42.92	1.89	
				1.41						
	Channichthyidae	<i>Chionodraco rastrospinosus</i>	h		43.65	2.28	5.22			
			h	1.83	43.90	2.40	5.47	43.78	2.34	
			h*		45.53	2.22	4.88			
			h*		44.90	2.22	4.94			
			h		44.27	2.10	4.74			
			c		43.80	2.08	4.75			
	Tripterygiidae	<i>Lepidonectes corallicola</i>	b		44.52	1.42	3.20			
			b		45.61	1.85	4.05			
	Pleuronectiformes	Labrisomidae	<i>Dialommus fuscus</i>	b		45.61	1.85	4.05		
<i>Paralichthys californicus</i>			b	0.80	46.50	1.58	3.40			
Tetraodontiformes	Tetraodontidae	<i>Tetraodon fluviatilis</i>	b		48.39	1.93	3.99			
		<i>Tetraodon nigroviridis</i>		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 10². 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 (ρ_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 μ l 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 μ l of 20 mM zinc sulfate were added to the samples followed by 10 U (units) of nuclease P1 (1 U/ μ l 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 μ l 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 KH₂PO₄ sterilized in autoclave and filtered through a Millipore GS-22 filter (0.22 μ m). Solvent B was 95% (v:v) methanol HPLC grade (J.T. Baker).

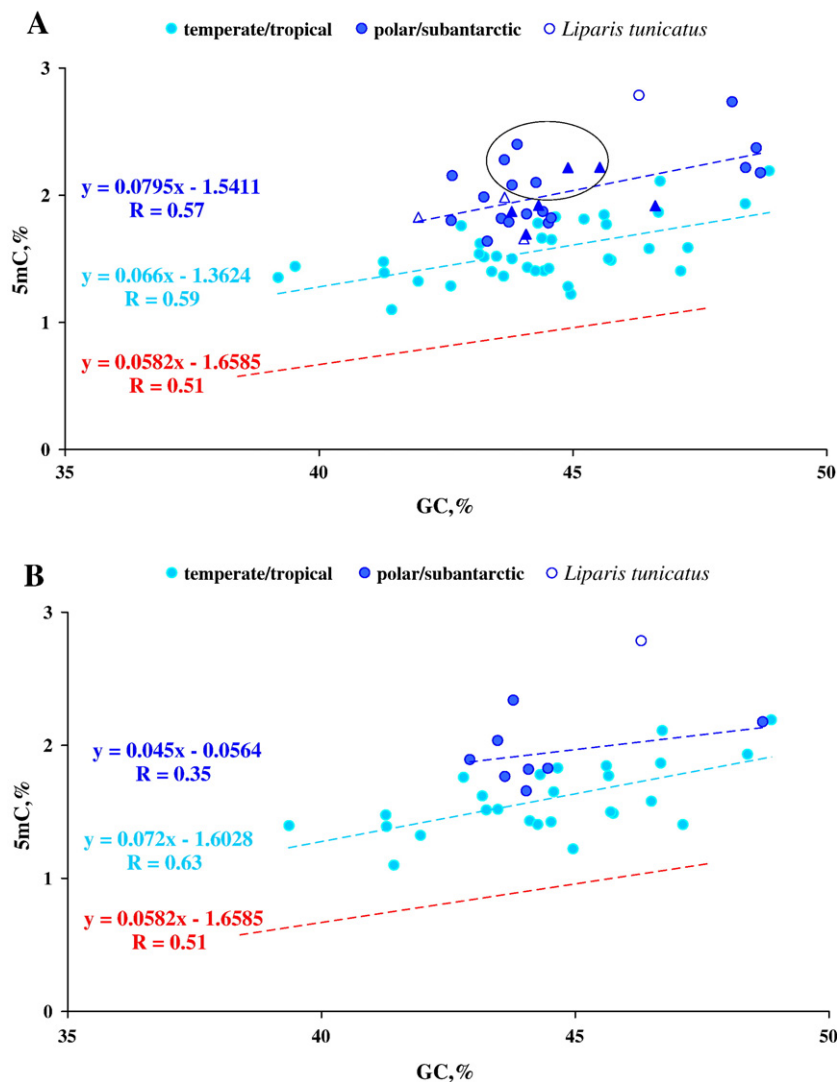


Fig. 1. A. Plot of 5mC levels against GC levels for the genomes of polar fishes (dark blue circles, dark blue triangles which concern sub-Antarctic fishes of polar origin and containing AFGPs, and empty triangles concerning native sub-Antarctic) and temperate/tropical fishes (light blue circles); the 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 λ_{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 μ g DNA samples from calf thymus dissolved in 20 μ l 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

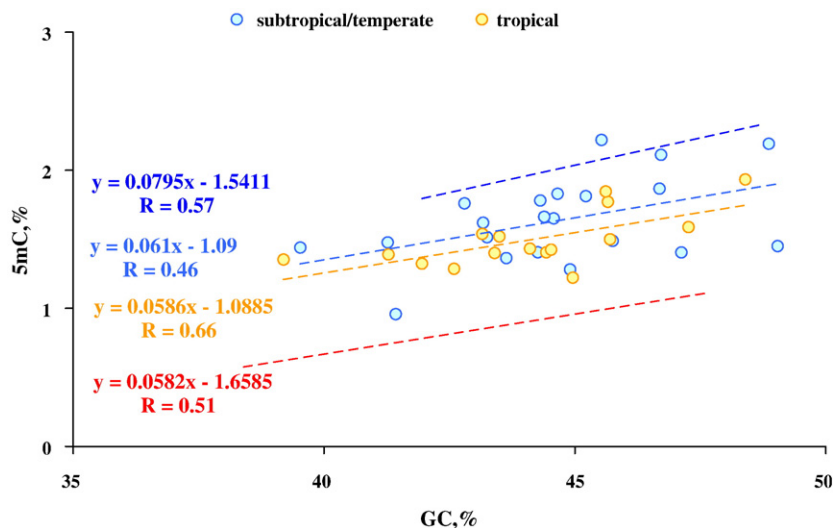


Fig. 2. Plot of 5mC levels against GC levels for the genomes of temperate and subtropical fishes (upper line) and of tropical fishes. Equations and correlation coefficients are given for the two groups as well as for polar fishes and mammals whose regression lines were given as reference. The classification 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, *Gobiidae*, *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

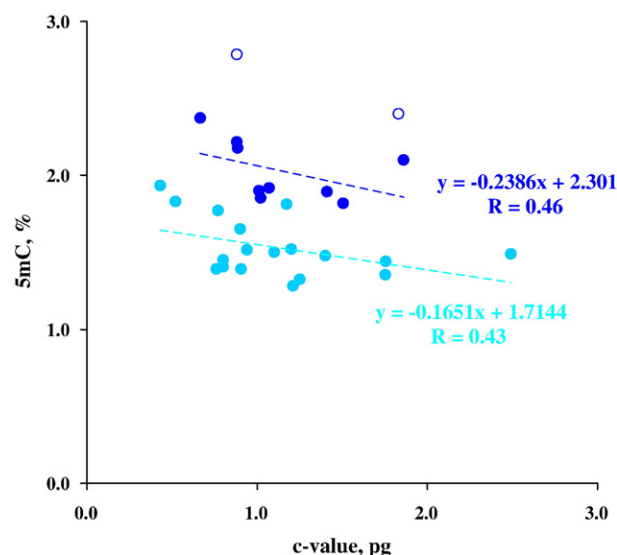


Fig. 3. Plot of 5mC against genome size for polar fishes (dark blue circles) and temperate/tropical fishes (light blue circles). The broken dark blue line refers to the correlation when all points were taken into account except for the two outliers (empty circles). (For interpretation of the references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

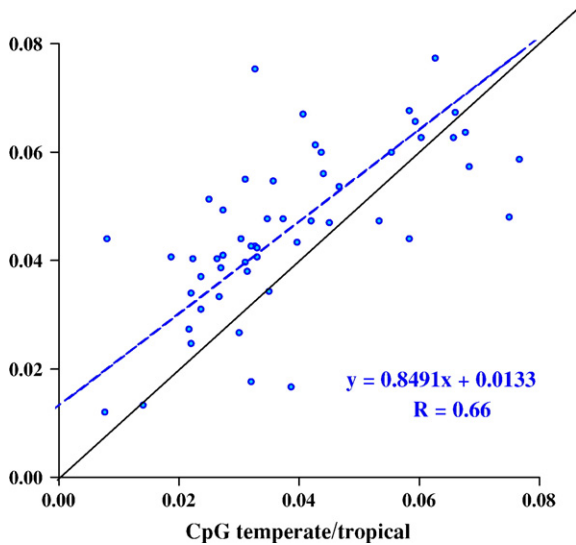


Fig. 4. CpG levels of genes from polar fishes are plotted against CpG levels of orthologous genes from temperate/tropical fishes listed in Supplementary Table 3. The solid line is the diagonal.

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; Aissani 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 *Notothernioidei*, 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,

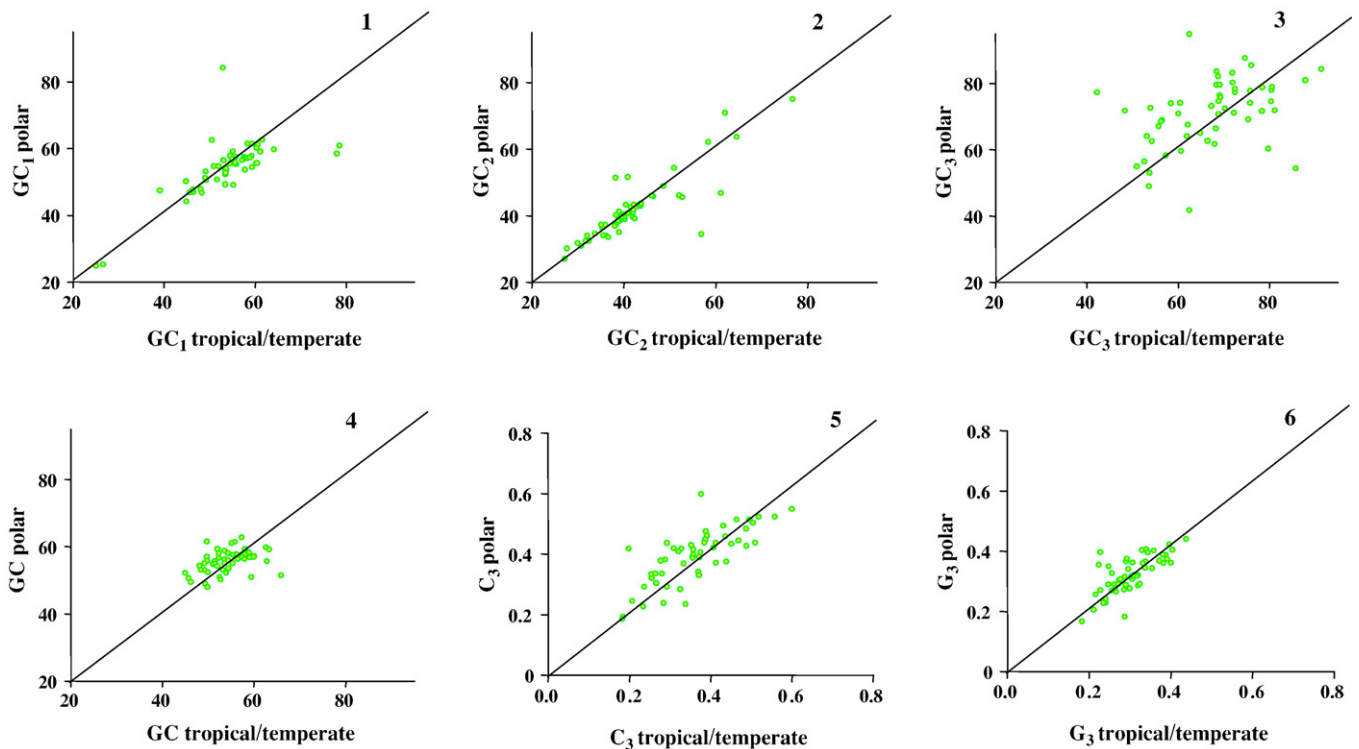


Fig. 5. GC, GC1, GC2, GC3, C3 and G3 levels from the genes from polar fishes are plotted against the corresponding levels of orthologous genes from temperate/tropical fishes listed in Supplementary Table 3. The solid lines are the diagonals.

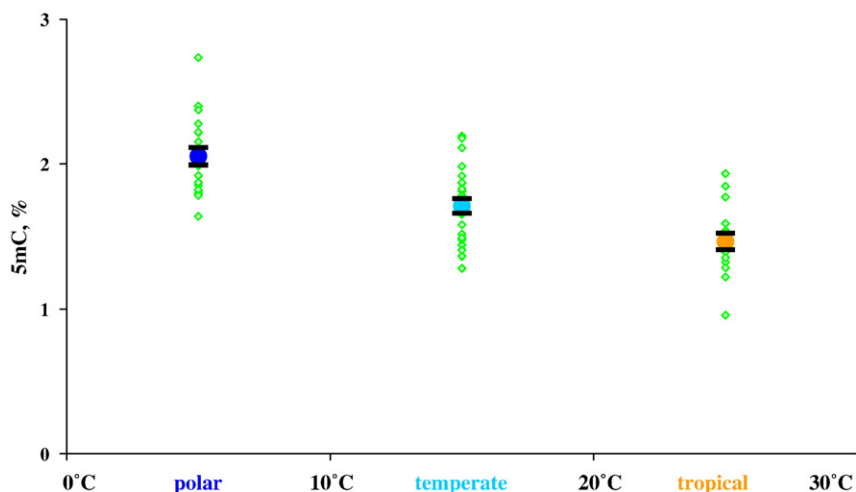


Fig. 6. Correlation between 5mC level and body temperature. The two extreme points correspond to the genomes of polar and tropical fishes, respectively; the middle point is the average body temperature of temperate fishes. Standard error bars are shown.

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

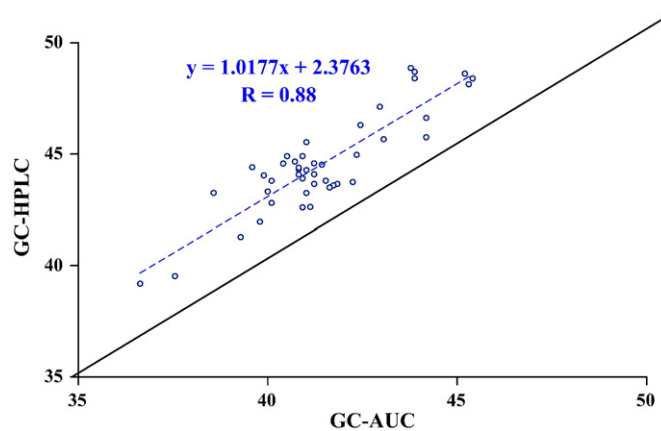


Fig. 7. Correlation between genomic GC levels as obtained by nucleoside analysis (HPLC) and by analytical ultracentrifugation (AUC) in fishes. Analytical ultracentrifuge profiles of fish DNA are reported in Supplementary Figs. 1 and 2.

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 *Gobiidae*, *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 Comeo 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 Tromsø, Norwegian College of Fishery Science, Tromsø, 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 constraints 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., Guernonprez, F., Desgrès, J., Bernardi, G., 1997. Methylation patterns in the isochores 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. *Cybius* 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., Desgrès, 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., in press. DNA methylation in reptiles. *Gene*. doi:10.1016/j.gene.2006.05.034.



DNA methylation in reptiles

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

Received by Ettore Olmo

Abstract

Very recent investigations have provided evidence for a higher DNA methylation level in polar and sub-antarctic fishes compared to temperate/tropical fishes, the latter being in turn higher than the DNA methylation level of warm-blooded vertebrates. These results confirm and extend the finding [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] that DNA methylation level of vertebrates is inversely related to body temperature. Here we studied the methylation level of reptilian genomes. The species previously analyzed exhibited methylation levels closer to those of mammals and birds rather than to those of fishes and amphibians. The sample was, however, too small to reach a final conclusion. Here we used Reversed-Phase-High-Performance Liquid Chromatography (RP-HPLC) to analyze the DNA methylation levels of 43 reptiles representing three out of four orders and 20 families. Such analysis has shown that snakes and lizards exhibit methylation levels covering the whole range comprised between those of temperate/tropical fish and mammals, while turtles, and, more so, crocodiles are close to mammals. We discuss some ecological and physiological data that explain these results.

© 2006 Elsevier B.V. All rights reserved.

Keywords: Body temperature; Homeotherms; Heterotherms; Life style

1. Introduction

Results presented in the preceding paper (Varriale and Bernardi, 2006-this issue) made clear that an inverse correlation holds between body temperature and DNA methylation in fishes. Here we approached the problem of DNA methylation in the case of reptiles. The reptiles previously studied (Jabbari et al., 1997) comprised one crocodile, two turtles, one varan and three snakes. This sample was small, heterogeneous (since they comprised three old data from Vanyushin et al. (1970, 1973)) and not representative from a phylogenetic point of view since the two turtles and two of the snakes belonged to the same genera. These reptiles showed DNA methylation levels that covered a two-fold range, a majority of values being closer to those of warm-blooded rather than to those of cold-blooded vertebrates. The results available so far suggested therefore that reptiles might be characterized by a

lower methylation level compared to fishes and amphibians, possibly indicating a hypothetical methylation transition at the emergence of the common ancestor of reptiles and mammals, namely at the appearance of amniotes.

We warned, however, that the number of reptilian DNAs analyzed was small, that rather different levels were found in the few reptiles investigated and that no estimate was available for the CpG levels (Jabbari et al., 1997). Under these circumstances, two possibilities should be considered, either that a reptilian ancestor had undergone a decrease in methylation, which was then transmitted by descent to extant reptiles, mammals and birds; or that the methylation changes occurred independently in reptiles (possibly, only in some of them), in mammals and in birds. Here we have expanded the sample to 43 reptiles representing three out of four orders and 18 families and have reached the conclusion that the second hypothesis is correct.

2. Materials and methods

For Materials and methods see preceding paper (Varriale and Bernardi, 2006-this issue). Modal buoyant densities of reptilian

Abbreviations: RP-HPLC, Reversed-Phase-High-Performance Liquid Chromatography; GC, molar ratio of guanosine+cytidine; 5mC, 5-methylcytosine.

* Corresponding author. Tel.: +39 081 5833215; fax: +39 081 2455807.

E-mail address: bernardi@szn.it (G. Bernardi).

DNAs are reported in Supplementary Table S2 and analytical ultracentrifuge profiles of reptilian DNAs in Supplementary Fig. S1.

3. Results

3.1. Analysis of 5mC level in reptiles

Table 1 presents the taxonomy and the sources of the genomes analyzed as well as the percent values of GC and 5mC obtained from the RP-HPLC analyses for each of the reptilian species, and the ratios 5mC/GC. The geographic distribution of the species is presented in Supplementary Table S1. Our dataset comprises 43 reptilian species covering three out of the four extant orders of reptiles, *Testudines* (turtles), *Squamata* (snakes and lizards), and *Crocodylia* (crocodiles). The fourth order, *Rhynchocephalia*, only comprises two living species belonging to the same family, *Sphenodontidae*. The only value from literature that we used was *Testudo graeca* because it was analyzed by our group with RP-HPLC (Jabbari et al., 1997). We did not use the other values from the literature because either they had been obtained with techniques different from RP-HPLC or they were reanalyzed in the present work.

The values of 5mC and GC are plotted in Fig. 1. It is immediately noticeable that the 5mC levels of reptiles show a scatter covering the whole range of values from mammals to tropical/temperate fishes if we consider both species (Fig. 1a) and families (Fig. 1b). In fact, the two groups of crocodylians and turtles are the GC-richest and the least methylated genomes, whereas squamates represent, in terms of DNA methylation, the most heterogeneous order among reptiles.

The different behavior of DNA methylation of reptiles as plotted in Fig. 1, compared to all other vertebrate classes and groups (including polar, temperate and tropical fishes; see Varriale and Bernardi, 2006-this issue), is stressed by the fact that no significant correlation ($R=0.04$) is obtained in this case. The correlation found when plotting 5mC vs. genome size (Fig. 2) is remarkably weaker ($R=0.12$) than those found for fishes (see Varriale and Bernardi, 2006-this issue). Moreover the slope is very slightly positive whereas it is negative for fishes. These findings indicate that in reptiles methylation is so variable that no effect of genome size can be detected.

3.2. Analysis of CpG doublets in reptiles

A comparison of the frequency of CpG doublets of coding sequences from reptiles with those of orthologous sequences from mammalian (mainly human) genes was attempted, but results were not meaningful for two reasons: scarcity of available data, and the contribution of CpG islands which are more abundant in mammalian compared to reptilian genomes (Aïssani and Bernardi, 1991a,b; and our unpublished observations).

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

As in the preceding paper, GC levels of reptilian DNAs were calculated from both modal buoyant densities and from nucleoside analysis (Fig. 3). As in the case of fish DNAs (see preceding paper), the former values were lower than the latter ones by 2–3% GC. This difference is due, at least in part, to DNA methylation. Indeed when GC is derived from buoyant density, the 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 a ~4% difference in 5mC, whereas it is about ~2% in fish DNAs and ~1.5% in reptilian DNAs. An additional contribution may come from the frequency of di- and tri-nucleotides that are different compared to the reference bacterial DNAs. An interesting observation in this connection is the similarity of the plots obtained for fishes and reptiles which will be discussed elsewhere. In any case, the correlation of Fig. 3 is useful because it allows correcting GC values directly derived from ρ_0 using the equation of Schildkraut et al. (1962).

4. Discussion

4.1. DNA methylation levels of reptiles: inter-ordinal differences

Reptiles tend to be tropical and a progressive decrease in species number occurs with increasing latitude (Heatwole and Taylor, 1987), *Sphenodon* having the lowest temperature (6–16 °C) when active (Halliday and Adler, 2002). Reptiles have evolved different forms of thermoregulation (see Huey, 1982, for a review). Behavioral modulation of the thermal flux is the main thermoregulator that allows reptiles to maintain body temperature within a certain range: strategies involve basking, hibernating, changing body posture, shading, or selectively exploiting the thermal variation in the environment. Thermal physiology also plays a role in the modulation of body temperature, but analyses of body temperatures demonstrated limited variation within most genera or species. Behavioral as well as physiological adjustments might contribute to thermal homeostasis, but in changing environmental conditions reptiles can undergo large variations in body temperature. This means that the correlation between body and environmental temperatures is not straightforward.

On the basis of the above complex situation and of the correlation between DNA methylation and body temperature over extended periods of time, one should expect quite a variability in the levels of DNA methylation of reptiles. This is what the results of Fig. 1 show. Indeed, as far as genome methylation level is concerned, reptiles are very heterogeneous in comparison with other vertebrates, in agreement with other findings on several features of reptilian genomes. These concern nucleotide composition, band asymmetry in cesium chloride, compositional heterogeneity (Aïssani and Bernardi, 1991a,b;

Table 1
List of the reptilian species analyzed ⁽¹⁾

Order	Family	Species	Source ⁽²⁾	c-value ⁽³⁾	Species			Families/ genera ⁽⁵⁾		
					GC	5mC	R ⁽⁴⁾	GC	5mC	
Testudines	Chelydridae	<i>Chelydra serpentina</i>	a	2.63	47.68	1.33	2.79	48.33	1.22	
		<i>Macrochelys temminckii</i>	a		48.98	1.11				2.26
	Emydidae	<i>Trachemys scripta elegans</i>		2.62	46.98	1.14	2.43	47.21	1.05	
		<i>Chrysemis picta</i>	b	2.59	47.44	0.96	2.02			
	Testudinidae	<i>Testudo graeca</i>			45.70	0.77	1.68	47.06	0.98	
		<i>Testudo hermanni</i>		3.52						
	Cheloniidae	<i>Chelonia mydas</i>	c	2.64	47.38	1.00	2.10	47.06	0.98	
		<i>Caretta caretta</i>	c		46.74	0.96	2.05			
	Squamata	Agamidae	<i>Chlamydosaurus kingii</i>	a		44.67	0.93	2.08	46.32	1.01
		Chamaeleonidae	<i>Furcifer oustaleti</i>	d		44.49	1.14	2.57		
Iguanidae		<i>Iguana iguana</i>	a	2.89	44.33	1.36	3.08			
Phrynosomatidae		<i>Sceloporus magister</i>	e		45.84	0.85	1.85			
Gekkonidae		<i>Tarentola mauritanica</i>	f	2.65	46.59	0.94	2.01			
		<i>Gekko gecko</i>		2.84	46.05	1.09	2.37			
Gerrhosauridae		<i>Zonosaurus madagascariensis</i>	d		47.11	1.19	2.53			
Teiidae		<i>Tupinambis teguixin</i>	g	2.70	45.60	0.86	1.88			
Lacertidae		<i>Podarcis sicula</i>	d	2.20	47.02	1.47	3.13			
		<i>Podarcis muralis</i>	g	2.36	48.18	1.34	2.79			
Scincidae		<i>Mabuya gravenhorstii</i>	d		45.94	0.83	1.81			
		<i>Mabuya sp.</i>		1.72						
Anguidae		<i>Anguis fragilis</i>	f	2.23	47.60	1.01	2.11			
Boidae		<i>Acrantophis dumerili</i>	a		44.41	0.94	2.12			
		<i>Boa constrictor</i>	a	1.90	41.95	0.68	1.62			
		<i>Python molurus</i>	a		43.18	0.81	1.88			
		<i>Python reticulatus</i>	a	1.60	42.38	0.65	1.54			
		Colubridae	<i>Hierophis viridiflavus</i>	f		44.16	1.32	3.00		
			<i>Elaphe lineata</i>	f		43.90	1.40	3.19		
	<i>Elaphe obsoleta</i>			2.14	43.31	1.19	2.75			
	<i>Elaphe guttata</i>	h		41.83	1.29	3.09				
	<i>Elaphe mandarina</i>	a		43.05	1.23	2.86				
	<i>Sibon longifrenis</i>	g		47.82	1.36	2.84				
	<i>Natrix tessellata</i>	g	1.83	44.12	1.03	2.32				
	<i>Liopholidophis lateralis</i>	d		44.36	1.41	3.19				
	<i>Geodipsas sp.</i>	d		43.28	1.38	3.19				
<i>Madagascarophis sp.</i>	d		42.71	1.10	2.58					
<i>Clelia rustica</i>	g		42.84	1.31	3.06					
Elapidae	<i>Walterinnesia aegyptia</i>	g		42.77	0.80	1.86				
	<i>Notechis scutatus</i>	g		42.44	1.05	2.46				
Viperidae	<i>Bothrops jararaca</i>		2.25	42.99	1.16	2.70				
	<i>Bothrops neuwiedi</i>			43.33	1.17	2.71				
	<i>Viper aspis francisredi</i>	f	2.91	43.40	1.20	2.76				
	<i>Daboia palestinae</i>	g		43.94	1.17	2.66				
	<i>Daboia russelii</i>		2.07							
	<i>Echis coloratus</i>	g		42.53	0.92	2.16				
Crocodylia	Crocodylidae	<i>Crocodylus niloticus</i>		3.40	48.44	0.85	1.75			
		<i>Alligator mississippiensis</i>	h	2.58	48.56	0.96	1.97			
	Alligatoridae					48.50	0.90			

⁽¹⁾ Taxonomy is from www.embl-heidelberg.de/~uetz/LivingReptiles.html. The underlined samples refer to reptiles only used for c-values.

⁽²⁾ Sources: (a) Romano Desensi, Italy; (b) Jon Costanzo, Department of Zoology, Miami University, Oxford OH 45056, USA (c) Fulvio Maffucci and Flegra Bentivegna, Aquariology Laboratory, SZN, Naples, Italy; (d) Rosaria Scudiero and Elio Parisi, IBP-CNR, Naples Italy; (e) Giorgio Bernardi, Laboratory of Molecular Evolution, SZN, Naples, Italy; (f) Gaetano Odierna, Università degli Studi di Napoli Federico II, Naples, Italy; (g) Dusan Kordis, Department of Biochemistry and Molecular Biology, Jozef Stefan Institute, Ljubljana, Slovenia; (h) Axel Janke, Department of Genetics, Division of Evolutionary Molecular Systematics, University of Lund, Lund, Sweden. *Testudo graeca* was analyzed by Jabbari et al. (1997). 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: *Anguis fragilis*, *Boa constrictor*, *Natrix tessellata*, *Podarcis muralis*, *Chrysemis picta*, *Trachemys scripta elegans*, *Testudo hermanni*.

⁽⁴⁾ R is the 5mC/GC ratio multiplied by 10².

⁽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 each other (indicated by thin vertical bars).

Hughes et al., 2002; see also Bernardi, 2004) and karyotypic evolutionary rates (Olmo et al., 2002; Olmo, 2005). A close look at the data of Fig. 1 reveals a number of interesting features. In terms of both high GC and low methylation level, the extreme case is that of crocodiles. Turtles are characterized by rather high GC levels and low methylation. Finally the squamates show the widest spread in both GC and methylation.

4.2. Crocodiles

Concerning body temperature, it has been observed that *Alligator mississippiensis* are mostly active in the 32–35 °C range, which is the most frequently recorded temperature. Indeed, physiologists demonstrated that crocodylians possess a high degree of thermoregulation (Smith, 1979; Seebacher et al., 1999) and are able to produce sufficient metabolic heat to elevate their body temperature above water temperature. A recent paper also stresses the similarities between the heart of

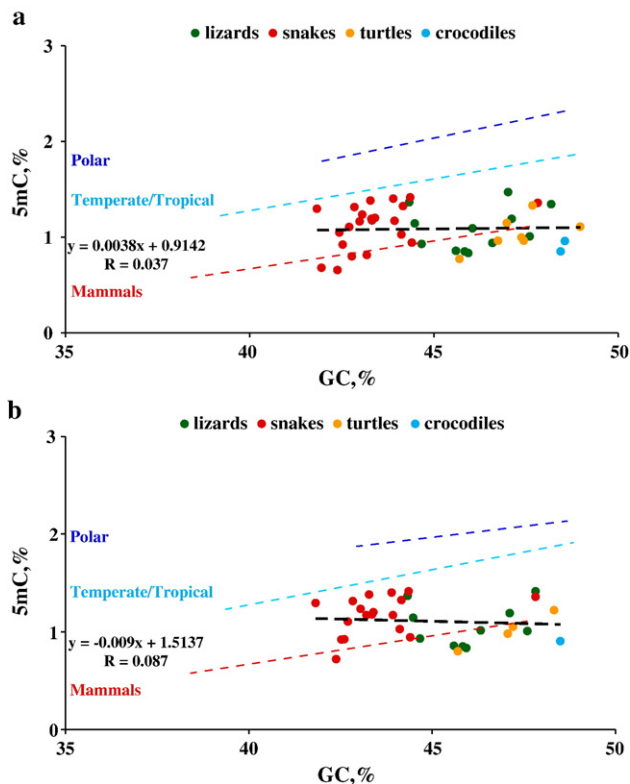


Fig. 1. a. Plot of 5mC levels against GC levels for the genomes of lizards (green circles), snakes (red circles), turtles (yellow circles) and crocodiles (blue circles). Values are listed in Table 1, column Species. The regression line, correlation coefficient (R) and equation for reptiles are shown (black line). Regression lines for polar fishes (dark blue line), temperate/tropical fishes (light blue line; Varriale and Bernardi, 2006-this issue), and mammals (red line; data from Jabbari et al., 1997, and from Varriale and Bernardi, in preparation) are also given as references. 1b. Plot of average 5mC levels against average GC levels for families/genera (listed in Table 1, column Families/genera). 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. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

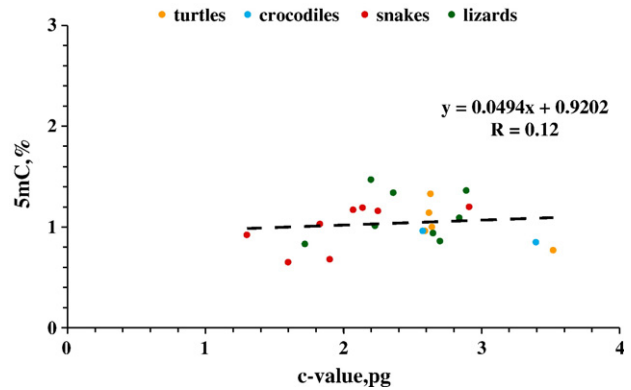


Fig. 2. Plot of 5mC against genome size for lizards (green circles), snakes (red circles), turtles (yellow circles) and crocodiles (blue circles). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

crocodylians and endotherms, showing the possibility of an endothermic ancestor of modern alligators and crocodiles (Summers, 2005).

4.3. Testudines

The case of a terrestrial turtle, *T. graeca*, was previously investigated (Thiery et al., 1976; Bernardi and Bernardi, 1990; Aïssani and Bernardi, 1991a). Its genome, already shown to be characterized by a remarkable compositional heterogeneity, exhibited one of the two lowest 5mC levels among all reptiles studied here. This may be explained by the rather high body temperature reached by some terrestrial turtles. Aquatic turtles also display a very low 5mC level which could possibly be explained by the high body temperature that they can reach. Examples are reported for a number of species and families. Freshwater turtles of the family *Emydidae* (such as *Trachemys scripta elegans* and *Chrysemys picta* analyzed here) achieve body temperatures significantly above both air and water temperatures when basking (Boyer, 1965; and references therein) This is a frequent behavioral way of heating up, found also for *Chelydra serpentina*, often observed in warm shallow water (Brattstrom, 1965). The two marine turtles that

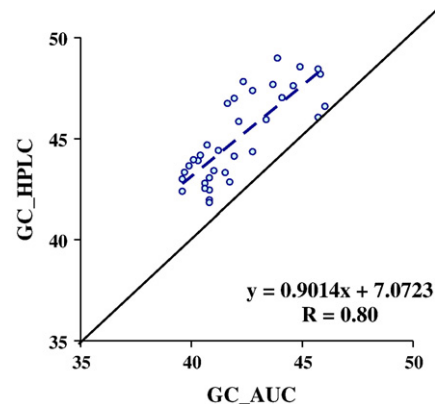


Fig. 3. Correlation between genomic GC levels as obtained by nucleoside analysis (HPLC) and by analytical ultracentrifugation (AUC) in reptiles.

we have analyzed in this study (*Caretta caretta* and *Chelonia mydas*) live in relatively warm waters. Moreover, body temperatures are often 1–3 °C above water temperatures and it can be assumed this is most likely the result of the accumulation of metabolically produced heat (Mrosovsky and Pritchard, 1971).

4.4. Squamata

Iguanidae represent a family in which many species, such as *Iguana iguana*, are diurnal baskers which are able to maintain relatively high and constant activity temperatures, within the range of 30–35 °C (Avery, 1982). A similar situation can be observed for many species belonging to the family *Agamidae* and *Lacertidae*, which are also diurnal baskers, but not for all species of *Anguidae* and *Chamaeleonidae*.

In snakes, the thermal relationship with the environment tends to be less precise than those of diurnal lizards because of their larger surface-to-volume ratios, conduction of heat and passive thermoregulation are relatively more important than basking in solar radiation. The elongate shape also creates problems of definition and measurement because temperature may vary considerably in different parts of the body.

4.5. DNA methylation levels of reptiles: intra-ordinal differences

The reason why the values of Table 1 for some species of reptiles were not averaged per families, as done in several cases for fishes (Varriale and Bernardi, 2006-this issue), is that significant differences were found not only within orders, but also within families and even within genera. Indeed, both GC and 5mC values were significantly different for the two *Python* and the four *Elaphe* species. While this may be due, at least in part, to the presence of satellite DNAs, the possible role of differences in body temperature cannot be ruled out.

5. Conclusion

This report, like the preceding one (Varriale and Bernardi, 2006-this issue), tried to connect biological and molecular data, highlighting the importance of temperature as an environmental factor able to act on genomes from ectotherms through natural selection. We focused here on the characteristics of life style of reptiles studied by ecologists and physiologists to find biological explanations for the results that we obtained at the genomic level. Physiological studies show that different reptiles follow different ways for increasing or decreasing body temperature, and can tolerate different ranges of temperatures. This biological diversity is reflected in genomic diversity. The results obtained in the present work provide a proof of such diversity. If the compositional transitions that took place in the genomes of vertebrates at the emergence of mammals and birds had appeared in the common ancestor of tetrapods, as proposed by Hughes et al. (1999), or in the common ancestor of reptiles and mammals, as proposed by Duret et al. (2002), we should have found identical or similar 5mC levels among reptiles (a situation found in the genomes of mammals and birds), whereas clearly this is not the case.

The new data support a strong influence of environment on reptilian genomes and are in full contrast with the conclusion of Duret et al. (2002) who proposed an “amniote hypothesis” about the GC increase transition in vertebrate genomes. This hypothesis might still have been compatible with the data obtained for reptiles by Jabbari et al. (1997), because the reptilian points were closer to points from warm-blooded vertebrates than to those from cold blooded. However, at that time only a small sample had been investigated, and it was not possible to reach any final conclusion. We found, now, that while DNA methylation is low in the case of turtles and crocodylians (which are also characterized by a relatively large degree of DNA heterogeneity; Aïssani and Bernardi, 1991a,b), namely in the case in which body temperature is relatively high, lizards and, more so, snakes, clearly show a strong variability in DNA methylation covering the whole 5mC range comprised between fish/amphibians and mammals/birds. This variability could be due to the different “average” body temperatures of these reptiles. One should not forget that reptiles represent a group different in many aspects both from thermoconformer ectotherms, such as fishes and amphibians, and from endotherms, such as mammals and birds. Reflecting their physiological peculiarity, their genomes display different DNA methylation levels that are intermediate between these two groups.

Interestingly, a positive correlation between body temperatures and GC level in some desert lizards as compared with some snakes was reported by Olmo (2003), who hypothesized that a higher GC level could help DNA to resist the destabilizing effect of temperature and genetic damage, which could derive from an increased metabolic rate correlated with an increased production of free radicals.

A final remark is that 5mC and CpG levels appear to integrate the effects of temperature and time of exposure to temperature. If this is true, more complex changes in the genome structure, such as the formation of GC-rich isochores, should be better correlated with DNA methylation rather than with body temperature. In the case of large-scale comparisons, such as that of fishes/amphibians with mammals/birds, temperature is a good reference because the exposure times under considerations are very large. In the case of reptiles, however, the formation of GC-rich isochores, the compositional heterogeneity and the compositional asymmetry should be related to 5mC and CpG levels, body temperature having a low significance in the absence of information of time of exposure.

Acknowledgements

We thank Ettore Olmo, Giacomo Bernardi and Kamel Jabbari for helpful comments. We are grateful to all colleagues who helped us with the gift of DNA or tissue samples, 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 and Sandra Hochscheid for references on reptilian physiology.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at [doi:10.1016/j.gene.2006.05.034](https://doi.org/10.1016/j.gene.2006.05.034).

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.
- Avery, R.A., 1982. Field studies of body temperatures and thermoregulation. In: Gans, C., Pough, F.H. (Eds.), *Biology of the Reptilia*, vol. 12 C. Academic Press, pp. 93–166.
- Bernardi, G., 2004. Structural and evolutionary genomics. *Natural Selection in Genome Evolution*. Elsevier, Amsterdam. Reprinted in 2005.
- Bernardi, G., Bernardi, G., 1990. Compositional patterns in the nuclear genome of cold-blooded vertebrates. *J. Mol. Evol.* 31, 265–281.
- Boyer, D.R., 1965. Ecology of the basking habit in turtles. *Ecology* 46, 99–118.
- Brattstrom, B.H., 1965. Body temperatures of reptiles. *Am. Midland Nat.* 73, 376–422.
- Duret, L., Semon, M., Piganeau, G., Mouchiroud, D., Galtier, N., 2002. Vanishing GC-rich isochores in mammalian genomes. *Genetics* 162, 1837–1847.
- Halliday, T., Adler, K., 2002. *The New Encyclopedia of Reptiles and Amphibians*. Oxford University Press, Oxford.
- Heatwole, H., Taylor, J., 1987. *Ecology of Reptiles*. Surrey Beatty and Sons, Chipping Norton.
- Huey, R.B., 1982. Temperature, physiology, and the ecology of reptiles. In: Gans, C., Pough, F.H. (Eds.), *Biology of the Reptilia*, vol. 12 C. Academic Press, pp. 25–85.
- Hughes, S., Zelus, D., Mouchiroud, D., 1999. Warm-blooded isochore structure in Nile crocodile and turtle. *Mol. Biol. Evol.* 16, 1521–1527.
- Hughes, S., Clay, O., Bernardi, G., 2002. Compositional patterns in reptilian genomes. *Gene* 295, 323–329.
- 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.
- Mrosovsky, N., Pritchard, P.C.H., 1971. Body temperatures of *Derموcheyls coriacea* and other sea turtles. *Copeia* 4, 624–631.
- Olmo, E., 2003. Reptiles: a group of transition in the evolution of genome size and of nucleotypic effect. *Cytogenet. Genome Res.* 101, 166–171.
- Olmo, E., 2005. Rate of chromosome changes and speciation in reptiles. *Genetica* 125, 185–203.
- Olmo, E., Capriglione, T., Odierna, G., 2002. Different genomic evolutionary rates in the various reptile lineages. *Gene* 295, 317–321.
- 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.
- Seebacher, F., Grigg, G.C., Beard, L.A., 1999. Crocodiles as dinosaurs: behavioral thermoregulation in very large ectotherms leads to high and stable body temperatures. *J. Exp. Biol.* 202, 77–86.
- Smith, E.N., 1979. Behavioral and physiological thermoregulation of crocodilians. *Am. Zool.* 19, 239–247.
- Summers, A.P., 2005. Warm-hearted crocs. *Nature* 434, 833–834.
- 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., 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-this issue. DNA methylation and body temperatures in fishes. *Gene*. [doi:10.1016/j.gene.2006.05.031](https://doi.org/10.1016/j.gene.2006.05.031).