Genome Organization and Chromosome Architecture

GIORGIO BERNARDI^{1,2}

¹*Science Department, Roma Tre University, 00146 Rome, Italy*

²*Stazione Zoologica Anton Dohrn, 80121 Naples, Italy*

Correspondence: gbernardi@uniroma3.it

How the same DNA sequences can function in the three-dimensional architecture of interphase nucleus, fold in the very compact structure of metaphase chromosomes, and go precisely back to the original interphase architecture in the following cell cycle remains an unresolved question to this day. The solution to this question presented here rests on the correlations that were found to hold between the isochore organization of the genome and the architecture of chromosomes from interphase to metaphase. The key points are the following: (1) The transition from the looped domains and subdomains of interphase chromatin to the 30-nm fiber loops of early prophase chromosomes goes through their unfolding into an extended chromatin structure (probably a 10-nm “beads-on-a-string” structure); (2) the architectural proteins of interphase chromatin, such as CTCF and cohesin subunits, are retained in mitosis and are part of the discontinuous protein scaffold of mitotic chromosomes; and (3) the conservation of the link between architectural proteins and their binding sites on DNA through the cell cycle explains the reversibility of the interphase to mitosis process and the “mitotic memory” of interphase architecture.

The organization of the eukaryotic genome and the architecture of eukaryotic chromosome were the subjects of several Cold Spring Harbor Symposia on Quantitative Biology over the past 60 years. This can be understood if one considers the biological relevance of these very difficult problems. Here I will propose a solution that is based on many years of research on genome and chromosome organization as well as on current investigations on interphase chromatin from many laboratories. As an introduction, I will briefly describe the approach used to study genome organization, some basic results, and then move on to the correlations between genome organization and chromosome architecture through the cell cycle.

The approach used, compositional genomics, is a minimalist, top-down strategy based on short-sequence frequencies in human DNA and/or GC levels as a proxy. In fact, because short sequences determine the fine structure of the double helix and the interactions of DNA with proteins (e.g., histones to build nucleosomes, transcription factors to interact with regulatory sequences), compositional genomics is based on genome structure and function. The approach initially relied on the fractionation of DNA in Cs₂SO₄ density gradient ultracentrifugation as run in the presence of sequence-specific ligands, such as Ag⁺ (Corneo et al. 1968), and on DNA sequences as soon as they became available in the late 1990s.

Compositional genomics, in its original ultracentrifugation approach, led to an unexpected discovery, namely the compositional compartmentalization of a mammalian genome (Filipinski et al. 1973). Indeed, DNA molecules 10–20 kb in size could be fractionated into a small number of families characterized by different base composition and different short-sequence frequencies (Devillers-

Thiery 1974). This compartmentalization was then also found in all the other eukaryotic genomes tested (Thiery et al. 1976). Moreover, the 10–20-kb DNA fragments were shown to derive by degradation during preparation from much larger, compositionally fairly homogeneous DNA regions characterized by different compositions, the isochores (Macaya et al. 1976), the basic elements of the compositional mosaic of the eukaryotic genome (Bernardi et al. 1985).

Isochores can be easily visualized as regions at least 200 kb in size by scanning the compositional profiles of chromosomes with a nonoverlapping window of 100 kb the compositional profiles of chromosomes (see Fig. 1). If isochores from the human genome are binned according to their GC levels, five families, L1, L2, H1, H2, and H3, ranging from 33% to 59% GC appear in the case of the human genome (see Table 1; Fig. 2). An analysis of isochores revealed the first rule of what was called the genomic code (Bernardi 1990; see also Bernardi and Bernardi 1991), a rule concerning the existence of strong correlations between the GC levels of contiguous coding and noncoding sequences as well as of the three codon positions of coding sequences (for review, see Bernardi 2004).

The second rule of the genomic code was that the GC levels of isochores are correlated with all tested structural/functional properties of the genome. Indeed, the L1, L2, and H1 families are different from the H2 and H3 families in many properties, such as gene density, replication timing, open/closed chromatin (see Table 2 for a longer, yet nonexhaustive list). In fact, the gene-rich compartment, called the genome core, and the gene-poor one, the genome desert, are characterized by opposite properties.

Figure 1. (A) Compositional profile of human chromosome 21 as seen through nonoverlapping 100-kb windows (Pavlicek et al. 2002). DNA stretches from isochore families L1 to H3 are represented here in different colors—deep blue, light blue, yellow, orange, and red, respectively. The ordinate values are the minima GC values (valleys) between isochore families. The red horizontal line at 41% GC separates the two (GC-poor and GC-rich) genome compartments. (B) Isochore profile of chromosome 21 using the IsoSegmenter program (Cozzi et al. 2015).

INTERPHASE CHROMOSOMES AND ISOCHORES

It is well established that (1) chromosomes occupy distinct territories in the interphase nucleus of eukaryotic cells (Cremer and Cremer 2010); (2) GC-rich, early replicating, transcriptionally active chromatin regions are located in the nuclear interior (Sadoni et al. 1999); and (3) the gene-richest regions display a much more spread-out, open conformation compared with the closed one of the gene-poorest regions (Saccone et al. 2002; Gilbert et al. 2004). Moreover, superresolution microscopy established the existence of a higher-order chromatin organization, the ~ 1 -Mb chromatin domains that may comprise smaller subdomains (Cremer and Cremer 2010).

These results were confirmed and expanded through other approaches. Indeed, the development of chromosome conformation capture (3C) technology (Dekker et al. 2002) and its variants led to a new insight into the three-dimensional chromatin organization of the interphase nucleus. The spatial proximity maps provided evidence for numerous domains that fall into two subchromosomal compartments, A and B, characterized by open and closed chromatin, respectively (Lieberman-Aiden et al. 2009).

At a lower size range, it appears that the three-dimensional structure of chromatin at interphase begins to be well understood as the result of investigations on mammalian cells and *Drosophila* concerning (1) the “lamina-associated domains” (LADs) and their borders (Guelen et al. 2008; Meuleman et al. 2013); (2) the different “chromatin states” of mammalian cells (Ernst and Kellis 2010; Julienne et al. 2013); (3) the different “chromatin types” of *Drosophila* (Filion et al. 2010); (4) the “topo-

logical domains” and the “topologically associating domains” (TADs) and their boundaries (Dixon et al. 2012; Nora et al. 2012), as well as the corresponding “physical domains” of *Drosophila* and their borders (Hou et al. 2012; Sexton et al. 2012); and (5) the subcompartments and the contact domains defined by the interaction patterns detected by in situ Hi-C (Rao et al. 2014; for reviews, see Ong and Corces 2014; Dekker and Heard 2015; Sexton and Cavalli 2015).

The architecture of interphase chromatin may be schematically visualized, at least for the purpose of this work, as a set of looped domains and boundaries, the TADs. In turn, looped domains essentially consist of subdomains, most of which are anchored by CTCF and by cohesin subunits (as shown in Fig. 3A).

Although the focus of the investigations just mentioned concerned the connections between chromatin structure and gene expression/regulation, here the main interest

Table 1. Isochore families in the human genome

	Amount (%)	GC maxima	GC minima	GC ranges
L1	19	35.5		33–37
L2	36	38.7	37	37–41
H1	31	43.0	41	41–46
H2	86 11	48.5	46	46–53
H3	3 14	55.0	53	53–59

Compiled from data of Costantini et al. 2006.
Values in bold are total amounts of L1 + L2 + H1 and H2 + H3 families, the genome desert, and the genome core.

Figure 2. Isochore families. The histogram displays the isochores from the human genome as pooled in bins of 1% GC. The Gaussian profile shows the distribution of isochore families, which are represented in different colors as in Figure 1. Gene densities (and all other structural and functional properties tested; see Table 2) define a genome desert, isochore families L1, L2, H1, and a genome core, isochore families H2, H3 (separated by a vertical broken red line). (From Bernardi 2015.)

was on the correlations between chromatin structure and compositional genome features in view of solving the mystery of the changes of chromosome architecture during the cell cycle.

Table 2. Structural and functional properties of the genome core versus the genome desert

GC level	+ ¹	GC heterogeneity	+ ^{17a}
LINE density	− ²	Intron, UTR size	− ¹⁸
SINE density	+ ³	Housekeeping genes	+ ¹
Gene density	+ ^{4,5}	Developmental genes	− ^{19,20}
Chromatin	Open ^{6,7}	Reverse bands	+ ^{17,21–23}
Gene expression	+ ^{8–11}	Nuclear location	Central ^{6,7,24}
4-strand	+ ^{12,13}	Translocations	+ ²⁵
B → Z	+ ¹³	Breakpoints	+ ²⁶
Bendability	+ ¹³	Fragile sites	+ ²⁷
Curvature	− ¹³	Proviral integration	+ ^{1,28}
GC-rich trinucleotides	+ ^{14,15}	Insertions/deletions	+ ²⁹
CpG, mC	+ ¹⁶	Recombination	+ ³⁰
CpG islands	+ ¹⁶	Point mutations	+ ³¹
Isochore size	− ¹⁷	Replication	Early ^{32–35}

In general, the properties of the genome core are just opposite to those of the genome desert. +/− signs indicate positive/negative differences of the properties of the genome core compared with those of the genome desert. Nucleosome positioning patterns (Frenkel et al. 2011) and chromatin states (Ernst and Kellis 2010; Julienne et al. 2013) also differ in the two genome spaces.

^aSee Table 1.
Sources: ¹Bernardi 2004; ²Meunier-Rotival et al. 1982; ³Soriano et al. 1983; ⁴Mouchiroud et al. 1991; ⁵Zoubak et al. 1996; ⁶Saccone et al. 2002; ⁷Gilbert et al. 2004; ⁸Caron et al. 2001; ⁹D’Onofrio 2002; ¹⁰Konu and Li 2002; ¹¹Versteeg et al. 2003; ¹²Biffi et al. 2013; ¹³Vinogradov 2003; ¹⁴Costantini and Bernardi 2008a; ¹⁵Arhondakis et al. 2011; ¹⁶Varriale and Bernardi 2009; ¹⁷Costantini et al. 2006; ¹⁸Duret et al. 1995; ¹⁹Kikuta et al. 2007; ²⁰Jabbari and Nurnberg 2016; ²¹Saccone et al. 1992; ²²Saccone et al. 2001; ²³Costantini et al. 2007; ²⁴Sadoni et al. 1999; ²⁵Fisher et al. 2005; ²⁶Lemaitre et al. 2009; ²⁷Berthelot et al. 2015; ²⁸Rynditch et al. 1998; ²⁹Costantini and Bernardi 2009; ³⁰Fullerton et al. 2001; ³¹Hodgkinson et al. 2012; ³²Watanabe et al. 2002; ³³Schmiegner et al. 2007; ³⁴Costantini and Bernardi 2008b; ³⁵Pope et al. 2014.

The results mentioned above are compared in Table 3 with the properties associated with the genome core and the genome desert. (In fact, several properties of the genome core listed in Table 2 could be added to the right column of Table 3.) This comparison leads to the conclusion that the properties of compartments and subcompartments, as well as those of chromatin domains and boundaries, match those of the isochores from the genome desert and the genome core, respectively, despite not always completely overlapping with each other because they are defined on the basis of different approaches. Indeed, although domain boundaries generally correspond to GC-rich isochores anchored by architectural proteins, such as CTCF and cohesin (present as a ring structure or as cohesin subunits), looped domains correspond to GC-poor isochores.

Interestingly, the interphase chromatin architecture is conserved across mammalian species (Dixon et al. 2012; Rao et al. 2014) and in syntenic regions (Vietri Rudan et al. 2015). These observations parallel the conservation of compositional landscapes of syntenic regions (Bernardi 2007) and of isochore families in mammalian genomes (Costantini et al. 2009). These findings indicate precise correlations between interphase chromatin architecture and the corresponding compositional landscapes and suggest that the loops and boundaries of interphase chromatin are correlated with the GC levels of isochores (K Jabbari and G Bernardi, in prep.).

INTERPHASE CHROMATIN TO PROPHASE BANDS

To understand the complex connection between isochores and chromosomal bands, a new approach was developed, compositional mapping (Bernardi 1989, 1995, 2004). This approach was initially based on assessments of GC levels around genome landmarks (e.g., genes localized on the physical map) of metaphase chromosomes

Figure 3. (A) A scheme of an interphase chromatin loop (a topologically associating domain, TAD, with three subdomains in this figure). The DNA framework of the loop is a large GC-poor isochore. The loop is closed by anchors (chromatin boundaries) that interact with two architectural proteins, CTCF (boxes) and cohesin (green oval). A number of subdomains have their loops anchored by CTCF (boxes) (see text). (B) Opening of the three-dimensional architecture of the domains and subdomains in a linear chromatin structure, possibly in a “beads-on-a string,” 10-nm conformation. Architectural proteins are visualized as still linked to their binding sites (see text). (C,D) Folding of the open structure into 30-nm fiber loops anchored by the architectural proteins and compaction into three early prophase, single-isochores bands R-G-R, the central one being a multiple-loop band, the flanking ones single-loop bands. (E) Coalescence of single-isochores bands into multiple-isochores bands. In the example shown, the R-G-R single-isochores bands coalesce into an R band because of a “majority rule” (2R vs. 1G). Architectural proteins contribute to form a discontinuous protein scaffold of the chromosome (see text). (From Bernardi 2015.)

(see Bernardi 2004), then on in situ hybridization of DNA from L1 and H3 isochores on metaphase and prometaphase chromosomes (Saccone et al. 1992, 2001), and finally on human genome sequences (Costantini et al. 2006, 2007). The latter results provided detailed information on the sizes and GC levels of isochores, of prometaphase bands, and of metaphase bands for all human chromosomes. Here, a new analysis of the supplementary data (approximately 120 pages) of these investigations (Costantini et al. 2006, 2007) was done, leading to new results and to a model for the critical transition from interphase chromatin to early prophase bands, which are presented below.

1. At the beginning of mitosis the three-dimensional organization of interphase chromatin disappears (Nau-mova et al. 2013), as expected, and is replaced in early prophase of human chromosomes by more than 3000 bands (Yunis 1981). This number approximately matches the number of isochores, about 3200 (Costantini et al. 2006). This preliminary indication that early prophase bands may correspond to individual isochores is now definitely supported by the observation (Fig. 4) that single-isochores bands represent ~8% of

metaphase bands, ~25% of prometaphase bands, and ~50% of mid-prophase bands. Indeed, the three relative amounts just quoted indicate, by extrapolation, that single-isochores bands represent the totality of early prophase bands when the number of the latter is approximately 3400, a value close to the total number of isochores (see Fig. 4).

2. The very early decrease in band numbers in early prophase (see Fig. 4) indicates a coalescence of the 30-nm fiber loops (a conventional definition because of a current debate) that form the bands. Obviously, the transition from the loops of interphase chromatin to the 30-nm fiber loops of early prophase needs a transient intermediary step. This can be visualized as the opening up of the three-dimensional structure of interphase chromatin (Fig. 3A) into an extended, essentially two-dimensional chromatin configuration, probably a 10-nm “beads-on-a-string” structure (the “beads” corresponding to nucleosomes), in which GC-rich and GC-poor isochores alternate (Fig. 3B). In the example of Figure 3, the anchors of an interphase chromatin domain (Fig. 3A) are opened up into two GC-rich isochores flanking a GC-poor isochore,

Table 3. Isochores and interphase chromatin

ISOCHORES ¹	
Genome desert (L1, L2, H1 families): gene-poor, LINES, closed chromatin, late replicating, 86% of the genome	Genome core (H2, H3 families): gene-rich, SINES, open chromatin, early replicating, CpG islands, HK genes (av. size 0.9 Mb)
COMPARTMENTS ²	
B: closed chromatin	A: open chromatin
DOMAINS AND BORDERS/BOUNDARIES	
Domains LADs (lamina-associated domains) ³ , gene-poor, H3K27me3 cLADs (constitutive LADs), GC-poor ⁴ Topological domains ⁵ , 91% of the genome (av. size 0.88 Mb) TADs (topologically associating domains) ⁶ Physical domains ⁷ , repressive marks	Borders CTCF, promoters, CpG islands, gene-rich, H3K4me2 ciLADs (constitutive inter-LADs), GC-rich Boundaries+“unorganized chromatin” (av. size 0.56 Mb) tRNAs, SINES, TSS Housekeeping genes Boundaries, active marks
CHROMATIN STATES ⁸	
Large-scale repressed Repeat-associated	Promoter-associated Transcription-associated
CHROMATIN TYPES ^{a9}	
Blue/green: heterochromatin Polycomb/H3K27me3, HPI/H3K9me2 Black: repressive chromatin all late replicating	Red, yellow: active chromatin H3K4me2, H3K79me3, both early replicating
SUBCOMPARTMENTS ^{b10}	
B1, H3K27me3; facultative heterochromatin; replication→middle S; B2,B3, replication→end S phase	A1, A2, gene-rich, high expression H3K36me3, H3K79me2, H3K27ac, H3K4me1, both early replicating ^b

HK, housekeeping; TSS, transcription start sites.
^a*Drosophila*. Other data concern mammalian cells.
^bSee also Bernardi 2015.
Sources: ¹Bernardi 2004; see also Table 2; ²Lieberman-Aiden et al. 2009; ³Guelen et al. 2008; ⁴Meuleman et al. 2013; ⁵Dixon et al. 2012; ⁶Nora et al. 2012; ⁷Sexton et al. 2012; ⁸Ernst and Kellis 2010; ⁹Filion et al. 2010; ¹⁰Rao et al. 2014.

the cohesin ring, if present, also being opened into its subunits. This opening process also concerns the anchors of the subdomains and the corresponding architectural proteins, CTCF and cohesin subunits (Fig. 3B).

3. The folding process of the extended chromatin structures into the 30-nm fiber loops just described does not take place at random locations on chromosomes. Indeed, the alternance of the GC-rich and the GC-poor bands in early prophase indicates that the folding specifically involves GC-rich and GC-poor isochores, respectively. This raises a question concerning which signals demarcate the GC-rich and the GC-poor isochores of the extended chromatin structures. The answer proposed here is that the signals are the same architectural proteins that were demarcating the anchors of the loops—CTCF and cohesin subunits, the former still being associated with their binding sites in mitotic chromosomes. This demarcation seems to apply differentially to the architectural proteins located at domain boundaries and to those that are associated with subdomains (see Zuin et al. 2014). In the first case (see Fig. 3), the two GC-rich isochores of the domain boundary may be folded in single loops, whereas the GC-poor isochore of the chromatin domain is folded in multiple loops that originate from

the subdomains (see Fig. 3C,D). Interestingly, the initial formation of chromatin loops may correspond to the packing of the linear array of nucleosomes of early prophase into a double helix of two interwound helical stacks with a straight linker DNA between successive nucleosomes, as observed by cryo-electron microscopy in reconstituted chromatin (Song et al. 2014, Travers 2014; see Fig. 5).

Figure 4. The percentages of single-isochore bands are plotted against the total number of bands at metaphase (400 bands), prometaphase (850 bands), and mid-prophase (1700 bands) and extrapolated to 100% single-isochore bands. (From Bernardi 2015.)

Figure 5. A 30-nm fiber loop is presented as a chromatin double helix. (Modified from Song et al. 2014, with permission from AAAS.)

4. The single-isochore bands (whether single- or multiple-loop) start coalescing very early into multiple-isochore bands, as indicated by the early decrease of band numbers (see Fig. 3). This coalescence appears to follow a precise rule in that it involves an odd number of single-isochore bands (most often three; Fig. 3E), to keep the GC alternation of bands in the increasing number of multiple-isochore bands. The bands originating by the coalescence process just mentioned will be G or R bands according to a “majority rule,” namely according to the number of G or R bands in the coalesced bands (see Fig. 3E).
5. A key point of this model is that the architectural proteins, CTCF and cohesin subunits, are retained after the interphase to mitosis transition, like a number of other proteins (see the section Metaphase to Interphase). This model is interesting for three main reasons: (i) the 30-nm fiber loops of chromosomes were estimated to be in the 100-kb range, a range that is approached by the median size of all chromatin loops, 185 kb (Rao et al. 2014); and (ii) the band coalescence may be driven by the increasing interactions among the architectural proteins, that, in fact, contribute to form the protein scaffold of mitotic chromosomes; and (iii) the model explains the recovery of the original three-dimensional architecture of interphase chromatin at the exit from mitosis.

PROPHASE TO PROMETAPHASE AND METAPHASE BANDS

A detailed analysis of the banding process leading from prophase to prometaphase and to metaphase bands is presented elsewhere (Bernardi 2015), so here only a few important points will be mentioned.

1. The progressive compaction from prophase to metaphase bands can be visualized as due to the coales-

cence of consecutive 30-nm loops (in agreement with Naumova et al. 2013).

2. The process is possibly driven by the increasing interactions of the discontinuous scaffold of architectural and other proteins.
3. The present model of metaphase chromosomes is a “loops-on-a-scaffold” model (Paulson and Laemmli 1977) in which, however, the chromatin loops are linked to the scaffold by architectural proteins, themselves part of the scaffold, and not by AT queues (<35% GC), 0.7 kb to several kilobases in size (Saitoh and Laemmli 1994), that simply do not exist in the H1, H2, and H3 isochores that form the R bands (see Bernardi 2015).

METAPHASE TO INTERPHASE

Although the transition from the three-dimensional architecture of interphase chromosomes to the increasingly compact one of metaphase chromosomes can be (wrongly) visualized just as a stochastic folding process, the reverse transition from metaphase to interphase is much more difficult to explain in simple terms. The reason is that the architecture of the new interphase chromosomes is such as to allow a quick reactivation of the original cell-specific programs, which implies that the original chromatin loops and anchors are precisely reformed. In fact, recent investigations have shown how sensitive chromatin functions are to changes in interphase architecture (Deng et al. 2014; Chandra et al. 2015; Dixon et al. 2015).

The model presented in Figure 2 can, however, solve this problem if read in the reverse from mitotic to interphase chromosomes. Needless to say, this explanation relies on the idea that architectural proteins, such as CTCF and cohesin subunits, remain associated with chromosomes during mitosis. This idea is totally acceptable if one takes into consideration that not only H3K9me3, H3K27, and Polycomb group proteins but also a fraction of transcription factors and of chromatin binding proteins (including cohesin; Yanagida 2009) are retained in mitotic chromosomes (Egli et al. 2008; Blobel et al. 2009; Follmer et al. 2012; Kadauke et al. 2012; Deng et al. 2014; Zaret 2014). It is then reasonable to believe that the “mitotic memory” of the interphase architecture concerns the entire three-dimensional architecture of interphase chromosomes, which allows the same cell-specific expression programs of the mother cells to be achieved thanks to the conserved link between architectural proteins and the corresponding binding sites on DNA. In fact, at the proof stage, it was found that evidence for the retention of CTCF in mitotic chromosomes was available (Burke et al. 2005).

CONCLUSION

This investigation leads to two major conclusions. The first one concerns the explanation of the reversible chang-

es of chromosome architecture through the cell cycle. The second has to do with the connection between genome organization and chromosome architecture.

Our understanding of chromosome architecture at interphase has recently made remarkable advances, essentially thanks to the development of chromosome conformation capture (3C) and derived approaches. In contrast, the old mystery surrounding the transitions of chromosome architecture from interphase to mitosis and from mitosis to interphase at the beginning of the following cycle has been waiting for a solution for a long time.

This mystery has now been solved by using a strategy that took advantage of our previous knowledge of genome organization in a critical size range (hundreds to thousands of kilobases) that encompasses isochores, interphase chromatin domains, and boundaries and chromosomal bands. A key point was to understand that the same architectural proteins, such as CTCF and cohesin subunits, not only could play a role at interphase but also could be retained in mitotic chromosomes and be reused in the interphase chromatin loops of the new cell cycle. The model of Figure 3 explains a most remarkable property of chromosome architecture through the cell cycle—namely, reversibility—that is, the fact that at the end of mitosis, the original interphase chromatin loop structure can be precisely recovered thanks to the retention of architectural proteins. Needless to say, the process is possible because it relies on unchanged genome sequences and on the consequent conserved locations of protein binding sites.

As far as the link between chromosome architecture and genome organization is concerned, we already knew that correlations existed (1) between the GC levels of contiguous coding and noncoding sequences and of the three codon positions and (2) between the GC levels of isochores and all structural/functional properties of the genome. These findings supported the idea that the genome is an integrated ensemble of isochores. A new, important point is now added to these correlations, which were called the genomic code, by finding that correlations also exist between the compositional properties of isochores and the structural properties of chromosomes through the cell cycle. The most remarkable correlation is that between the architecture of interphase chromatin and the isochore organization of the genome (G Bernardi, in prep.), because this new point considerably extends the significance of the genomic code and leads to a unifying view of genome organization and chromosome architecture. In fact, the genomics code is the fourth and last pillar of Molecular Biology, the first three being (1) the double helical structure of DNA (1951–1953); (2) the regulation of gene expression in *Escherichia coli* and the discovery of messenger RNA (1957–1961); and (3) the genetic code (1961–1966). The genomic code was a much more difficult problem as indicated by the fact that it took 55 years (1960–2015) to be solved against 15 years for the other three issues together.

In general terms, the present results fulfill the old prophecy that “order must be in chromosomes” (Rabl 1885), confirm the notion that isochores represent “a fun-

damental level of genome organization” (Eyre-Walker and Hurst 2001), and represent a break through in our understanding of the eukaryotic genome.

ACKNOWLEDGMENTS

The author is indebted to the Department of Science of Roma Tre University (especially Paolo Ascenzi and Riccardo Angelini) for hospitality, to Caterina Nuvoli for excellent technical help, to Giacomo Bernardi and Oliver Clay for critical reading, detailed comments, and discussions, and to Fernando Alvarez-Valin, Maria Costantini, Paolo Cozzi, Kamel Jabbari, Gabriel Macaya, and Ettore Olmo for useful suggestions. This paper is dedicated to the memory of Max Birnstiel, Eric Davidson, Walter Gehring, Jerzy Jurka, and Emile Zuckerkandl, old friends from the Golden Age of Molecular Biology.

REFERENCES

- Arhondakis S, Auletta F, Bernardi G. 2011. Isochores and the regulation of gene expression in the human genome. *Gen Biol Evol* **3**: 1080–1089.
- Bernardi G. 1989. The isochore organization of the human genome. *Annu Rev Genet* **23**: 637–661.
- Bernardi G. 1990. Le génome des vertébrés: Organisation, fonction et évolution. *Biofutur* **94**: 43–46.
- Bernardi G. 1995. The human genome: Organization and evolutionary history. *Annu Rev Genet* **29**: 445–476.
- Bernardi G. 2004 (reprinted in 2005). *Structural and evolutionary genomics. Natural selection in genome evolution*. Elsevier, Amsterdam (out of print but freely available at www.giorgiobernardi.eu).
- Bernardi G. 2007. The neo-selectionist theory of genome evolution. *Proc Natl Acad Sci* **104**: 8385–8390.
- Bernardi G. 2015. Chromosome architecture and genome organization. *PLoS One* **10**: e0143739.
- Bernardi G, Bernardi G. 1991. Compositional properties of nuclear genes from cold-blooded vertebrates. *J Mol Evol* **33**: 57–67.
- Bernardi G, Olofsson B, Filipinski J, Zerial M, Salinas J, Cuny G, Meunier-Rotival M, Rodier F. 1985. The mosaic genome of warm-blooded vertebrates. *Science* **228**: 953–957.
- Berthelot C, Muffato M, Abecassis J, Crollius HR. 2015. The 3D organization of chromatin explains evolutionary fragile genomic regions. *Cell Rep* **10**: 1913–1924.
- Biffi G, Tannahill D, McCafferty J, Balasubramanian S. 2013. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* **5**: 182–186.
- Blobel GA, Kaudake S, Wang E, Lau AW, Xuber J, Chou MM, Vakoc CR. 2009. A reconfigured pattern of MLL occupancy within mitotic chromatin promotes rapid transcriptional reactivation following mitotic exit. *Mol Cell* **36**: 970–983.
- Burke LJ, Zhang R, Bartkuhn M, Tiwari VK, Tavosoidana G, Kurukuti S, Weth C, Leers J, Galjart N, Ohlsson R, et al. 2005. CTCF binding and higher order chromatin structure of the H19 locus are maintained in mitotic chromatin. *EMBO J* **24**: 3291–3300.
- Caron H, van Schaik B, van der Mee M, Baas F, Riggins G, van Sluis P, Hermus MC, van Asperen R, Boon K, Voute PA, et al. 2001. The human transcriptome map: Clustering of highly expressed genes in chromosomal domains. *Science* **291**: 1289–1292.
- Chandra T, Ewels PA, Schoenfelder S, Furlan-Magaril M, Wingett SW, Kirschner K, Thuret JY, Andrews S, Fraser P, Reik W. 2015. Global reorganization of the nuclear landscape in senescent cells. *Cell Rep* **10**: 471–483.

- Corneo G, Ginelli E, Soave C, Bernardi G. 1968. Isolation and characterization of mouse and guinea pig satellite DNA's. *Biochemistry* **7**: 4373–4379.
- Costantini M, Bernardi G. 2008a. The short-sequence designs of isochores from the human genome. *Proc Natl Acad Sci* **105**: 13971–13976.
- Costantini M, Bernardi G. 2008b. Replication timing, chromosomal bands, and isochores. *Proc Natl Acad Sci* **105**: 3433–3437.
- Costantini M, Bernardi G. 2009. Mapping insertions, deletions and SNPs on Venter's chromosomes. *PLoS One* **4**: e5972.
- Costantini M, Clay O, Auletta F, Bernardi G. 2006. An isochore map of human chromosomes. *Genome Res* **16**: 536–541.
- Costantini M, Clay O, Federico C, Saccone S, Auletta F, Bernardi G. 2007. Human chromosomal bands: Nested structure, high-definition map and molecular basis. *Chromosoma* **116**: 29–40.
- Costantini M, Cammarano R, Bernardi G. 2009. The evolution of isochore patterns in vertebrate genomes. *BMC Genomics* **10**: 146–162.
- Cozzi P, Milanesi L, Bernardi G. 2015. Segmenting the human genome into isochores. *Evol Bioinform* **11**: 253–261.
- Cremer T, Cremer M. 2010. Chromosomal territories. *Cold Spring Harb Perspect Biol* **2**: a003889.
- Dekker J, Heard E. 2015. Structural and functional diversity of topologically associating domains. *FEBS Lett* **589**: 2877–2884.
- Dekker J, Rippe K, Dekker M, Kleckner N. 2002. Capturing chromosome conformation. *Science* **295**: 1306–1311.
- Deng W, Rupin JW, Krivega I, Breda L, Motta I, Jahn KS, Reik A, Gregory PD, Rivella S, Dean A, et al. 2014. Reactivation of developmentally silenced globin genes by forced chromatin looping. *Cell* **158**: 849–860.
- Devillers-Thiery A. 1974. Utilisation des endonucleases dans l'étude des sequences des ADN. Thesis, Université Paris VII.
- Dixon JR, Selvaraj S, Yue F, Kim A, Li Y, Shen Y, Hu M, Liu JS, Ren B. 2012. Topological domains in mammalian genomes identified by analysis of chromatin interactions. *Nature* **85**: 376–380.
- Dixon JR, Jung I, Selvaraj S, Shen Y, Antosiewicz-Bourget JE, Lee AY, Ye Z, Kim A, Rajagopal N, Xie W, et al. 2015. Chromatin architecture reorganization during stem cell differentiation. *Nature* **518**: 331–336.
- D'Onofrio G. 2002. Expression patterns and gene distribution in the human genome. *Gene* **300**: 179–187.
- Duret L, Mouchiroud D, Gautier C. 1995. Statistical analysis of vertebrate sequences reveals that long genes are scarce in GC-rich isochores. *J Mol Evol* **40**: 308–317.
- Egli D, Birkhoff G, Eggan K. 2008. Mediators of reprogramming: Transcription factors and transitions through mitosis. *Nat Rev* **9**: 505–516.
- Ernst J, Kellis M. 2010. Discovery and characterization of chromatin states for systematic annotation of the human genome. *Nature* **28**: 817–825.
- Eyre-Walker A, Hurst LD. 2001. The evolution of isochores. *Nat Rev Genet* **2**: 549–555.
- Filion GJ, van Bommel JG, Braunschweig U, Talhout W, Kind J, Ward LD, Brugman W, de Castro IJ, Kerkhoven RM, Bussemaker HJ, et al. 2010. Systematic protein location mapping reveals five principal chromatin types in *Drosophila* cells. *Cell* **15**: 212–224.
- Filipski J, Thierry JP, Bernardi G. 1973. An analysis of the bovine genome by Cs₂SO₄/Ag⁺ density gradient centrifugation. *J Mol Biol* **80**: 177–197.
- Fisher AM, Strike P, Scott C, Moorman AV. 2005. Breakpoints of variant 9;22 translocations in chronic myeloid leukaemia locate preferentially in the CG-rich regions of the genome. *Genes Chromosomes Cancer* **43**: 383–389.
- Follmer NE, Wani AH, Francis NJ. 2012. A Polycomb group protein is retained at specific sites on chromatin mitosis. *PLoS Genet* **8**: e1003135.
- Frenkel ZN, Bettecken T, Trifonov EN. 2011. Nucleosome DNA sequence structure of isochores. *BMC Genomics* **12**: 203–209.
- Fullerton SM, Carvalho AB, Clark AG. 2001. Local rates of recombination are positively correlated with GC content in the human genome. *Mol Biol Evol* **18**: 1139–1142.
- Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. 2004. Chromatin architecture of the human genome: Gene-rich domains are enriched in open chromatin fibers. *Cell* **118**: 555–566.
- Guelen L, Pagie L, Brasset E, Meuleman W, Faza MB, Talhout W, Eussen BH, de Klein A, Wessels L, de Laat W, et al. 2008. Domain organization of human chromosomes revealed by mapping of nuclear lamina interactions. *Nature* **453**: 948–951.
- Hodgkinson A, Chen Y, Eyre-Walker A. 2012. The large-scale distribution of somatic mutations in cancer genomes. *Hum Mutat* **33**: 136–143.
- Hou C, Li L, Qin ZS, Corces VG. 2012. Gene density, transcription, and insulators contribute to the partition of the *Drosophila* genome into physical domains. *Cell* **48**: 471–484.
- Jabbari K, Nurnberg P. 2016. A genomic view on epilepsy and autism candidate genes. *Genomics* (in press).
- Julienne H, Zoufir A, Audit B, Arneodo A. 2013. Human genome replication proceeds through four chromatin states. *PLoS Comput Biol* **9**: e1003233.
- Kadauke S, Udugama MI, Pawlicki JM, Achtman JC, Jain DP, Cheng Y, Hardison RC, Blobel GA. 2012. Tissue-specific mitotic bookmarking by hematopoietic transcription factor GATA1. *Cell* **150**: 725–737.
- Kikuta H, Laplante M, Navratilova P, Komisarczuk AZ, Engström PG, Fredman D, Akalin A, Caccamo M, Sealy I, Howe K, et al. 2007. Genomic regulatory blocks encompass multiple neighboring genes and maintain conserved synteny in vertebrates. *Genome Res* **17**: 545–555.
- Konu O, Li MD. 2002. Correlations between mRNA expression levels and GC contents of coding and untranslated regions of genes in rodents. *J Mol Evol* **54**: 35–41.
- Lemaître C, Zaghloul L, Sagot MF, Gautier C, Arneodo A, Tannier E, Audit B. 2009. Analysis of fine-scale mammalian evolutionary breakpoints provides new insight into their relation to genome organization. *BMC Genomics* **10**: 335.
- Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, et al. 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* **326**: 289–293.
- Macaya G, Thierry JP, Bernardi G. 1976. An approach to the organization of eukaryotic genomes at a macromolecular level. *J Mol Biol* **108**: 237–254.
- Meuleman W, Peric-Hupkes D, Kind J, Beaudry JB, Pagie L, Kellis M, Reinders M, Wessels L, van Steensel B. 2013. Constitutive nuclear lamina–genome interactions are highly conserved and associated with A/T-rich sequence. *Genome Res* **23**: 270–280.
- Meunier-Rotival M, Soriano P, Cuny G, Strauss F, Bernardi G. 1982. Sequence organization and genomic distribution of the major family of interspersed repeats of mouse DNA. *Proc Natl Acad Sci* **79**: 355–359.
- Mouchiroud D, D'Onofrio G, Aïssani B, Macaya G, Gautier C, Bernardi G. 1991. The distribution of genes in the human genome. *Gene* **100**: 181–187.
- Naumova N, Imakaev M, Fudenberg G, Zhan Y, Lajoie BR, Mirny LA, Dekker J. 2013. Organization of the mitotic chromosome. *Science* **342**: 948–953.
- Nora EP, Lajoie BR, Schulz EG, Giorgetti L, Okamoto I, Servant N, Piolot T, van Berkum NL, Meisig J, Sedat J, et al. 2012. Spatial partitioning of the regulatory landscape of the X-inactivation centre. *Nature* **485**: 381–385.
- Ong CT, Corces VG. 2014. CTCF: An architectural protein bridging genome topology and function. *Nat Rev Genet* **15**: 234–246.

- Paulson JR, Laemmli UK. 1977. The structure of histone-depleted metaphase chromosomes. *Cell* **12**: 817–828.
- Pavlicek A, Paces J, Clay O, Bernardi G. 2002. A compact view of isochores in the draft human genome sequence. *FEBS Lett* **511**: 165–169.
- Pope BD, Ryba T, Dileep V, Yue F, Wu W, Denas O, Vera DL, Wang Y, Hansen RS, Canfield TK, et al. 2014. Topologically associating domains are stable units replication-timing resolution. *Nature* **515**: 403–405.
- Rabl C. 1885. Über Zelltheilung. *Morphol Jahrb* **10**: 214–230.
- Rao SSP, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, Sanborn AL, Machol I, Omer AD, Lander ES, et al. 2014. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* **159**: 1665–1680.
- Rynditch A, Zoubak S, Tsyba L, Tryapitsina-Guley N, Bernardi G. 1998. The regional integration of retroviral sequences into the mosaic genomes of mammals. *Gene* **222**: 1–16.
- Saccone S, De Sario A, Della Valle G, Bernardi G. 1992. The highest gene concentrations in the human genome are in T bands of metaphase chromosomes. *Proc Natl Acad Sci* **89**: 4913–4917.
- 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.
- Saccone S, Federico C, Andreozzi L, D'Antoni S, Bernardi G. 2002. Localization of the gene-richest and the gene-poorest isochores in the interphase nuclei of mammals and birds. *Gene* **300**: 169–178.
- Sadoni N, Langer S, Fauth C, Bernardi G, Cremer T, Turner BM, Zink D. 1999. Nuclear organization of mammalian genomes: Polar chromosome territories build up functionally distinct higher order compartments. *J Cell Biol* **146**: 1211–1226.
- Saitoh Y, Laemmli UK. 1994. Metaphase chromosome structure: Bands arise from a differential folding path of the highly AT-rich scaffold. *Cell* **76**: 609–622.
- Schmegner C, Hoegel J, Vogel W, Assum G. 2007. The rate, not the spectrum, of base pair substitutions changes at a GC-content transition in the human *NF1* gene region: Implications for the evolution of the mammalian genome structure. *Genetics* **175**: 421–428.
- Sexton T, Cavalli G. 2015. The role of chromosome domains in shaping the functional genome. *Cell* **160**: 1049–1059.
- Sexton T, Yaffe E, Kenigsberg E, Bantignies F, Leblanc B, Hoichman M, Parrinello H, Tanay A, Cavalli G. 2012. Three-dimensional folding and functional organization principles of the *Drosophila* genome. *Cell* **148**: 1–15.
- Song F, Chen P, Sun D, Wang M, Dong L, Liang D, Xu RM, Zhu P, Li G. 2014. Cryo-EM study of the chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* **344**: 376–379.
- Soriano P, Meunier-Rotival M, Bernardi G. 1983. The distribution of interspersed repeats in non-uniform and conserved in the mouse and human genomes. *Proc Natl Acad Sci* **80**: 1816–1820.
- Thiery JP, Macaya G, Bernardi G. 1976. An analysis of eukaryotic genomes by density gradient centrifugation. *J Mol Biol* **108**: 219–235.
- Travers A. 2014. The 30-nm fiber redux. *Science* **344**: 370–372.
- Varriale A, Bernardi G. 2009. Distribution of DNA methylation, CpGs, and CpG islands in human isochores. *Genomics* **95**: 25–28.
- Versteeg R, van Schaik BD, van Batenburg MF, Roos M, Monajemi R, Caron H, Bussemaker HJ, van Kampen AH. 2003. The human transcriptome map reveals extremes in gene density, intron length, GC content, and repeat pattern for domains of highly and weakly expressed genes. *Genome Res* **13**: 1998–2004.
- Vietri Rudan M, Barrington C, Henderson S, Ernst C, Odom DT, Tanay A, Hadjur S. 2015. Comparative Hi-C reveals that CTCF underlies evolution of chromosomal domain architecture. *Cell Rep* **10**: 1297–1309.
- Vinogradov AE. 2003. DNA helix: The importance of being GC-rich. *Nucleic Acid Res* **31**: 1838–1844.
- Watanabe Y, Fujiyama A, Ichiba Y, Hattori M, Yada T, Ikemura T. 2002. Chromosome-wide assessment of replication timing for human chromosomes 11q and 21q: Disease-related genes in timing-switch regions. *Hum Mol Genet* **11**: 13–21.
- Yanagida M. 2009. Clearing the way for mitosis: Is cohesin a target? *Nat Rev* **10**: 489–496.
- Yunis JJ. 1981. Mid-prophase human chromosomes. The attainment of 2000 bands. *Hum Genet* **56**: 293–298.
- Zaret KS. 2014. Genome reactivation after the silence in mitosis: Recapitulating mechanisms of development? *Dev Cell* **29**: 132–134.
- Zoubak S, Clay O, Bernardi G. 1996. The gene distribution of the human genome. *Gene* **174**: 95–102.
- Zuin J, Dixon JR, van der Reijden MI, Ye Z, Kolovos P, Brouwer RW, van de Corput MP, van de Werken HJ, Knoch TA, van Ijcken WF, et al. 2014. Cohesin and CTCF differentially affect chromatin architecture and gene expression in human cells. *Proc Natl Acad Sci* **111**: 996–1001.

Publisher's Comment

A reader raised the issue of similarity between this *Cold Spring Harbor Symposia on Quantitative Biology* (SQB) review:

Bernardi G. 2015. Genome organization and chromosome architecture. *Cold Spring Harb Symp Quant Biol*. Published in Advance January 22, 2016, doi: 10.1101/sqb.2015.80.027318

and an article published by *PLoS One* in November 2015:

Bernardi G. 2015. Genome organization and chromosome architecture. *PLoS One* **10**: e0143739. doi: 10.1371/journal.pone.0143739.

There is some overlap in content between the two publications; however, there are also differences in focus and scope. While the PLOS submission was under consideration, Professor Bernardi informed PLOS of his intention to publish another manuscript in SQB that put the work in a larger perspective. Professor Bernardi informed the publisher of SQB of the article submitted to *PLoS One*, he cited the *PLoS One* article, and he attributed the reproduced figures to the *PLoS One* article in the figure legends.

Comment

Cold Spring Harb Symp Quant Biol **80** (2015) doi: 10.1101/sqb.2015.80.027318

Publisher's Comment

A reader raised the issue of similarity between this *Cold Spring Harbor Symposia on Quantitative Biology* (SQB) review:

Bernardi G. 2015. Genome organization and chromosome architecture. *Cold Spring Harb Symp Quant Biol*. Published in Advance January 22, 2016, doi: 10.1101/sqb.2015.80.027318

and an article published by *PLoS One* in November 2015:

Bernardi G. 2015. Genome organization and chromosome architecture. *PLoS One* **10**: e0143739. doi: 10.1371/journal.pone.0143739.

There is some overlap in content between the two publications; however, there are also differences in focus and scope. While the PLOS submission was under consideration, Professor Bernardi informed PLOS of his intention to publish another manuscript in SQB that put the work in a larger perspective. Professor Bernardi informed the publisher of SQB of the article submitted to *PLoS One*, he cited the *PLoS One* article, and he attributed the reproduced figures to the *PLoS One* article in the figure legends.

doi: 10.1101/sqb.2015.80.030197

Cold Spring Harbor Symposia on Quantitative Biology

Genome Organization and Chromosome Architecture

Giorgio Bernardi

Cold Spring Harb Symp Quant Biol 2015 80: 83-91 originally published online January 22, 2016
Access the most recent version at doi:[10.1101/sqb.2015.80.027318](https://doi.org/10.1101/sqb.2015.80.027318)

Related Content

Publisher's Comment

[Cold Spring Harb Symp Quant Biol UNKNOWN](#) , 2015 80:

References

This article cites 84 articles, 22 of which can be accessed free at:
<http://symposium.cshlp.org/content/80/83.full.html#ref-list-1>

Articles cited in:

<http://symposium.cshlp.org/content/80/83.full.html#related-urls>

License

Email Alerting Service

Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

To subscribe to *Cold Spring Harbor Symposia on Quantitative Biology* go to:
<http://symposium.cshlp.org/subscriptions>