



Minireview

Genome dynamics over evolutionary time: “C-value enigma” in light of chromosome structure

Predrag Slijepcevic

Department of Life Sciences, College of Health and Life Sciences, Brunel University London, UB8 3PH, Uxbridge, UK

ARTICLE INFO

Keywords:

Chromosome rearrangements
Genome evolution
C-value enigma
Nucleotype
Chromothripsis
Informatics metaphor

ABSTRACT

Eukaryotic genome evolution integrates processes behind (i) chromosome plasticity (change in chromosome structure and number), (ii) genome stability maintenance (perfect stability would prevent adaptive processes) and (iii) genome size. Relationships between these variables remain enigmatic, hence the term “C-value enigma”. This term reflects an apparent lack of correlation between genome size and perceived organismal complexity, replacing an older term “C-value paradox”. A useful concept for explaining the enigma is the nucleotypic function, a pluralistic approach unifying a range of phenomena not covered by the conventional genotype and phenotype concepts. In this paper I expand the nucleotype function by adding two additional elements. First element is the “informatics metaphor” according to which genomes act as information-processing entities integrating “hardware” (structural DNA + epigenetic-related DNA) and “software” (protein-coding DNA) components of the genome into a single unit behind organismal fitness. Second element is gross chromosome restructuring, or chromothripsis, as a novel process behind evolutionary chromosome plasticity.

I believe that all the studies that we carried out during the last three decades have demonstrated that DSBs are primarily responsible for the formation of chromosomal aberrations.

A.T. Natarajan, 2012. Reflections on a lifetime in cytogenetics [1].

1. Introduction

Evolutionary genome dynamics is described accurately, but somewhat cryptically, by the phrase “genome in pieces” which refers to chromosomes [2]. If we define chromosomes as genome compartments consisting of DNA and proteins (histone and non-histone) equipped with two functional elements, centromeres and telomeres, bacteria and archaea lack chromosomes [3]. Thus, chromosomes are emergent structures formed at the point of a major prokaryote to eukaryote evolutionary transition [4,5]. In terms of DNA profiles chromosomes are mosaics consisting of repetitive DNA elements, transposons, autonomous retrotransposons, non-autonomous retrotransposons and a small fraction of protein-coding DNA [6].

The key event that shaped the eukaryotic genome at the point of a symbiotic merger between an archaeal host and an α -proteobacterial partner, was the fragmentation of the host’s circular genome into multiple linear fragments (genome in pieces) caused by the massive invasion of the partner’s mobile group II introns [5,7]. The newly formed linear chromosomes were unstable until the processes for their

stabilization emerged. These processes were first mediated by retrotransposons and later by properly functioning telomeres synthesized by the enzyme telomerase, a former retrotransposon itself [8]. The membrane structure separating the genome from the rest of the cell emerged in parallel, leading to formation of the cell nucleus.

According to one scenario, the end-stabilizing structures of newly formed chromosomes also served as centromeres by virtue of being attractors of the tubulin-based spindle apparatus [5]. This scenario predicts that the first eukaryotic chromosomes were telocentric: the end structures played the role of telomeres and centromeres simultaneously. The karyotype evolution that ensued was mediated by telomeres and centromeres interchanging functionally and thus acting as factors behind the evolutionary chromosome plasticity [7].

Thus, the phrase “genomes in pieces” reflects the evolutionary point at which the true chromosomes emerged. The composite structure of chromosomal DNA (viral, mobile genetic elements etc.) parallels the composite structure of the eukaryotic cell in which formerly independent prokaryotes merged into the symbiotic partnership [4].

While the replication of circular prokaryotic genomes is relatively simple, the process behind the eukaryotic genome replication and division is more complex. It requires formation of the specialized structure, the mitotic spindle, initiated by the pair of centrosomes. The spindle serves as the structure for (i) attracting the chromosomes via centromeres in mitosis and meiosis, (ii) splitting them into two double chromatid (meiosis I) or single chromatid (meiosis II and mitosis)

E-mail address: predrag.slijepcevic@brunel.ac.uk.

<https://doi.org/10.1016/j.mrgentox.2018.05.005>

Received 15 February 2018; Received in revised form 28 March 2018; Accepted 3 May 2018
Available online 05 May 2018

1383-5718/ © 2018 Elsevier B.V. All rights reserved.

chromosomes and (iii) subsequent chromosome segregation into daughter cells. The entire process, when viewed under time-lapse microscopy, is reminiscent of the coordinated “chromosome dance” which conforms to the rules of cellular mechanics. For example, the longest chromosome arm must not exceed half of the average length of the spindle axis at telophase to prevent premature cut of chromatids by cell walls and subsequent breakage-fusion-bridge cycles [9].

The aim of this paper is to look at processes by which the composite eukaryotic genome, or “the genome in pieces”, changes over evolutionary time. The traditional term for these dynamic changes is chromosome rearrangements: numerical (polyploidy, dysploidy, aneuploidy) or structural (e.g. translocations). Structural chromosomal rearrangements result from DNA double strand breaks (DSBs) [10–12]. I will argue that evolutionary chromosome rearrangements emerge from an interplay between processes that set the species-specific genome size and their counterparts behind chromosome plasticity (change in chromosome structure/number). I will also argue that chromosomes heralded a major evolutionary shift. In addition to the information-processing role of DNA, dominating genomes of chromosome-free prokaryotes, chromosomal DNA acquired the role of the information-free building material.

I start by outlining the relationship between genome size and chromosome number. This provides the platform for introducing the concept of nucleotide function and “informatics metaphor”. In the final part I integrate aspects of genome stability maintenance, nucleotide function and “informatics metaphor” to provide a new angle for understanding processes behind evolutionary chromosome rearrangements.

2. Evolutionary genome size dynamics and theories behind it

The question whether genome size reflects organismal complexity occupied biologists for decades. It started as the “C-value paradox”. C-value is the amount of DNA in unreplicated gametic (haploid) nuclei. C-values can be expressed in (i) picograms (pg) and (ii) number of base pairs (1 pg = 0.978 Gb) [13]. Early studies revealed that human genomic DNA exceeds that of bacteria by a thousand times. However, some species of fish and flowering plants, for example, have 30–50 times more genomic DNA than humans [14]. This lack of correlation between genome size and the intuitively perceived organismal complexity was termed “C-value paradox” [15].

Based on data for ~15,000 species [16,17], the size of eukaryotic genomes varies by a factor of 65,000 [reviewed in 13]. The smallest known eukaryotic genome, that of the microsporidean *Encephalitozoon intestinalis*, has been estimated at 0.0023 Gb. The largest genome is found in the flowering plant *Paris japonica* at 148.8 Gb. Early estimates, according to which some species of protists have genomes > 200 Gb, have been discarded as technically inaccurate [13].

Given that one nucleotide is estimated to be ~0.34 nm in length, this translates into following lengths: ~1.5 mm for the smallest genome, and ~100 m for the largest genome. The human genome, by comparison, is estimated at ~3 Gb and ~2 m length. A summary of genome sizes across the species range is presented in Fig. 1. Extra-large or extra-small genomes are rare. Most eukaryotic genomes fall within the 1–100 Gb range.

Genome size is usually measured by cytological methods such as Feulgen densitometry and flow cytometry [13]. In recent years whole genome sequencing data have become available for a large number of species. This offers an opportunity to compare cytological estimates with the whole genome sequencing-based calculations of genome size. In the first study of this sort a total of 501 species were analysed including: 148 species of animals, 81 land plants, 202 fungi and 70 protists [18]. Analysis revealed a good correspondence between cytological estimates of genome size and their sequence-based counterparts. However, some discrepancies have been identified. For example, there was a general tendency of sequence-based calculations data to provide a

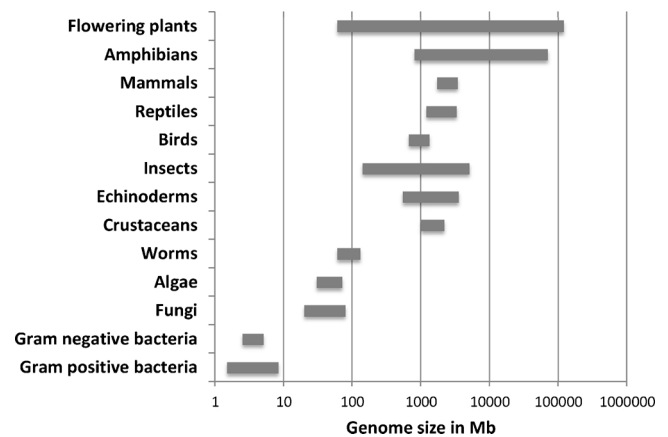


Fig. 1. Genome size across the evolutionary species range. Adapted from Wikipedia.

lower estimate of genome size than cytological methods. The mismatch becomes more pronounced with increasing genome size. This could be due to inherent difficulty associated with sequencing heterochromatic (highly repetitive) regions of genomes leading towards systematic underestimation of genome content.

The question arises as to which evolutionary factors influence species-specific genome size. In recent literature this question is described as the “C-value enigma” [14]. Several theories have been proposed to explain the genome size diversity in eukaryotes. The “mutation pressure” theories argue that large proportions of the genome are “junk” or “selfish” DNA in the form of transposable elements, retroviruses, bacterial plasmids etc. The first proponent of this view was Susumu Ohno [19] who suggested that the genome is a repository of extinct genes, which represent a true genetic “junk”. However, extinct genes or “pseudogenes” form only a small proportion of eukaryotic genomes. Later proponents of “mutation pressure” theories argued that “junk” DNA is any sequence that lacks coding or regulatory function [20]. Some authors criticized the tendency of evolutionary biologists towards “adaptationism” – every component of the genome must account for an organismal function. Instead, they thought that DNA within eukaryotic genomes can expand for its own benefit, hence “selfish” or parasitic, a process called either “non-phenotypic selection” or “intra-genomic selection” [21,22]. The expansion of “selfish” DNA would become a burden to the host cell only when other physiological processes within the cell are affected [23]. Thus, the C-value of a species is a product of balancing between mutation pressure, which acts to expand genome size and physiological tolerance factors within the host cell. Following this logic, all “mutation pressure” theories argue that the relationship between the cell volume and DNA content is coincidental.

On the other hand, “optimal DNA” theories argue that there is a strong correspondence between DNA content and cell and nuclear volumes. These theories are governed by different principles at the heart of which is the view that DNA plays not only qualitative roles in determining organismal fitness, but also secondary or quantitative roles. For example, a secondary or quantitative role of DNA is its ability to act as the building material for chromosomes independently of its information-carrying role (see below). Thus, C-value may be the product of selection forces acting via secondary or quantitative roles. For example, “nucleoskeletal theory” suggests that there is a co-evolutionary interaction between cell and nuclear volumes [24]. In this scenario, DNA content or C-value is a secondary product of an evolutionary compromise between “selection for cell size and for developmental rates”.

Similarly, but more pluralistically Gregory [14] argued that a hierarchical approach is required to fully understand the C-value enigma. This approach is best described by the term “nucleotide” function [25,26] which introduces a new hierarchical level positioned between

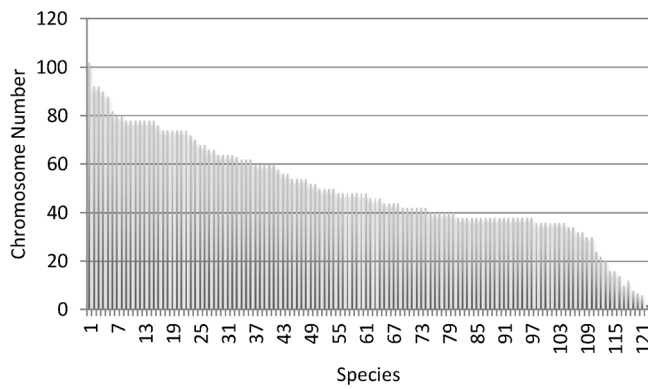


Fig. 2. Chromosome numbers in > 100 species of animals and plants listed in a database [31].

the genotype and phenotype levels. This new level of biological hierarchy allows for pluralistic inclusion of factors ignored by the classical genotype and phenotype hierarchies. The new factors include: nuclear structure and cell size, cell division times and developmental rates, chromosome territories within the interphase nuclei, 3 D nuclear topology, mechanical forces acting on the cell, selfish gene level selection, supraorganismal level selection such as group selection and genetic drift [24,27,28]. I will argue later that this pluralistic approach is the most productive way for understanding the interplay between genome size and chromosome plasticity.

3. Is there a correlation between genome size and chromosome number?

The number of genomic pieces (chromosomes) per eukaryotic genome is highly variable. The lowest chromosome number of $n = 2$ (one pair) is found in an ant species *Myrmeca pilosula* [29]. The highest chromosome number reported in animals is $n = 372$ found in North American shortnose sturgeon *Acipenser brevirostrum* [30]. Many plant species have chromosome numbers of $n = > 200$ [13]. Large chromosome numbers are rare and represent the consequence of polyploidy. Fig. 2 shows more typical chromosome numbers in a group of > 100 animal and plant species listed in [31].

An interesting question is how chromosome number/structure in a given complement or karyotype, relates to genome size. Evolutionarily close species may have either minor karyotype differences as exemplified by only one chromosome rearrangement in the human karyotype relative to the ape karyotype [32], or dramatic differences as exemplified by the karyotypes of two closely related deer species, one of which has the lowest chromosome number in mammals (Indian muntjac, $2n = 6$ or 7) and the other one (Chinese muntjac) a more typical mammalian chromosome number ($2n = 46$) [33]. This is also paralleled by a significant change in the C-value, which is reduced from 2.7 pg to 2.1 pg (30% difference) in Indian muntjac relative to Chinese muntjac. There is no significant difference between C-values of human and ape genomes.

The relationship between chromosome number and genome size has been investigated by only a few studies. For example, the study of genome size analysis of 501 species assessed by sequencing-based calculations revealed a weak but significant positive correlation between genome size and chromosome number [18]. In this study chromosome number varied from $n = 3$ in various animal species to $n = 84$ in sea lamprey *Petromyzon marinus*. The average genome size was 1.154 Gb for animals, 1.066 Gb for land plants and 0.035 Gb for fungi.

However, in a more recent study covering 6052 species no correlation between genome size and chromosome number was found [34]. The study targeted all species for which (i) genome size has been estimated by cytological methods and (ii) chromosome number is known.

These two examples suggest that the relationship between genome size and chromosome number is either weak or non-existent.

4. Nucleotype function and “informatics metaphor”

Eukaryotic genomes exist only in three-dimensional space of the cell nucleus. This highlights importance of the nucleotype level hierarchy [14]. In line with this, the old view of a random arrangement of genomic pieces in the nucleus, reminiscent of “spaghetti in a can”, has gradually been replaced by the notion of a precise nuclear topology in which individual chromosomes occupy distinct territories, typically spherical in shape and 2–4 μm in diameter [28]. In larger genomes chromosome territories are compact and interact with each other only through their peripheries. By contrast, smaller genomes (e.g. yeast) have less well-defined chromosome territories which interact with each other more readily.

Invention of new techniques such as chromosome conformation capture (3C) and its derivatives such as genome wide 3C (Hi-C) enabled precise analysis of sub-chromosomal region interactions within the 3 D space of the nucleus at the sequence level. A series of studies revealed individual chromosome domains termed topologically associated domains or TADs [35]. TADs can be distinguished on the basis of their expression activity, replication timing and histone markers. They have sharp boundaries separated by chromatin insulator proteins. In addition to TADs, chromatin domains interacting with non-DNA structures, such as the nuclear lamina, termed LADs (lamina associated domains) have also been identified [35]. LADs are characterized by genetic silence and late replication timing. Thus, chromosome structure can be described as a non-random arrangement of TADs and LADs, which must be equipped with two essential functional elements, centromeres and telomeres.

Chromosome topology within the 3 D nuclear space is tissue specific and reproducible after each cell division suggesting that a form of cell memory is at play. This notion is further supported by the discovery that the nucleus is precisely reconstructed after each cell division by the BAF mediated stiff chromatin surface that specifies nuclear geometry in a reproducible manner [36].

The general picture emerging from high-resolution studies based on Hi-C technologies is in line with the pluralistic principles of the nucleotype level hierarchy. On one hand, there is a clear link between the genome organization and its function as exemplified by the fact that gene positioning within the nucleus is likely to affect its expression. On the other hand, a precise position of a gene within the 3 D space of the nucleus is not an absolute requirement for its proper function indicating a degree of functional redundancy.

This degree of redundancy should be assessed in light of “informatics metaphor” [37,38]. Genomes are viewed as sophisticated information processing and information storing entities. In contrast to conventional man-made information processing machines or computers, in which hardware and software components are separate and independent from each other, genomes integrate and unify these two components. This is visible from the key difference between the universal Turing machine and genomes as biological information processors. The universal Turing machine does not change in the process of computation. By contrast, genomes do change in the process of computation, thus enabling genome and organismal evolution [38].

Here is a brief description of genome composition in light of “informatics metaphor”. The phenotype-determining information is stored in genomic parts which code for proteins (**genetic information**). This represents biological software or programme required to establish organismal phenotype via RNA and proteins (the central dogma of molecular biology). The hardware required to run this programme is a combinatorial function of DNA and histone proteins (**epigenetics and histone code**) operated by the 3 D nuclear structure of chromosomes in which repetitive and genetically redundant forms of DNA dominate (**structural role of DNA**).

To illustrate quantitative contribution to the three-component

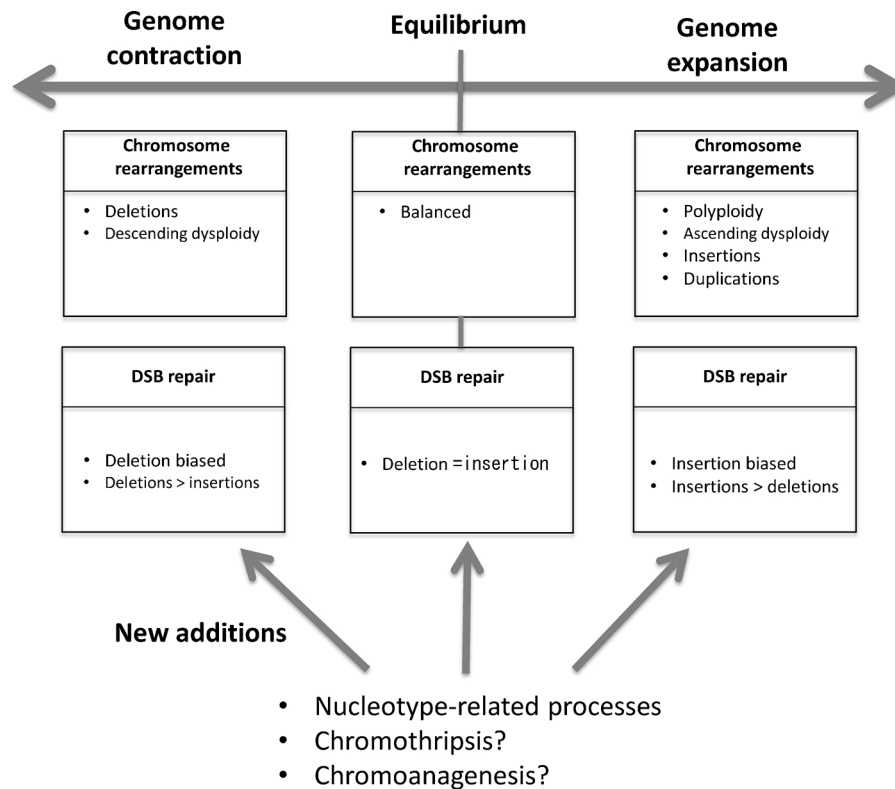


Fig. 3. Processes behind genome evolution.

genome we can use the human genome. Only 1.5% of the human genome represents genetic information or protein coding genes. Assessing the amount of DNA involved in epigenetic or regulatory function is more difficult. In spite its enormous length DNA containing chromatin occupies only 15% of the nuclear volume [39]. Others estimated that an upper limit of functional DNA in the human genome is 25% [40]. The remaining 75% represents structural DNA with no informationally or regulatory useful properties.

Thus, DNA has a secondary or structural role required for the “hardware” component of the genome. This scenario is largely in line with the nucleotide level hierarchy, which does not dismiss the notion of “junk” DNA as a form of informational redundancy utilized for chromosome structure.

To express this stance more simplistically it may be useful to borrow a sentence from Arthur Koestler used in a different context: “... the chemical analysis of bricks and mortar will tell you next to nothing about the architecture of a building.” [41] In the context of the nucleotide theory “bricks and mortar” represent structural DNA. The “building” represents organismal phenotype.

Thus, chromosomes are emergent structures, different from bacterial nucleoids, in which DNA molecules acquired a structural role, in addition to their genetic information-carrying role.

5. Discussion

Genome size and chromosome plasticity are independent variables. Processes behind their regulation are critically important for the full understanding of genome evolution. Occasionally, there are disagreements as to how the processes should be interpreted. A representative example is a critique of the way the ENCODE (ENCyclopedia Of DNA Elements) project interprets genome functionality. Doolittle [27] and Graur et al. [42] criticized the ENCODE’s stance that 80% of the human genome is functional. This contradicted earlier estimates that only 5% of mammalian genomes is under the evolutionary constraint. Furthermore, Doolittle [27] and Graur et al. [42] argued that the notion of

“selfish” or “junk” DNA cannot be discarded as long as some DNA in the genome is neutral in terms of selection effects. Doolittle [27] suggested that: “A larger theoretical framework, embracing informational and structural roles for DNA, neutral as well as adaptive causes of complexity, and selection as a multilevel phenomenon, is needed.”

Perhaps the first step in this direction is to recall the three component “informatics metaphor” [37,38] and recognize that (i) the only organisms in which majority of genomic DNA codes for proteins and therefore plays exclusively “genetic information” role are prokaryotes (ii) prokaryotic and eukaryotic genomes differ dramatically with latter containing emergent structures, chromosomes, absent in the former and (iii) with the emergence of chromosomes DNA acquired two more roles with respect to “informatics metaphor”, epigenetic and chromosome structure-supporting role, which play the hardware part of the genome.

In line with the pluralistic approach of the nucleotide theory which generally embraces the above assumptions, it seems likely that at least some DNA in eukaryotic genomes is neutral in terms of selection effects and thus can justifiably be called “junk”. The exact proportion of this class of DNA is unknown. The ENCODE project, when completed, should be able to provide a precise estimate [43].

The presence of selection-neutral DNA in the eukaryotic genomes is needed for chromosome structure. In line with the “informatics metaphor” genomes may be viewed as biological information processors in which the hardware component is represented by the chromosome structure and epigenetic properties of the genome, whereas the software component containing instructions for protein synthesis is represented by the unique DNA containing the organism-specific genetic information (genetic code). The hardware component must have non-adaptive structural redundancy (“junk DNA”) to enable genome evolution.

Once the above considerations are taken into account we can proceed to assess processes behind evolutionary chromosome rearrangements from a new angle. The conventional view is that genome evolution is a balancing process between genome size and plasticity with three broad outcomes: expansion, contraction and equilibrium [44]

(Fig. 3). Processes behind genome expansion include chromosome gains (insertions and duplications), ascending dysploidy (increase in chromosome number) and whole genome duplications (polyploidy). Processes behind genome contraction include chromosome deletions and descending dysploidy (chromosome number reduction). Processes that preserve genomic equilibrium include neutral chromosome rearrangements lacking major DNA gains or losses (balanced rearrangements).

Molecular and cellular processes behind chromosome rearrangements include: (i) deletion-biased DSB repair leading to genome contraction, (ii) insertion-biased DSB repair including retroelement dispersion leading to genome expansion, (iii) balanced outcomes of DSB repair processes (e.g. erroneous repair leading to translocations) with neutral effects on genome size and (iv) constraints of cellular division mechanics (size of individual chromosome arms; see above) which lead to genome contraction [9–12,44] (Fig. 3).

The conventional view should be revised to take account of the nucleotype theory but also the fact that chromosome structure as the hardware part of eukaryotic genomes, different from the classical Turing machine, may have some autonomy in terms of evolutionary restructuring (Fig. 3). In line with this possibility gross chromosome restructuring termed chromothripsis or chromosome shattering has recently been discovered [reviewed in [45]]. Chromothripsis is characterized by the massive chromosome fragmentation and subsequent chromosome repair through rejoining fragments randomly by DSB repair processes. Mechanisms behind chromothripsis are not fully understood but include: (i) micronuclei formation and subsequent aberrant DNA replication or premature chromosome condensation, (ii) telomere-mediated chromothripsis and (iii) exposure to ionizing radiation [46–48].

Even though chromothripsis has originally been identified in cancer cells, research has shown that chromothripsis operates in human germline cells [49,50], thus opening a possibility that this might be an evolutionary relevant process. In line with this, chromothripsis-like events have been discovered in the small photosynthetic eukaryote, *Ostreococcus tauri* (Chlorophyta, Mamiellophyceae) [51]. Massive chromosome rearrangements in *O. tauri*, such as insertions and translocations are reminiscent of chromothripsis-mediated chromosome fragmentation and subsequent fragment rearrangements.

Furthermore, chromothripsis has been shown to act as a mechanism for genome reduction in the plant *A. thaliana* [52]. Interestingly, it is mediated by CENH3 (centromeric histone 3) which, if mutated, causes formation of micronuclei containing whole chromosomes. The micronuclei-bound chromosomes suffer catastrophic fragmentation. Some fragments are then stitched together by Ligase IV and the rest are lost. The newly emerging chromosomes can be inherited by subsequent generations, thus leading to genome contraction in *A. thaliana*. This example illustrates that chromothripsis can shape evolutionary chromosome rearrangements. A slightly different form of chromothripsis termed chromoanagenesis [45] virtually means rebirth or creation of new chromosomes. It remains to be established whether chromothripsis and chromoanagenesis are widespread evolutionary phenomena (Fig. 3).

Finally, a recent study shows clear aspects of nucleotype-related mechanisms at work during genome reduction in flowering plants [53]. It has been argued that the apparent evolutionary success of angiosperms or flowering plants, measured by their dominance in terrestrial ecosystems, can be explained by a dramatic reduction in genome size which enabled reduction in cell size. This in turn created a maximum space for packing more stomata and veins into their leaves leading to maximization of angiosperm primary productivity in terrestrial ecosystems.

In conclusion, this paper integrates aspects of genome stability maintenance and chromosome aberration mechanisms [10–12], cell mechanics [7,9], nucleotype-related processes [14] and information theory [37,38] to provide a new angle for understanding processes behind genome evolution.

Conflict of interest

None.

Acknowledgement

Supported by Brunel University London.

References

- [1] A.T. Natarajan, Reflections on a lifetime in cytogenetics, *Mutat. Res.* 751 (2012) 1–6.
- [2] M. Garavís, C. González, A. Villasante, On the origin of the eukaryotic chromosome: the role of noncanonical DNA structures in telomere evolution, *Genome. Biol. Evol.* 5 (2013) 1142–1150.
- [3] L. Margulis, Serial endosymbiotic theory (SET) and composite individuality: transition from bacterial to eukaryotic genomes, *Microbiol. Today* 31 (2004) 172–174.
- [4] L. Margulis, M.F. Dolan, R. Guerrero, The chimeric eukaryote: origin of the nucleus from the karyomastigote in amitochondriate protists, *Proc. Natl. Acad. Sci. U. S. A.* 97 (2000) 6954–6959.
- [5] A. Villasante, J.P. Abad, M. Méndez-Lago, Centromeres were derived from telomeres during the evolution of the eukaryotic chromosome, *Proc. Natl. Acad. Sci. U. S. A.* 104 (2007) 10542–10547.
- [6] H.H. Kazazian Jr., Mobile elements: drivers of genome evolution, *Science* 303 (2004) 1626–1632.
- [7] P. Slijepcevic, Mechanisms of the evolutionary chromosome plasticity: integrating the 'Centromere-from-Telomere' hypothesis with telomere length regulation, *Cytogenet. Genome Res.* 148 (2016) 268–278.
- [8] T.M. Nakamura, T.R. Cech, Reversing time origin of telomerase, *Cell* 92 (1998) 587–590.
- [9] I. Schubert, J.L. Oud, There is an upper limit of chromosome size for normal development of an organism, *Cell* 88 (1997) 515–520.
- [10] A.T. Natarajan, G. Obe, Molecular mechanisms involved in the production of chromosomal aberrations. I. Utilization of *Neurospora* endonuclease for the study of aberration production in G2 stage of cell cycle, *Mutat. Res.* 52 (1978) 137–149.
- [11] A.T. Natarajan, G. Obe, Molecular mechanisms involved in the production of chromosomal aberrations. III. Restriction endonucleases, *Chromosoma* 90 (1984) 120–127.
- [12] P.E. Bryant, Enzymatic restriction of in situ mammalian cell DNA using PvuII and BamHI: Evidence for the double strand break origin of chromosomal aberrations, *Int. J. Radiat. Biol.* 46 (1984) 57–65.
- [13] O. Hidalgo, J. Pellicer, M. Christenhusz, H. Schneider, A.R. Leitch, L.J. Leitch, Is there an upper limit to genome size? *Trends Plant Sci.* 22 (2017) 567–573.
- [14] T.R. Gregory, Coincidence coevolution, or causation? DNA content, cell size, and the C-value enigma, *Biol. Rev.* 76 (2001) 65–101.
- [15] C.A. Thomas, The genetic organization of chromosomes, *Annual Rev. Genet.* 5 (1971) 237–256.
- [16] Animal Genome Size Database. <http://www.genomesize.com/>. (Accessed 10 February 2018).
- [17] Plant DNA C-values Database. <http://data.kew.org/cvalues/>. (Accessed 10 February 2018).
- [18] T.R. Gregory, T.A. Elliott, What's in a genome? The C-value enigma and the evolution of eukaryotic genome content, *Phil. Trans. R. Soc. B* 370 (2015) 20140331.
- [19] S. Ohno, So much 'junk' DNA in our genome, in: H.H. Smith (Ed.), *Evolution of Genetic Systems*, Gordon & Breach, New York, 1972, pp. 366–370.
- [20] M. Pagel, R.A. Johnstone, Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox, *Proc. R. Soc. London B* 249 (1992) 119–124.
- [21] W.F. Doolittle, C. Sapienza, Selfish genes, the phenotype paradigm and genome evolution, *Nature* 284 (1998) 601–603.
- [22] L.E. Orgel, F.H.C. Crick, Selfish DNA: the ultimate parasite, *Nature* 284 (1980) 604–607.
- [23] B. Charlesworth, P. Sniegowski, W. Stephan, The evolutionary dynamics of repetitive DNA in eukaryotes, *Nature* 371 (1994) 215–220.
- [24] T. Cavalier-Smith, Nuclear volume control by nucleoskeletal DNA: selection for cell volume and cell growth rate and the solution of the DNA C-value paradox, *J. Cell Sci.* 34 (1978) 247–278.
- [25] M.D. Bennett, The duration of meiosis, *Proc. R. Soc. London B* 178 (1971) 277–299.
- [26] M.D. Bennett, Nuclear DNA content and minimum generation time in herbaceous plants, *Proc. Royal Soc. London B* 181 (1972) 109–135.
- [27] F.W. Doolittle, Is junk DNA bunk? a critique of ENCODE, *Proc. Natl. Acad. Sci. U. S. A.* 110 (2013) 5294–5300.
- [28] G. Cavalli, T. Misteli, Functional implications of genome topology, *Nat. Struct. Mol. Biol.* 20 (2013) 290–299.
- [29] M.W.J. Crosland, R.H. Crozier, *Myrmeciapilosula*, an ant with only one pair of chromosomes, *Science* 231 (1986) 1278.
- [30] D.S. Kim, Y.K. Nam, J.K. Noh, C.H. Park, F.A. Chapman, Karyotype of North American shortnose sturgeon *Acipenser brevirostrum* with the highest chromosome number in the Acipenseriformes, *Ichthyol. Res.* 52 (2005) 94–97.
- [31] List of organisms by chromosome count. https://en.wikipedia.org/wiki/List_of_organisms_by_chromosome_count. (Accessed 10 February 2018).
- [32] J.W.I. Jdo, A. Baldini, D.C. Ward, S.T. Reeders, R.A. Wells, Origin of human chromosome 2: an ancestral telomere-telomere fusion, *Proc. Natl. Acad. Sci. U. S. A.*

- A. 88 (1991) 9051–9055.
- [33] V. Tshipouri, M.G. Schueler, S. Hu, N.I.S.C. Comparative Sequencing Program, A. Dutra, E. Pak, H. Riethman, E.D. Green, Comparative sequence analyses reveal sites of ancestral chromosomal fusions in the Indian muntjac genome, *Genome Biol.* 9 (2008) R155.
- [34] A.M. El-Shehawi, M.M. Elseehy, Genome size and chromosome number relationship contradicts the principle of Darwinian evolution from common ancestor, *J. Phylogenetics Evol. Biol.* 5 (2017) 179.
- [35] A. Gonzalez-Sandoval, S.M. Gasser, On TADs and LADs Spatial control over gene expression, *Trends Genet.* 32 (2016) 485–495.
- [36] M. Samwer, M.W.G. Schneider, R. Hoefler, P.S. Schmalhorst, J.G. Jude, J. Zuber, D.W. Gerlich, DNA cross-bridging shapes a single nucleus from a set of mitotic chromosomes, *Cell* 170 (2017) 956–972.
- [37] J.A. Shapiro, R. Sternberg, Why repetitive DNA is essential to genome function, *Biol. Rev.* 80 (2005) 1–24.
- [38] E. Ben-Jacob, Bacterial wisdom, Gödel's theorem and creative genomic webs, *Physica A* 248 (1998) 57–76.
- [39] J. Dekker, T. Misteli, Long-Range chromatin interactions, *Cold Spring Harb. Perspect. Biol.* 7 (2015) a019356.
- [40] D. Graur, An upper limit on the functional fraction of the human genome, *Genome Biol. Evol.* 9 (2017) 1880–1885.
- [41] A. Koestler, *The Ghost in the Machine*, Picador, London, 1975.
- [42] D. Graur, Y. Zheng, N. Price, R.B. Azevedo, R.A. Zufall, E. Elhaik, On the immortality of television sets: function in the human genome according to the evolution-free gospel of ENCODE, *Genome Biol. Evol.* 5 (2013) 578–590.
- [43] M. Kellis, B. Woldc, M.P. Snyderd, B.E. Bernsteinb, A. Kundajea, G.K. Marinovc, L.D. Ward, E. Birneyg, G.E. Crawford, J. Dekkeri, I. Dunham, L.L. Elnitski, P.J. Farnham, E.A. Feingold, M. Gerstein, M.C. Giddings, D.M. Gilbert, T.R. Gingeras, E.D. Green, R. Guigo, T. Hubbard, J. Kent, J.D. Lieb, R.M. Myers, M.J. Pazin, B. Ren, J.A. Stamatoyannopoulos, Z. Weng, K.P. White, R.C. Hardison, Defining functional DNA elements in the human genome, *Proc. Natl. Acad. Sci. U. S. A.* 111 (2014) 6131–6138.
- [44] I. Schubert, G.T.H. Vu, Genome stability and evolution: attempting a holistic view, *Trends Plant Sci.* 21 (2016) 749–757.
- [45] P. Ly, D.W. Cleveland, Rebuilding chromosomes after catastrophe: emerging mechanisms of chromothripsis, *Trends Cell Biol.* 27 (2017) 917–930.
- [46] C.Z. Zhang, A. Spektor, H. Cornils, J.M. Francis, E.K. Jackson, S. Liu, M. Meyerson, D. Pellman, Chromothripsis from DNA damage in micronuclei, *Nature* 522 (2015) 179–184.
- [47] J. Maciejowski, Y. Li, N. Bosco, P.J. Campbell, T. de Lange, Chromothripsis and kataegis induced by telomere crisis, *Cell* 163 (2015) 1641–1654.
- [48] M. Morishita, T. Muramatsu, Y. Suto, M. Hirai, T. Konishi, S. Hayashi, D. Shigemizu, T. Tsunoda, K. Moriyama, J. Inazawa, Chromothripsis-like chromosomal rearrangements induced by ionizing radiation using proton microbeam irradiation system, *Oncotarget* 7 (2016) 10182–10192.
- [49] W.P. Kloosterman, V. Guryev, M. van Roosmalen, K.J. Duran, E. de Bruijn, S.C. Bakker, T. Letteboer, B. van Nesselrooij, R. Hochstenbach, M. Poot, E. Cuppen, Chromothripsis as a mechanism driving complex de novo structural rearrangements in the germline, *Hum. Mol. Genet.* 20 (2011) 1916–1924.
- [50] M. Fukami, H. Shima, E. Suzuki, T. Ogata, K. Matsubara, T. Kamimaki, Catastrophic cellular events leading to complex chromosomal rearrangements in the germline, *Clin. Genet.* 91 (2017) 653–660.
- [51] R. Blanc-Mathieu, M. Krasovec, M. Hebrard, S. Yau, E. Desgranges, J. Martin, W. Schackwitz, A. Kuo, G. Salin, C. Donnadiou, Y. Desdevises, S. Sanchez-Ferandin, H. Moreau, E. Rivals, I.V. Grigoriev, N. Grimsley, A. Eyre-Walker, G. Piganeau, Population genomics of picophytoplankton unveils novel chromosome hypervariability, *Sci. Adv.* 3 (2017) e1700239.
- [52] E.H. Tan, I.M. Henry, M. Ravi, K.R. Bradnam, T. Mandakova, M.P. Marimuthu, I. Korf, M.A. Lysak, L. Comai, S.W. Chan, Catastrophic chromosomal restructuring during genome elimination in plants, *Elife* 15 (2015) e06516.
- [53] K.A. Simonin, A.B. Roddy, Genome downsizing, physiological novelty, and the global dominance of flowering plants, *PLoS Biol.* 16 (2018) e2003706.