Marine Genomics xxx (2015) xxx–xxx

Contents lists available at ScienceDirect

Marine Genomics

journal homepage: www.elsevier.com/locate/margen

Review

An overview on genome organization of marine organisms

Maria Costantini *

Department of Biology and Evolution of Marine Organisms, Stazione Zoologica Anton Dohrn, Villa Comunale, 80121 Naples, Italy

ARTICLE INFO

Article history:
Received 18 February 2015
Received in revised form 17 March 2015
Accepted 17 March 2015
Available online xxxx

Keywords:
Compartmentalization
Genome
Heterogeneity
Isochores
Marine organisms

ABSTRACT

In this review we will concentrate on some general genome features of marine organisms and their evolution, ranging from vertebrate to invertebrates until unicellular organisms. Before genome sequencing, the ultracentrifugation in CsCl led to high resolution of mammalian DNA (without seeing at the sequence). The analytical profile of human DNA showed that the vertebrate genome is a mosaic of isochores, typically megabase-size DNA segments that belong in a small number of families characterized by different GC levels. The recent availability of a number of fully sequenced genomes allowed mapping very precisely the isochores, based on DNA sequences. Since isochores are tightly linked to biological properties such as gene density, replication timing and recombination, the new level of detail provided by the isochore map helped the understanding of genome structure, function and evolution. This led the current level of knowledge and to further insights.

Contents

1. Introduction

Well before genome sequencing, ultracentrifugation in Cs_2SO_4 density gradients in the presence of sequence-specific DNA ligands (e.g., Ag+) was shown to lead to a high resolution of mammalian DNAs according to base composition (Corneo et al., 1968). These findings opened a new inroad in the study of the organization of eukaryotic genomes, superseding DNA reassociation kinetics (Britten and Kohne, 1968), which was based on the separation of single- and double-stranded DNAs on hydroxyapatite (Bernardi, 1965). According to the new density gradient approach the genomes of warm-blooded vertebrates were characterized by a striking long-range compositional heterogeneity (neglecting satellite DNAs; Filipski et al., 1973; Macaya et al., 1976; Thiery et al., 1976). Indeed, these genomes are mosaics of isochores, long (major than 300 kb), compositionally fairly homogeneous regions that belong to a small number of families characterized by different average GC levels (Macaya et al., 1976). A quarter of a century after the original studies that had defined the approximate sizes and compositions of isochores as well as the compositions and relative amounts of isochore families, it was reported that isochores could not be identified in the draft sequence of the human genome (Lander et al., 2001), starting a debate that is still ongoing. The different computational approaches used to disprove or redefine isochores (Eyre-Walker and Hurst, 2001; Häring and Kypr, 2001; Lander et al., 2001; Nekrutenko and Li, 2001; Cohen et al., 2005) were, however, shown to be inadequate (Bernardi, 2001; Clay and Bernardi, 2001a,b, 2005; Li, 2002; Oliver et al., 2002, 2004; Li et al., 2003), even if some of them led to a partial identification of isochores. This debate prompted us to map the isochores, as originally defined (Macaya et al., 1976), in the finished sequence of the human genome (International Human Genome Sequencing Consortium, 2004). Average GC levels were, therefore, assessed over long DNA stretches (>200 kb), while GC variation was estimated by measuring standard deviations of GC over such stretches using a 100-kb moving window (Costantini et al., 2006). In fact, we demonstrated that if one scans the GC profiles of human

* Tel.: +39 081 5833315; fax: +39 081 7641355.
E-mail address: maria.costantini@szn.it.

http://dx.doi.org/10.1016/j.margen.2015.03.015
1874-7787/

chromosomes from any starting point using a fixed window of 100 kb, one can find a mosaic of sequences ranging from 200 kb to several megabases that are characterized by different GC levels and by a remarkable compositional homogeneity. Our results were in agreement with previous results obtained by equilibrium sedimentation and confirm the existence of five isochore families, but they go much farther in that they directly identify and map isochores on chromosomes, thus leading to a resolution of >3000 chromosomal bands. According to our estimate, the total number of isochores bands in the human genome is 3159, a number very close to the maximum number (3000) mentioned by Yunis et al. (1977) for chromosomal bands at the highest resolution. The isochore pattern is, expectedly, different from chromosome to chromosome. However, when isochores were pooled in bins of 1% GC (Fig. 1), isochore families stand out. This was evident for the GC-poor isochore families L1, L2, and H1, but also visible for the GC-rich H2 and H3 families, which are present in small amounts in the genome. The relative amounts of DNA in isochore families were 19%, 37%, 31%, 11%, and 3% for L1, L2, H1, H2, and H3 isochores, respectively, again in fair agreement with previous results (Macaya et al., 1976; Cuny et al., 1981).

Localizing genes in separate isochores led to the discovery of an unexpected and strikingly nonrandom distribution of genes (Mouchiroud et al., 1991; Zoubak et al., 1996), which were found in two "gene spaces." The "genome core," composed of the isochore families H2 and H3, comprises more than half of the genes even though they represent only 15% of the genome, whereas the "genome desert" (the isochore families L1, L2, and H1) is made up of large expanses with low and often extremely low gene densities (Fig. 2). These two gene spaces are characterized by several different properties (for review, see Bernardi, 2004), the most remarkable ones being the correlations of isochore families not only with gene density but also with replication timing (Costantini and Bernardi, 2008), recombination, and location and chromatin structure in interphase nuclei, chromatin being "open" in the genome core and "closed" in the genome desert (Saccone et al., 2002). Finally, the isochore families corresponded to peaks in "gene landscapes" (Cruvellier et al., 2004), formed by the distribution of coding sequences according to the GC levels of second and third codon positions (GC_2 and GC_3).

In this review we will deal with some general features on the genome compartmentalization of marine organisms, approaching the general problem of the organization and evolution of genomes at the sequence level (ranging from vertebrates to invertebrates until unicellular organisms).

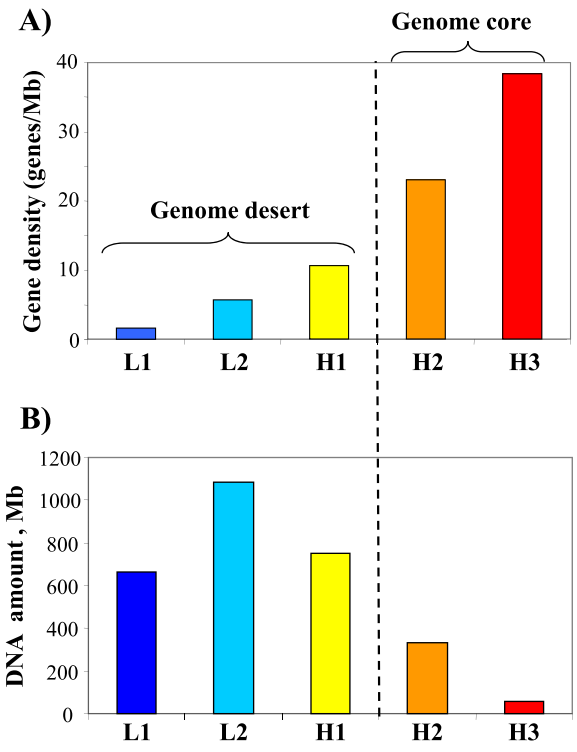

Fig. 2. The histograms represent in A) the gene density (genes/megabases) and in B) DNA amount in megabases of the five isochore families, calculated on the basis of DNA sequences of the human genome.

2. Vertebrates

Investigations published 30 years ago led to the discovery of phylogenetic differences at the macromolecular level in eukaryotic genomes (Thiery et al., 1976). In particular, a major compositional difference was found between the genomes of warm- and cold-blooded vertebrates. While the former were very heterogeneous, as first observed in the "main band" (satellite DNAs are neglected here) of calf DNA (Filipski et al., 1973), the latter were characterized by a much lower compositional heterogeneity. Previous work from our laboratory showed that (i) vertebrate genomes are mosaics of isochores, typically megabase-size DNA segments that are fairly homogeneous in base

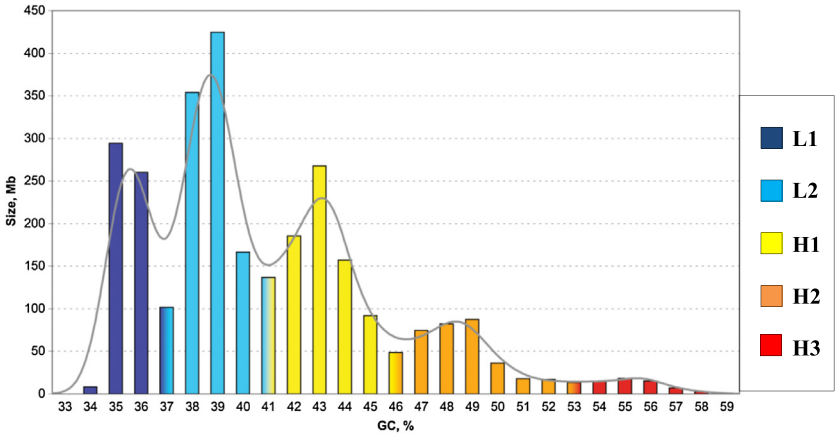

Fig. 1. Distribution of isochores in human genome, according to GC levels. The histogram shows the distribution (by weight) of isochores as pooled in bins of 1% GC. Colors represent the five isochore families. Values at minima (histogram bars with mixed colors) were split between the two neighboring families. The Gaussian profile shows the distribution of isochores as estimated directly by the "density" function in R (bandwidth 0.7% GC) (Silverman, 1986). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
(Modified from Costantini et al., 2006)

Please cite this article as: Costantini, M., An overview on genome organization of marine organisms, Mar. Genomics (2015), http://dx.doi.org/10.1016/j.margen.2015.03.015