Integrative rDNAomics—Importance of the Oldest Repetitive Fraction of the Eukaryote Genome

Radka Symonová

Faculty of Science, Department of Biology, University of Hradec Králové, 500 03 Hradec Králové, Czech Republic; radka.symonova@gmail.com

Abstract: Nuclear ribosomal RNA (rRNA) genes represent the oldest repetitive fraction universal to all eukaryotic genomes. Their deeply anchored universality and omnipresence during eukaryotic evolution reflects in multiple roles and functions reaching far beyond ribosomal synthesis. Merely the copy number of non-transcribed rRNA genes is involved in mechanisms governing e.g., maintenance of genome integrity and control of cellular aging. Their copy number can vary in response to environmental cues, in cellular stress sensing, in development of cancer and other diseases. While reaching hundreds of copies in humans, there are records of up to 20,000 copies in fish and frogs and even 400,000 copies in ciliates forming thus a literal subgenome or an rDNAome within the genome. From the compositional and evolutionary dynamics viewpoint, the precursor 45S rDNA represents universally GC-enriched, highly recombining and homogenized regions. Hence, it is not accidental that both rDNA sequence and the corresponding rRNA secondary structure belong to established phylogenetic markers broadly used to infer phylogeny on multiple taxonomical levels including species delimitation. However, these multiple roles of rDNAs have been treated and discussed as being separate and independent from each other. Here, I aim to address nuclear rDNAs in an integrative approach to better assess the complexity of rDNA importance in the evolutionary context.

Keywords: nuclear rDNA; rRNA; GC-content; secondary structure; nucleolus

1. The Eukaryotic rDNAome

RNA is essential for information flow from DNA to protein being the dominant macromolecule in protein synthesis [1]. Of the major RNA types, mRNA, tRNA, rRNA, and numerous short non-coding snRNAs, our focus here is on the nuclear rRNA encoded by ribosomal DNA (rDNA), i.e., by rRNA genes. In eukaryotic cells, up to 80% of RNA synthesis belongs to rRNA transcription indispensable to the preservation of ribosome biogenesis and protein synthesis [2]. There are about 1.5–3 million ribosomes per eukaryotic cell [3]. Hence, ribosome biogenesis consumes a tremendous amount of cellular energy and rRNA synthesis is tightly linked to cell growth and proliferation, and as such, it is responsive to general metabolism and environmental challenges [4]. In Eukaryotes, rRNA genes consist of several distinct multigene families tandemly arrayed as repeats composed of tens to hundreds or even thousands of copies. Beside two mitochondrial rRNAs, i.e., the 12S and 16S rRNA in eukaryotes, there are two fractions of nuclear rDNAs—a large, nucleolus-forming 45/47S rDNA unit and a substantially smaller extra-nucleolar 5S rDNA (Figure 1). Both the 45S and 5S rDNAs are organized into clusters of repeats often enabling their cytogenetic visualization on chromosomes [5]. The 5S rDNA can also (co)exist scattered separately within the genome e.g., in the spotted gar as shown in Figure 2a and in other organisms in S1–S4. In budding yeast, the rDNA has a very peculiar organization—the 5S rDNA unit is present in the intergenic spacers (IGSs) of the 45S rDNA and thus, alternating with 45S rDNA units [6]. The coding rDNA sequence is highly conserved among eukaryotes, while the IGSs (Figure 1) that separate the proper units of the 45S rDNA cluster can differ in length and sequence.

In budding yeast, where rDNAs are particularly well-described, IGSs contain three unique elements that are common: an origin of replication, a replication fork blocking site and a promoter that directs the synthesis of noncoding transcripts [7]. In mammals, IGSs contain regulatory regions called UCE (upstream control element), CP (core promoter) and T (termination of transcription site) [8]. Only a fraction of the numerous rDNA copies is transcribed into rRNA. The non-transcribed rDNA copies are extremely important for integrity of the entire genome [7]. In yeast, strains with artificially reduced rDNA copy numbers became sensitive to DNA damage by chemicals and ultraviolet light. This sensitivity further increased as the number of rDNA repeats decreased [7]. In rats, mice, and clawed frog *Xenopus*, the IGS contains one or more RNA polymerase I (Pol I) promoters with high homology to the core region of the main rDNA promoter [9]. Transcripts originating from spacer promoters are co-directional with pre-rRNA synthesis and enhance transcription from the main rDNA promoter, possibly by releasing Pol I [10,11]. Intergenic spacers rRNA have a crucial function in rDNA silencing. In mice, intergenic transcripts originating from a promoter located approximately 2 kb upstream from the pre-rRNA start site are processed into a heterogeneous population of 150–250 nucleotide RNAs, dubbed promoter RNA (pRNA) as their sequence matches the rDNA promoter [6,8,9,11].
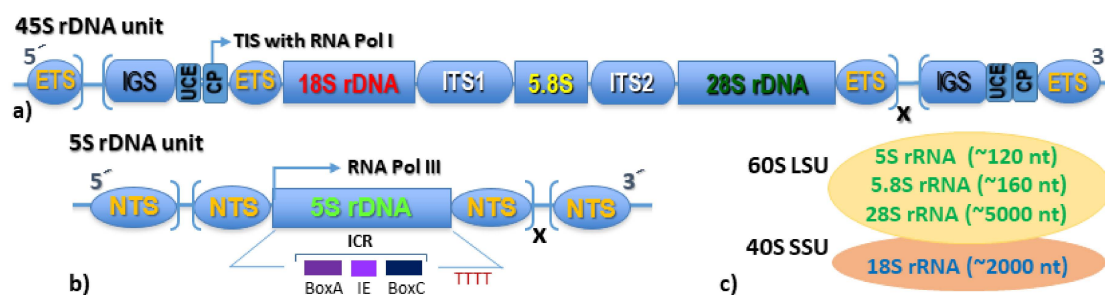

Figure 1. Brief guide to eukaryotic rDNAome—the genomic organization of the rDNA loci. (**a**) Structural organization of the 45S rDNA gene cluster (or rRNA transcription unit); the repeating or single clusters of rDNA can be found scattered throughout genome, they form the precursor pre-rRNA since ribonucleases remove spacers and release separate rRNA molecules in nucleolus—the site of ribosome biogenesis to polysome ribosome formation; (**b**) Structural organization of the 5S rDNA unit (the 5S rDNA can be also dispersed in the genome in many species); (**c**) 80S eukaryotic ribosome composed of the large subunit (LSU) and the small subunit (SSU) with outlined rRNAs. CP—core promoter, ETS—external transcribed spacer, ICR—internal control region, IE—internal element, IGS—intergenic spacer, ITS1, ITS2—internal transcribed spacer 1 and 2, RNA Pol I and III—RNA polymerase I and III, LSU—large (ribosomal) subunit, nt—nucleotides, NTS—non-transcribed spacer, SSU—small subunit, TIS—transcription initiation site, TTTT—polyT transcription termination site, UCE—upstream control element.

2. The Multifaceted Nucleolus

Multiple copies of rRNA gene clusters form nucleolar organizer regions (NORs), the NORs, around which nucleoli are built in the interphase nucleus. The nucleolar 45S rDNAs are transcribed by RNA polymerase I into rRNAs, further processed and assembled with ribosomal proteins into ribosomes [12]. Nucleoli form at the end of mitosis and persist until the onset of the next mitosis. Active nucleoli, where the pre-rRNA transcription takes place, can be visualized in nuclei by silver impregnation, the argyrophilic Ag-NOR staining [13]. From the ultrastructural viewpoint, avian and mammalian nucleoli contain three components (fibrillar centers, dense fibrillar component, and granular component) and differ from all other eukaryotes that possess bipartite nucleoli (i.e., a network of fibrillary strands embedded within granules) [14]. Interestingly, both types of nucleolar arrangement occur among living reptiles: a bicompartmentalized nucleolus in turtles and a tricompartmentalized nucleolus in lizards, crocodiles, and snakes [15]. From the functional viewpoint, the nucleolus was long regarded as a mere ribosome-producing factory. However, during recent decades numerous and crucial non-ribosomal roles were described for the nucleolus [4]. Now, there is a still growing body

of evidence that the nucleolus is central to cellular processes as varied as stress response, cell cycle regulation, RNA modification, cell metabolism, and genome stability and integrity [7]. All organisms sense and respond to stressing conditions by downregulating the transcription of rDNA to rRNA and ribosome biogenesis as these processes are extremely energy-consuming [4].

3. The Nucleolus Forming 45S rDNA

The 45S rDNA transcription unit forms a precursor pre-rRNA consisting of 18S, 5.8S, and 28S rRNAs separated by two internal transcribed spacers (ITS1, ITS2) that are removed during the rRNAs maturation process. The entire unit is further delimited by external transcribed spacers (ETS). Intergenic spacers (IGS; Figure 1a) separate each such unit with both of its sides bearing important regulatory elements [16]. This nomenclature applies to the Animal Kingdom. In plants, there is a 25S rDNA gene (instead of the 28S rDNA) within the large nucleolar rDNA multigene family 35S rDNA (instead of the 45S rDNA [17]). In unicellular organisms, where budding yeasts are the most important model system, the 35S rDNA consists of 25S, 5.8S, and 18S together with the 5S rDNA localized into the intergenic spacer within the 35S rDNA [18]. By addition of some 50–60 ribosomal proteins, the 25S/28S, 5.8S, and 5S rRNAs are fashioned into the large ribosomal subunit, 60S LSU. The 18S rRNA associates with 30–40 ribosomal proteins to form the small ribosomal subunit, 40S SSU (Figure 1c). Molecular cytogenetic localization of the rDNAs 28S rDNA fraction of the 45S rDNA unit on chromosomes is shown on Figure 2b.

The importance of nucleolus-forming rDNA and its proper functioning can be seen in the phenomenon of nucleolar dominance. Nucleolar dominance (or a nucleolus under-development due to expression of rDNA from just one parent) is a dramatic disruption in the formation of nucleoli and epigenetically controlled silencing of 45S rDNA in one of the progenitors in an interspecies hybrid. It is characteristic of some plant and animal distant hybrids and represents an example of a non-mammalian maternal imprinting of 45S rDNA [19,20]. Among animals it has been so far evidenced in details in intra-generic hybrids of *Xenopus* [20,21], in inter-generic hybrids of cyprinid fish [22,23] and in two lines of mouse-human somatic hybrids, where the human ribosomal genes were repressed, and only mouse ribosomal genes were expressed [24]. In the species *Drosophila melanogaster*, a special example of allelic inactivation resembling nucleolar dominance exists [25]. *D. melanogaster* carries its rDNA array on the X and on the Y chromosome [26], but the entire X chromosome rDNA array is normally silenced in *D. melanogaster* males, while the Y chromosome rDNA array is dominant and expressed [25].
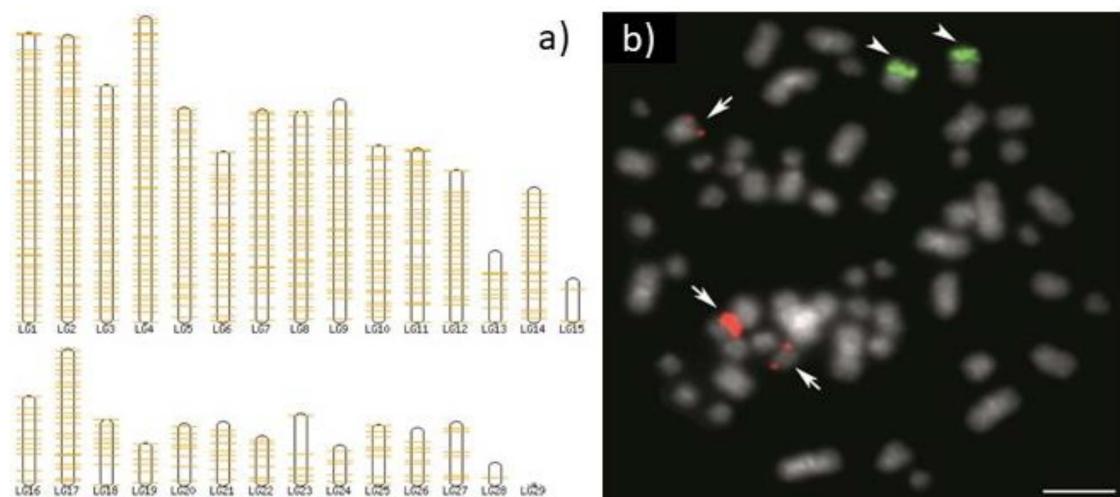
Figure 2. Comparison of two approaches of localization of rDNA on linkage groups and chromosomes in an ancient non-teleost ray-finned fish, spotted gar *(Lepisosteus oculatus)*. (**a**) An in silico approach of visualization of the genomic position of rDNA loci utilizing the Ensembl genome browser tool BioMart to map 5S rDNA on linkage groups (LGs); (**b**) molecular cytogenetic localization of 5S (green, arrowheads) and 28S rDNA (red, arrows) on chromosomes by means of *fluorescence* in situ hybridization (FISH). Bar equals 5 µm (From [27], online Supplementary Material). This comparison shows the sensitivity of the in silico approach. The method enables detection of a single 5S rDNA molecule. It is possible to visualize dispersed molecules across the genome and their pseudogenes in this case. The FISH approach is limited only to huge clusters of accumulated rDNAs and has been utilized for decades particularly in cytotaxonomy in fishes, where most other cytogenetic markers work poorly. Both approaches have their own importance and justification and limits of their mutual interconnection at the current level of genomic data quality—chromosome pairs have not yet been assigned to their corresponding LGs in still too many of the sequenced species. More examples of 5S rDNA localization across LGs are available in Supplementary Materials Figures S1–S4.

4. The Extra-Nucleolar 5S rDNA

The 5S rDNA is much shorter and far less complex within its tandem array structure in comparison with the 45S rDNA. The 5S rDNA consists of a highly conserved sequence of about 120 bp coding for the 5S rRNA and including following functional elements: Box A, IE, Box C [28]. This transcribed sequence is separated at both of its ends from other transcriptional units by a highly variable non-transcribed spacer (NTS; Figure 1b). The participation of 5S rRNA on the ribosome structure is shown on Figure 1c. The 5S rRNA enhances protein synthesis by stabilizing the ribosomal structure and the peptidyl transferase activity, and potentially transmits and coordinates functional centres of the ribosome [29,30]. Two ways for visualization of 5S rDNA sites in silico on linkage groups (LGs) utilizing genomic data and employing methods of molecular cytogenetics on chromosomes are shown in Figure 2. More examples of in silico visualization of 5S rDNA on LGs are in the Supplementary Materials Figures S1–S4.

Two tissue and developmentally specific types of 5S rDNA exist in lower vertebrates, namely the somatic and the oocytes-specific ones. In fish, the oocyte type is lost during development completely (e.g., [31]). Whereas in a frog, it is lost largely [32]. The oocyte repeat comprises a 120 bp oocyte-type 5S rRNA gene placed within the few hundred bp long native AT-rich flanks, whereas the somatic repeat, i.e., a similar 120 bp somatic-type 5S rRNA gene is placed within native GC-rich flanks [33]. Instability of the oocyte 5S rRNA gene transcription complex contributes to the inactivation of the oocyte 5S rRNA gene during embryogenesis [34].

Moreover, in bony and cartilaginous fish, two types of co-occurring 5S rDNA can be distinguished. These probably paralogous type I and II were described in bony fish [35,36] and in elasmobranchs [28]. These types differ by the length and the sequence of their NTS region whereby the longer version is designated as type II [28]. Three functional variants of 5S rRNA genes exist in all life stages of common

sea urchin *Paracentrotus lividus* [37]. Systematic study of 5S rDNA sequence diversity in 97 metazoan species [38] describe several paralogous 5S rDNA sequences in 58 of the examined organisms and a flexible genome organization of 5S rDNA in animals. This study also describes three different types of termination signals and variable distances between the coding regions and the typical termination signal. Importantly, a consensus sequence and secondary structure of metazoan 5S rRNA is presented in this study [38], which can be very useful in more detailed future studies of both 5S rDNA and rRNA.

5. Copy Number Really Matters

Gene duplication is an important and frequent evolutionary process [39] and the resulting copy number variation (CNV) is the most frequent type of genetic variation per base pair in the population [40]. Although alteration of gene copy number or gene dosage has deleterious effects for a significant fraction of the genome, changes in dosage are well tolerated in many genes (reviewed by [41]). CNV of rDNA is highly studied in rDNAomics since it provides a mechanism for cellular homeostasis and for rapid and above all reversible adaptation [42–44]. Due to the tandem repetitive structure of rDNA, the repeat number can be easily reduced by homologous recombination among the repeats. However, there is a finely tuned ´gene amplification system´ compensating for these losses and another highly sophisticated system controlling the ´proper´ rDNA copy number [7]. Moreover, these systems are capable of linking external nutrients availability with rDNA copy number [45] that proves the role of rDNA in the cellular energy metabolism as described for nucleolus above. This illustrates how crucial the right copy number of rDNA is for each cell. Moreover, as uncovered at least in humans, the CNV of rDNA represents a novel and cryptic source of hypervariable genomic diversity with far-reaching global regulatory consequences [42]. However, we have accumulated only limited understanding of the immense importance of these seemingly passive and simple phenomena tightly linked with regulation of nuclear as well as mitochondrial genes expression [42] and probably with many more essential cellular mechanisms. There is an inconsistency in the quantification of rDNA copy number even in the human genome. One important study states that the rDNA copy number varies among healthy humans as a result of natural genetic diversity between 14–410 copies of the 45S rDNA unit per genome [41]. A more recent study reported that the number of rDNA repeats varies from 250 to 670 copies per diploid genome [46]. Therefore, data provided by Gibbons et al. [41,45] should be treated with caution as the low limit number of 14 copies has not been otherwise found in mammals. The genomes of higher eukaryotes harbour hundreds and thousands of copies and only prokaryotic genomes can carry fewer copies of ribosomal genes. Moreover, given that there are five pairs of clusters of ribosomal genes, located on five pairs of human acrocentric chromosomes (13, 14, 15, 21, 22), it is logically impossible that ten clusters could totally count 14 copies, i.e., just 1.4 copies per cluster. These obviously underestimated values can be explained as artifacts caused by a poor suitability of PCR-based techniques for the quantification of GC-rich moderate repeats used by Gibbons et al., 2014, 2015 [42,47] compared to more suitable nonradioactive quantitative hybridization (NQH) used by Chestkov et al., 2018 [46]. The reason is that rDNA is a specific region often forming non-canonical hairpin and loop structures and prone to oxidation in vivo and after extraction from cells [46]. Phenotypic effects of rDNA copy number were recently summarized by [48]. CNV of rDNA (loss as well as amplification) is linked to tumorigenesis [49,50].

A substantial intra-species CNV of rDNA was, among others, evidenced in a freshwater microcrustacean *Daphnia* [51]. rDNA copy number does change among tissues and during ontogenesis in multicellular organisms [31,52] and it is age-dependent in single-cell yeasts [7]. Substantial differences in rDNA CN and number of their sites on chromosomes have been repeatedly recorded in vertebrates and invertebrates on the inter-population and inter-species level of comparison [5]. Such genomic differences might also contribute to genome diversifications, reproductive barriers formation, speciation events and finally to an increased biodiversity, e.g., [53,54]. Numerous examples from fish cytogenetics show that the variation in rDNA repeats prove to be highly informative as it is subject to a more relaxed regulation than in higher vertebrates [55]. Here, traditional cytogenetics meet the currently booming

genomics to mutual usage and benefit from each other. The Animal rDNA database currently contains 539 records on fish rDNAs, namely 5S rDNA in 417 species and 45S rDNA in 479 species [5]. However, a detailed analysis of rDNA sequence organization and variation and CNV on the molecular level exists only for a handful of fish species including both 5S and 45S rDNA of zebrafish [31,51], 5S rDNA and only partial 45S rDNA of pike [56], only 5S rDNA of tilapia [57], molecular organization of the 5S rDNA type II of elasmobranchs (i.e., sharks, rays, and skates [28]) and cichlids [58]. In *Drosophila* germline stem cells, rDNA copy number decreases during aging and this age-dependent decrease in rDNA copy number is transgenerationally heritable. However, young animals are capable of recovering the normal rDNA copy number [59]. The copy number obviously plays a functional role: in *Xenopus*: the somatic 5S rDNA has about 400 copies, while the oocyte 5S rDNA has about 20,000 copies [32]. Locati et al. [31] detected about 9000 5S rRNA genes in the zebrafish genome assembly GRCz10 [31] and Symonova et al. detected about 20,000 copies of 5S rRNA genes in the Northern pike *Esox lucius* and its congener *E. cisalpinus* [56]. However, the record holders are currently protists, namely ciliates: Oligotrichia and Peritrichia [60] and representatives of the ciliate group Spirotrichea - *Oxytricha nova* with about 200,000 rDNA copies [61] and *Stylonychia lemnae* with estimated 400,000 copies of rDNA [62]. The already mentioned single-cell ciliate protozoan *Tetrahymena* amplifies its rDNA 9000-fold during development of the somatic macronucleus [63]. Whereas the copy number of 45S and 5S rDNA units is tightly coupled in mouse and human [47], such a control is apparently missing in fish [52,56,64]. This fact together with the aforementioned difference in nucleolar organization between higher and lower vertebrates and also other genomic traits (e.g., genomic GC heterogeneity) indicate that another major evolutionary transition *sensu* [65] occurred in evolution from anamniotes towards amniotes. This huge copy number variation might be linked to (or might have resulted in) the heterogeneity in rRNA genes and their variants that had been considered a peculiarity of some plants. Only recently, this heterogeneity was proved also in animal ribosomal genes, including human and mouse, where variant rRNA alleles exhibit tissue-specific expression and ribosomes bearing variant rRNA alleles are present in the actively translating ribosome pool [66].

One special topic of the rDNA CNV is based on molecular cytogenetic localization of both 5S and 45S rDNAs using FISH. FISH with rDNAs represents one of the most important chromosomal markers particularly in non-model organisms and especially in cold-blooded vertebrates, where methods like R-banding do not yield any usable and reproducible pattern. For these reasons, a heavy body of literature on molecular cytogenetics of rDNA has accumulated (for plants [17], for animals [5]). Since rDNA was omitted from many genome sequencing projects due to issues with its assembling [50,67], any precise quantification of rDNA copies is mostly still impossible. On the other hand, the still increasing availability of long-read sequencing can overcome assembling issues and provides opportunity to link the numerous results from molecular cytogenetics with genomics as was successfully demonstrated in fish cytogenomics [54,56].

A very special chapter of the rDNAomics book deals with rDNA of eukaryotic microorganisms [67]. In their 2010 review, Torres-Machorro et al. present available information on both rDNA fractions from about hundred microbial eukaryotes and show an unexpected diversity in their genomic organization [68]. Later, Drouin and Tsang [69] focus their review of 5S rDNA in protists on adaptive potential of its organization. Microbial eukaryotic rDNAs may be coded alone, in tandem repeats, linked to each other or linked to other genes. They exist in the chromosome or extrachromosomally in linear or circular units and rDNA coding regions may contain introns, sequence insertions, protein-coding genes, or additional spacers [68]. The atypical structures of rDNA have been considered as exceptions. However, it is rather likely that these organisms have preserved variations in the organization of these versatile genes that may be considered as living records of evolution [68]. A huge step in establishing the functional significance of rDNA in evolution and in ecology of organisms has been performed in protists [60].

6. Overview of Important Facts about rDNA

The most import facts about rDNA can be summarized as follows: rDNA is ubiquitous and universal across Prokaryotes, Archaea, and Eukaryotes [70]. It has a high degree of functional and sequence conservation of rDNA genes [71]. At the same time, rDNA belongs to the most copy number-hypervariable genomic segments [42] and the tandemly repeated rDNA arrays are among the most evolutionary dynamic loci of eukaryotic genomes in terms of copy number. Due to its heavy transcription, repetitive structure, and programmed replication fork pauses, the rDNA is one of the most unstable regions in the genome [7,18]. Their high genomic copy number relative to other genes appears to be much larger than required, however, unlike protein-coding genes, rDNA cannot undergo additional rounds of amplification via translation when organisms require more rRNA transcripts [72]. Copy number of 45S units is balanced with that of the 5S rDNA in mouse and human [47] but not in fish, summarized by [5]. These multiple copies of rDNA evolve in a highly coordinated manner, through unequal crossing over and/or gene conversion, two mechanisms related to homologous recombination [73]. The rRNA gene repeats use a unique gene amplification system to restore the copy number after this has been reduced due to recombination [7]. The RFB (replication fork barrier) coordinates replication and recombination, and through the latter, mediates a possible increase in the number of rDNA repeats. rDNA loci are dynamic genetic elements, their copy number changes dynamically and transgenerationally yet is maintained through a recovery mechanism in the germline (for *Drosophila* see [59]). In plants, extensive variation can exist in both rDNA copy number and rRNA expression. Among maize inbred lines, thousands of genes co-regulate with rRNA expression, including genes participating in ribosome biogenesis and other functionally relevant pathways [74]. Not only the rDNA copy number [45] but also the rRNA expression variation is a valuable source of functional diversity that affects gene expression variation and field-based phenotypic changes [74]. The intra-genomic homogenization of rDNA mostly occurs through 'concerted evolution' [75]. rDNA also shows high rates of meiotic recombination [75,76] and rDNA sites are hotspots for genome rearrangements [77]. Copy number of rDNA arrays modulates genome-wide expression of hundreds to thousands of genes and subtle changes in rDNA copy number between individuals may contribute to biologically relevant phenotypic variation also in humans [78]. rDNA contributes to global chromatin regulation and thus to a balance between heterochromatin and euchromatin in the nucleus [79]. The enormous variation in the number of rDNA copies per eukaryotic genome correlates with genome size [80] and the copy number of the 45S rDNA fraction was shown to negatively correlate with mtDNA abundance [42]. Hence, rDNA copy number variation, CNV ("rDNA dosage") is a major determinant of naturally occurring genome-wide gene expression variation in humans [42]. Ribosomal RNAs (rRNAs) account for up to 80% of all RNAs in eukaryotic cells [50]. Growth-activated rRNA synthesis may be mediated by the up-regulation of individual rDNA units, in addition to the activation of silent gene copies see e.g., Banditt et al. [81]. In mammals, 5S and 45S rDNA arrays are non-homologous, physically unlinked, transcribed by different RNA Polymerases and encode functionally interdependent RNA components of the ribosome [47]. Clusters of the 45S rDNA unit give origin to the nucleolus, the nuclear organelle that is the site of pre-45S rDNA transcription and ribosome biogenesis, see e.g., [8,82]. The rDNA contact map shows that 5S and 45S arrays each have thousands of contacts in the folded genome, with rDNA-associated regions and genes dispersed across all chromosomes [83,84]. Due to its highly repetitive nature, rDNA has been excluded from most mammalian genome-wide studies because of challenges associated with its analysis, and thus remains understudied. There is an unusual and universal GC richness of the 45S rDNA fraction in cold- as well as warm-blooded vertebrates (more details below) [72,84]. An extensive range of epigenetic modifications regulating rRNA genes transcription [67,85–87] results in only a mere subset of the multiple copies being transcribed however with far reaching implications for the entire genome (elucidation of the epigenetics of rDNA in sufficient detail would require a lot of research). On top of it, rDNA loci serve as a specialized niche for mobile elements [88].

rDNA units (so called 'rDNA-like signal') can be found scattered throughout the genome in humans [89]. These units can be described as follows: 1) highly degraded, but near full length, rDNA units, including both 45S and Intergenic Spacer (IGS), can be found at multiple sites in the human genome on chromosomes without rDNA arrays; 2) these rDNA sequences have a propensity for being centromere proximal; and 3) sequence at all human functional rDNA array ends is divergent from canonical rDNA to the point that it is pseudogenic. For this in fish, see Figure 2a and in other chordates see the Supplementary Materials Figures S1–S4.

rDNA represents a cryptic source of hypervariable genomic diversity with global regulatory consequences (ribosomal quantitative trait loci (eQTL)) in humans. The variation provides a mechanism for cellular homeostasis and for rapid and reversible adaptation [42,47].

7. GC Content of rDNA

The 45S rDNA gene clusters form the GC-richest genomic fraction particularly in Eukaryotes [90] with humans having 60%–80% GC in different parts of the rDNA [91], whereas the median genomic GC is 40.9% (NCBI, human genome assembly). This GC-richness is ascribed to the recombination rate based process known as GC-biased gene conversion [73]. On the other hand, some studies link the extremely high GC levels in rDNAs to particular requirements for stem-and-loop systems in rRNA that have an effect on the composition of the corresponding genes to thermal adaptation [92]. This line of explanations belongs to the Thermodynamic Stability Hypothesis attempting to account for the overall AT/GC heterogeneity in birds and mammals and the AT/GC homogeneity in the remaining vertebrates and the other Eukaryotes [90]. However, although Wang et al., 2006 showed support for such a thermal adaptation in Bacteria and Archaea, they did not find any for warm-blooded birds and mammals with only a slightly higher GC content of 18S (55.7%) versus cold-blooded fishes and amphibians with approximately 53.5% of GC. Their partitioning of the GC content across the 18S rRNA sequences into stem and loop regions demonstrated [93] that the differences are not concentrated in the paired stem regions as expected by Bernardi [90]. Interesting and relevant aspects of rDNA GC content exist in so-called expansion segments (ES) in 28S and 18S rRNA molecules [94,95]. Expansion of the 28S rRNA shows a clear phylogenetic increase, with a dramatic rise in mammals and especially in hominids. Here, a GC- or AU-biased expansion of rRNAs has developed in both plants and metazoans, with the GC-bias largely being preferred in extremely GC-rich ES of vertebrate 28S rRNA. This compositional bias towards GC is linked to potential roles of GC-rich rRNA during protein synthesis [96,97] and could contribute to the discussion whether the genomic GC content is driven by neutral versus selective processes. An interesting explanation of the universal GC richness of 45S rDNA comes from the GC biology—these multicopy genes should all be in a DNA region with a homogenous GC composition to allow concerted evolution and to prevent divergence through generations [98].

8. Phylogeny, Species Delimitation, and Secondary Structure of rRNAs—The Way How to Determine in Silico Whether Two Lineages Can Successfully Cross

The ITS2 sequence already belongs to the most popular and well established phylogenetic and DNA barcoding markers [99]. The rDNA sequence and its corresponding rRNA secondary structure is one of the few universal features of life without any known case of horizontal transfer and above all, identifying the organism to a unique species, making it uniquely suited to assess phylogenetic relationships [100,101]. The secondary structure of the ITS regions is well known for a wide variety of eukaryotes and have been used to aid in the alignment of these sequences for phylogenetic comparisons [101]. The RNA sequence of the ITS2 possesses another special trait so far not fully examined, namely compensatory base changes (CBCs, Figure 3). CBCs are mutations occurring simultaneously on both sides of a nucleotide pair in the ITS2 secondary structure with retention of the paired nucleotide bond, whereas hemi-CBC is a mutation of a single nucleotide of the pair still retaining the bond [102]. CBC analyses have been primarily performed in fungi and plants [92–94]. This is the reason why the majority of literature references, including methods descriptions, are on

plants (e.g., estimating structure-based phylogenetic trees from ITS2 data by [103–105]). CBCs analyses have been already successfully used to verify taxonomy of closely related species and to distinguish morphologically indistinct species in insects [102]. That shows the huge potential of mining for CBCs in ITS2 rRNA secondary structures also in the Animal Kingdom.

Figure 3. Visualization of compensatory base changes (CBCs) on a hypothetical internal transcribed spacer 2 (ITS2) secondary structure. The left panel depicts a conserved helix segment in which a CBC occurs, where both nucleotides of the pair underwent a mutation that resulted in retaining the paired nucleotide bond. The right panel shows a hemi-CBC, where only one nucleotide of the pair, i.e., C to A, underwent a mutation while the pair retained the nucleotide bond. Redrawn according to [106].

However, any detailed and particularly systematic survey of rRNA secondary structure in the Animal Kingdom is still in its infancy although it would be highly desirable in numerous areas of biology. Moreover, analyses of rRNA secondary structure could represent another intersection between rDNAomics based on molecular cytogenetics and on genomics since molecular cytogenetic studies frequently provide DNA sequences of rDNA fragments used in FISH experiments and further DNA sequences, in the meanwhile, became available in NCBI GenBank or could be retrieved from whole-genome datasets. Ideally, such integrative studies could contain cytogenetic results accompanied by details on rDNA/rRNA sequence and rRNA secondary structure as shown in Figure 4 to better explore any potential sequence polymorphism.
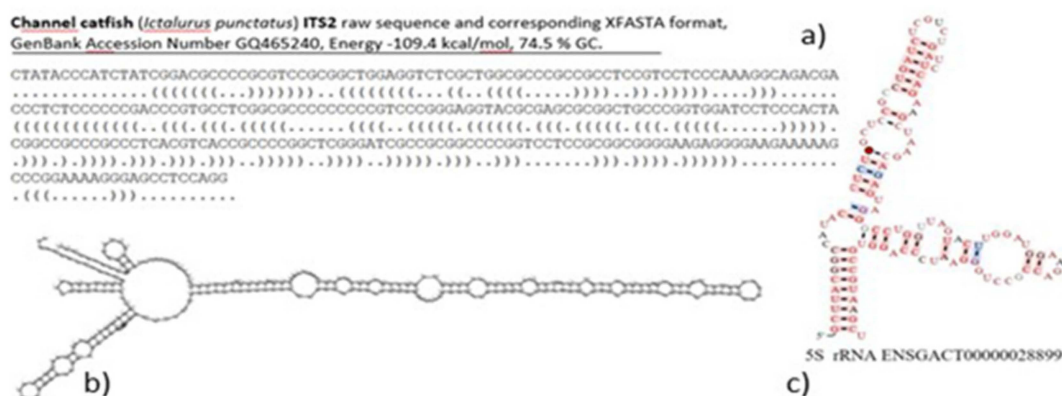

Figure 4. Text and graphic representation of prediction of rRNA secondary structure. (**a**) the Xfasta or "dot-bracket notation" way of representation of the ITS2 rRNA secondary structure of channel catfish, (**b**) visualization of the corresponding secondary structure of the sequence in (**a**) with one longest helix and four short helices, (**c**) stickleback, prediction of 5S rRNA secondary structure, Ensemble.

9. Concluding Remarks

The field of rDNAomics is extremely rapidly evolving further and has far-reaching implications for numerous areas of current biology and medicine. Medical aspects represent another crucial chapter

of rDNAomics that already exceeds the scope of this review. On the other hand, being aware of this fact might help scientists from the area of fundamental research working on non-model organisms to provide justification of their work. There are numerous diseases associated with rDNA dysfunction, particularly cancer [39,40,61,66]. Ribosomopathies are diseases caused by abnormalities in the structure or function of ribosomal component proteins or rRNA genes, or other genes whose products are involved in ribosome biogenesis [107]. Not only sequence, but also copy number of rDNAs is of particular importance in cancer—human cancer genomes show a loss of copies, accompanied by global copy number co-variation [50]. Even more relevant is the fact that rDNA repeat instability coincides with predisposition to cancer, premature aging and neurological impairment in ataxia-telangiectasia and Bloom syndrome (Warmerdam and Wolthuis, 2018). Additionally, it was shown that cancers undergo coupled 5S rDNA array expansion and 45S rDNA loss that is accompanied by increased proliferation rate and nucleolar activity. Somatic changes in rDNA copy number can exceed 10-fold the naturally occurring copy number variation across individuals [49]. Malfunction of nucleoli can be the cause of several human conditions called nucleolopathies [108]. The nucleolus is being investigated as a target for cancer chemotherapy [109,110]. Moreover, rDNA copy number may be a simple and useful indicator of whether a cancer will be sensitive to DNA damaging treatments [50]. Hence, it is desirable to understand rDNA organization, function, and its impact on the entire nucleolus and other genes' regulation (as briefly outlined here) in a broader evolutionary context.

Supplementary Materials: The following are available online at http://www.mdpi.com/2073-4425/10/5/345/s1, Figure S1: *Tetraodon nigroviridis*, karyogram with labeled 5S rDNA loci (red arrowheads). Figure S2: Zebrafish (*Danio rerio*), complete karyogram with labeled 5S rDNA loci (pink lines). Figure S3: A tunicate sea squirt (*Ciona intestinalis*), C ≈ 0.2, the thirty-five 5S rDNA site detected on four chromosomes could be assigned to a linkage group. Figure S4: Human genome illustrates here the best assembled vertebrate genome, C ≈ 3.5. Genomic localization of annotated 5S rDNA showing 5S rDNA scattered throughout fish genomes visualized on karyograms of species assembled to the chromosome level (5S rDNA sequences were filtered using the BioMart tool and the Ensemble Genes Database version 92 from Ensembl.org version 92). For comparison, two model mammalian genomes (Hsa, Mmu) and one avian genome are shown.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Bernhardt, H.S.; Tate, W.P. A Ribosome Without RNA. *Front. Ecol. Evol.* **2015**, *3*, 1–6. [CrossRef]

2. Boisvert, F.-M.; van Koningsbruggen, S.; Navascués, J.; Lamond, A.I. The multifunctional nucleolus. *Nat. Rev. Mol. Cell Biol.* **2007**, *8*, 574–585. [CrossRef] [PubMed]

3. Porokhovnik, L.; Gerton, J.L. Ribosomal DNA-connecting ribosome biogenesis and chromosome biology. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **2019**. [CrossRef] [PubMed]

4. Grummt, I. The nucleolus—guardian of cellular homeostasis and genome integrity. *Chromosoma* **2013**, *122*, 487–497. [CrossRef] [PubMed]

5. Sochorová, J.; Garcia, S.; Gálvez, F.; Symonová, R.; Kovařík, A. Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database. *Chromosoma* **2018**, *127*, 141–150. [CrossRef]

6. James, S.A.; O'Kelly, M.J.T.; Carter, D.M.; Davey, R.P.; van Oudenaarden, A.; Roberts, I.N. Repetitive sequence variation and dynamics in the ribosomal DNA array of Saccharomyces cerevisiae as revealed by whole-genome resequencing. *Genome Res.* **2009**, *19*, 626–635. [CrossRef] [PubMed]

7. Kobayashi, T. Ribosomal RNA gene repeats, their stability and cellular senescence. *Proc. Jpn. Acad. Ser. B* **2014**, *90*, 119–129. [CrossRef]

8. Russell, J.; Zomerdijk, J.C.B.M. RNA-polymerase-I-directed rDNA transcription, life and works. *Trends Biochem. Sci.* **2005**, *30*, 87–96. [CrossRef] [PubMed]

9. Santoro, R.; Schmitz, K.-M.; Sandoval, J.; Grummt, I. Intergenic transcripts originating from a subclass of ribosomal DNA repeats silence ribosomal RNA genes in trans. *EMBO Rep.* **2010**, *11*, 52–58. [CrossRef] [PubMed]

10. Grimaldi, G.; Di Nocera, P.P. Multiple repeated units in Drosophila melanogaster ribosomal DNA spacer stimulate rRNA precursor transcription. *Proc. Natl. Acad. Sci. USA* **1988**, *85*, 5502–5506. [CrossRef]

11. Mayer, C.; Neubert, M.; Grummt, I. The structure of NoRC-associated RNA is crucial for targeting the chromatin remodelling complex NoRC to the nucleolus. *EMBO Rep.* **2008**, *9*, 774–780. [CrossRef]

12. Németh, A.; Längst, G. Genome organization in and around the nucleolus. *Trends Genet. TIG* **2011**, *27*, 149–156. [CrossRef] [PubMed]

13. Howell, W.M.; Black, D.A. Controlled silver-staining of nucleolus organizer regions with a protective colloidal developer: A 1-step method. *Experientia* **1980**, *36*, 1014–1015. [CrossRef] [PubMed]

14. Thiry, M.; Lafontaine, D.L.J. Birth of a nucleolus: The evolution of nucleolar compartments. *Trends Cell Biol.* **2005**, *15*, 194–199. [CrossRef]

15. Thiry, M.; Lamaye, F.; Lafontaine, D.L.J. The nucleolus: When 2 became 3. *Nucleus* **2011**, *2*, 289–293. [CrossRef]

16. Henras, A.K.; Plisson-Chastang, C.; O'Donohue, M.-F.; Chakraborty, A.; Gleizes, P.-E. An overview of pre-ribosomal RNA processing in eukaryotes. *Wiley Interdiscip. Rev. RNA* **2015**, *6*, 225–242. [CrossRef] [PubMed]

17. Garcia, S.; Garnatje, T.; Kovařík, A. Plant rDNA database: Ribosomal DNA loci information goes online. *Chromosoma* **2012**, *121*, 389–394. [CrossRef] [PubMed]

18. Kobayashi, T.; Sasaki, M. Ribosomal DNA stability is supported by many 'buffer genes'—introduction to the Yeast rDNA Stability Database. *FEMS Yeast Res.* **2017**, *17*. [CrossRef]

19. Pikaard, C.S. Nucleolar dominance: Uniparental gene silencing on a multi-megabase scale in genetic hybrids. *Plant Mol. Biol.* **2000**, *43*, 163–177. [CrossRef] [PubMed]

20. Michalak, K.; Maciak, S.; Kim, Y.B.; Santopietro, G.; Oh, J.H.; Kang, L.; Garner, H.R.; Michalak, P. Nucleolar dominance and maternal control of 45S rDNA expression. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*, 20152201. [CrossRef]

21. Maciak, S.; Michalak, K.; Kale, S.D.; Michalak, P. Nucleolar Dominance and Repression of 45S Ribosomal RNA Genes in Hybrids between *Xenopus borealis* and *X. muelleri* (2n = 36). *Cytogenet. Genome Res.* **2016**, *149*, 290–296. [CrossRef] [PubMed]

22. Xiao, J.; Hu, F.; Luo, K.; Li, W.; Liu, S. Unique nucleolar dominance patterns in distant hybrid lineage derived from Megalobrama Amblycephala × Culter Alburnus. *BMC Genet.* **2016**, *17*. [CrossRef]

23. Cao, L.; Qin, Q.; Xiao, Q.; Yin, H.; Wen, J.; Liu, Q.; Huang, X.; Huo, Y.; Tao, M.; Zhang, C.; et al. Nucleolar Dominance in a Tetraploidy Hybrid Lineage Derived From Carassius auratus red var. () × Megalobrama amblycephala (). *Front. Genet.* **2018**, *9*. [CrossRef] [PubMed]

24. Onishi, T.; Berglund, C.; Reeder, R.H. On the mechanism of nucleolar dominance in mouse-human somatic cell hybrids. *Proc. Natl. Acad. Sci. USA* **1984**, *81*, 484–487. [CrossRef] [PubMed]

25. Greil, F.; Ahmad, K. Nucleolar Dominance of the *Y* Chromosome in *Drosophila melanogaster*. *Genetics* **2012**, *191*, 1119–1128. [CrossRef] [PubMed]

26. Tautz, D.; Hancock, J.M.; Webb, D.A.; Tautz, C.; Dover, G.A. Complete sequences of the rRNA genes of Drosophila melanogaster. *Mol. Biol. Evol.* **1988**, *5*, 366–376. [PubMed]

27. Symonová, R.; Majtánová, Z.; Arias-Rodriguez, L.; Mořkovský, L.; Kořínková, T.; Cavin, L.; Pokorná, M.J.; Doležálková, M.; Flajšhans, M.; Normandeau, E.; et al. Genome Compositional Organization in Gars Shows More Similarities to Mammals than to Other Ray-Finned Fish: CYTOGENOMICS OF GARS. *J. Exp. Zoolog. B Mol. Dev. Evol.* **2017**, *328*, 607–619. [CrossRef]

28. Castro, S.I.; Hleap, J.S.; Cárdenas, H.; Blouin, C. Molecular organization of the 5S rDNA gene type II in elasmobranchs. *RNA Biol.* **2016**, *13*, 391–399. [CrossRef]

29. Szymanski, M.; Barciszewska, M.Z.; Erdmann, V.A.; Barciszewski, J. 5S Ribosomal RNA Database. *Nucleic Acids Res.* **2002**, *30*, 176–178. [CrossRef] [PubMed]

30. Dinman, J.D. 5S rRNA: Structure and Function from Head to Toe. *Int. J. Biomed. Sci. IJBS* **2005**, *1*, 2–7.

31. Locati, M.D.; Pagano, J.F.B.; Ensink, W.A.; van Olst, M.; van Leeuwen, S.; Nehrdich, U.; Zhu, K.; Spaink, H.P.; Girard, G.; Rauwerda, H.; et al. Linking maternal and somatic 5S rRNA types with different sequence-specific non-LTR retrotransposons. *RNA* **2017**, *23*, 446–456. [CrossRef]

32. Peterson, R.C.; Doering, J.L.; Brown, D.D. Characterization of two xenopus somatic 5S DNAs and one minor oocyte-specific 5S DNA. *Cell* **1980**, *20*, 131–141. [CrossRef]

33. Tomaszewski, R.; Jerzmanowski, A. The AT-rich flanks of the oocyte-type 5S RNA gene of Xenopus laevis act as a strong local signal for histone H1-mediated chromatin reorganization in vitro. *Nucleic Acids Res.* **1997**, *25*, 458–466. [CrossRef] [PubMed]

34. Wolffe, A.P.; Brown, D.D. Developmental regulation of two 5S ribosomal RNA genes. *Science* **1988**, *241*, 1626–1632. [CrossRef]

35. Martins, C.; Galetti, P.M. Two 5S rDNA arrays in neotropical fish species: is it a general rule for fishes? *Genetica* **2001**, *111*, 439–446. [CrossRef] [PubMed]

36. *Fish Cytogenetics*; Pisano, E. (Ed.) Science Publishers: Enfield, NH, USA, 2007; ISBN 978-1-57808-330-5.

37. Dimarco, E.; Cascone, E.; Bellavia, D.; Caradonna, F. Functional variants of 5S rRNA in the ribosomes of common sea urchin Paracentrotus lividus. *Gene* **2012**, *508*, 21–25. [CrossRef] [PubMed]

38. Vierna, J.; Wehner, S.; Höner zu Siederdissen, C.; Martínez-Lage, A.; Marz, M. Systematic analysis and evolution of 5S ribosomal DNA in metazoans. *Heredity* **2013**, *111*, 410–421. [CrossRef] [PubMed]

39. Conrad, B.; Antonarakis, S.E. Gene Duplication: A Drive for Phenotypic Diversity and Cause of Human Disease. *Annu. Rev. Genomics Hum. Genet.* **2007**, *8*, 17–35. [CrossRef]

40. The Wellcome Trust Case Control Consortium; Conrad, D.F.; Pinto, D.; Redon, R.; Feuk, L.; Gokcumen, O.; Zhang, Y.; Aerts, J.; Andrews, T.D.; Barnes, C.; et al. Origins and functional impact of copy number variation in the human genome. *Nature* **2010**, *464*, 704–712. [CrossRef]

41. Rice, A.M.; McLysaght, A. Dosage sensitivity is a major determinant of human copy number variant pathogenicity. *Nat. Commun.* **2017**, *8*, 14366. [CrossRef] [PubMed]

42. Gibbons, J.G.; Branco, A.T.; Yu, S.; Lemos, B. Ribosomal DNA copy number is coupled with gene expression variation and mitochondrial abundance in humans. *Nat. Commun.* **2014**, *5*. [CrossRef]

43. Long, E.O.; Dawid, I.B. Repeated genes in eukaryotes. *Annu. Rev. Biochem.* **1980**, *49*, 727–764. [CrossRef] [PubMed]

44. Oakes, M.; Siddiqi, I.; Vu, L.; Aris, J.; Nomura, M. Transcription factor UAF, expansion and contraction of ribosomal DNA (rDNA) repeats, and RNA polymerase switch in transcription of yeast rDNA. *Mol. Cell. Biol.* **1999**, *19*, 8559–8569. [CrossRef]

45. Jack, C.V.; Cruz, C.; Hull, R.M.; Keller, M.A.; Ralser, M.; Houseley, J. Regulation of ribosomal DNA amplification by the TOR pathway. *Proc. Natl. Acad. Sci.* **2015**, *112*, 9674–9679. [CrossRef] [PubMed]

46. Chestkov, I.V.; Jestkova, E.M.; Ershova, E.S.; Golimbet, V.E.; Lezheiko, T.V.; Kolesina, N.Y.; Porokhovnik, L.N.; Lyapunova, N.A.; Izhevskaya, V.L.; Kutsev, S.I.; et al. Abundance of ribosomal RNA gene copies in the genomes of schizophrenia patients. *Schizophr. Res.* **2018**, *197*, 305–314. [CrossRef] [PubMed]

47. Gibbons, J.G.; Branco, A.T.; Godinho, S.A.; Yu, S.; Lemos, B. Concerted copy number variation balances ribosomal DNA dosage in human and mouse genomes. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 2485–2490. [CrossRef]

48. Porokhovnik, L.N.; Lyapunova, N.A. Dosage effects of human ribosomal genes (rDNA) in health and disease. *Chromosome Res.* **2019**, *27*, 5–17. [CrossRef]

49. Wang, M.; Lemos, B. Ribosomal DNA copy number amplification and loss in human cancers is linked to tumor genetic context, nucleolus activity, and proliferation. *PLoS Genet.* **2017**, *13*, e1006994. [CrossRef]

50. Xu, B.; Li, H.; Perry, J.M.; Singh, V.P.; Unruh, J.; Yu, Z.; Zakari, M.; McDowell, W.; Li, L.; Gerton, J.L. Ribosomal DNA copy number loss and sequence variation in cancer. *PLoS Genet.* **2017**, *13*, e1006771. [CrossRef] [PubMed]

51. Eagle, S.H.; Crease, T.J. Copy number variation of ribosomal DNA and Pokey transposons in natural populations of Daphnia. *Mob. DNA* **2012**, *3*, 4. [CrossRef] [PubMed]

52. Locati, M.D.; Pagano, J.F.B.; Girard, G.; Ensink, W.A.; van Olst, M.; van Leeuwen, S.; Nehrdich, U.; Spaink, H.P.; Rauwerda, H.; Jonker, M.J.; et al. Expression of distinct maternal and somatic 5.8S, 18S, and 28S rRNA types during zebrafish development. *RNA* **2017**, *23*, 1188–1199. [CrossRef] [PubMed]

53. Dion-Cote, A.-M.; Symonova, R.; Rab, P.; Bernatchez, L. Reproductive isolation in a nascent species pair is associated with aneuploidy in hybrid offspring. *Proc. R. Soc. B Biol. Sci.* **2015**, *282*, 20142862. [CrossRef] [PubMed]

54. Dion-Côté, A.-M.; Symonová, R.; Lamaze, F.C.; Pelikánová, Š.; Ráb, P.; Bernatchez, L. Standing chromosomal variation in Lake Whitefish species pairs: The role of historical contingency and relevance for speciation. *Mol. Ecol.* **2017**, *26*, 178–192. [CrossRef]

55. Symonová, R.; Howell, W. Vertebrate Genome Evolution in the Light of Fish Cytogenomics and rDNAomics. *Genes* **2018**, *9*, 96. [CrossRef]

56. Symonová, R.; Ocalewicz, K.; Kirtiklis, L.; Delmastro, G.B.; Pelikánová, Š.; Garcia, S.; Kovařík, A. Higher-order organisation of extremely amplified, potentially functional and massively methylated 5S rDNA in European pikes (Esox sp.). *BMC Genomics* **2017**, *18*. [CrossRef]

57. Martins, C.; Wasko, A.P.; Oliveira, C.; Porto-Foresti, F.; Parise-Maltempi, P.P.; Wright, J.M.; Foresti, F. Dynamics of 5S rDNA in the tilapia *(Oreochromis niloticus)* genome: Repeat units, inverted sequences, pseudogenes and chromosome loci. *Cytogenet. Genome Res.* **2002**, *98*, 78–85. [CrossRef] [PubMed]

58. Nakajima, R.T.; Cabral-de-Mello, D.C.; Valente, G.T.; Venere, P.C.; Martins, C. Evolutionary dynamics of rRNA gene clusters in cichlid fish. *BMC Evol. Biol.* **2012**, *12*, 198. [CrossRef]

59. Lu, K.L.; Nelson, J.O.; Watase, G.J.; Warsinger-Pepe, N.; Yamashita, Y.M. Transgenerational dynamics of rDNA copy number in Drosophila male germline stem cells. *eLife* **2018**, *7*. [CrossRef] [PubMed]

60. Gong, J.; Dong, J.; Liu, X.; Massana, R. Extremely High Copy Numbers and Polymorphisms of the rDNA Operon Estimated from Single Cell Analysis of Oligotrich and Peritrich Ciliates. *Protist* **2013**, *164*, 369–379. [CrossRef]

61. Prescott, D.M. The DNA of ciliated protozoa. *Microbiol. Rev.* **1994**, *58*, 233–267. [CrossRef] [PubMed]

62. Heyse, G.; Jönsson, F.; Chang, W.-J.; Lipps, H.J. RNA-dependent control of gene amplification. *Proc. Natl. Acad. Sci. USA* **2010**, *107*, 22134–22139. [CrossRef]

63. Pan, W.-C.; Orias, E.; Flacks, M.; Blackburn, E.H. Allele-specific, selective amplification of a ribosomal RNA gene in tetrahymena thermophila. *Cell* **1982**, *28*, 595–604. [CrossRef]

64. Symonová, R.; Majtánová, Z.; Sember, A.; Staaks, G.B.; Bohlen, J.; Freyhof, J.; Rábová, M.; Ráb, P. Genome differentiation in a species pair of coregonine fishes: An extremely rapid speciation driven by stress-activated retrotransposons mediating extensive ribosomal DNA multiplications. *BMC Evol. Biol.* **2013**, *13*, 42. [CrossRef]

65. Szathmáry, E.; Smith, J.M. The major evolutionary transitions. *Nature* **1995**, *374*, 227–232. [CrossRef] [PubMed]

66. Parks, M.M.; Kurylo, C.M.; Dass, R.A.; Bojmar, L.; Lyden, D.; Vincent, C.T.; Blanchard, S.C. Variant ribosomal RNA alleles are conserved and exhibit tissue-specific expression. *Sci. Adv.* **2018**, *4*, eaao0665. [CrossRef]

67. Bughio, F.; Maggert, K.A. The peculiar genetics of the ribosomal DNA blurs the boundaries of transgenerational epigenetic inheritance. *Chromosome Res. Int. J. Mol. Supramol. Evol. Asp. Chromosome Biol.* **2018**. [CrossRef]

68. Torres-Machorro, A.L.; Hernández, R.; Cevallos, A.M.; López-Villaseñor, I. Ribosomal RNA genes in eukaryotic microorganisms: Witnesses of phylogeny? *FEMS Microbiol. Rev.* **2010**, *34*, 59–86. [CrossRef]

69. Drouin, G.; Tsang, C. 5S rRNA Gene Arrangements in Protists: A Case of Nonadaptive Evolution. *J. Mol. Evol.* **2012**, *74*, 342–351. [CrossRef]

70. Mallatt, J.; Chittenden, K.D. The GC content of LSU rRNA evolves across topological and functional regions of the ribosome in all three domains of life. *Mol. Phylogenet. Evol.* **2014**, *72*, 17–30. [CrossRef] [PubMed]

71. Agrawal, S.; Ganley, A.R.D. The conservation landscape of the human ribosomal RNA gene repeats. *PLOS ONE* **2018**, *13*, e0207531. [CrossRef] [PubMed]

72. Kobayashi, T. A new role of the rDNA and nucleolus in the nucleus—rDNA instability maintains genome integrity. *BioEssays* **2008**, *30*, 267–272. [CrossRef]

73. Escobar, J.S.; Glémin, S.; Galtier, N. GC-Biased Gene Conversion Impacts Ribosomal DNA Evolution in Vertebrates, Angiosperms, and Other Eukaryotes. *Mol. Biol. Evol.* **2011**, *28*, 2561–2575. [CrossRef]

74. Li, B.; Kremling, K.A.G.; Wu, P.; Bukowski, R.; Romay, M.C.; Xie, E.; Buckler, E.S.; Chen, M. Coregulation of ribosomal RNA with hundreds of genes contributes to phenotypic variation. *Genome Res.* **2018**, *28*, 1555–1565. [CrossRef]

75. Ganley, A.R.D.; Kobayashi, T. Highly efficient concerted evolution in the ribosomal DNA repeats: Total rDNA repeat variation revealed by whole-genome shotgun sequence data. *Genome Res.* **2007**, *17*, 184–191. [CrossRef]

76. Stults, D.M.; Killen, M.W.; Pierce, H.H.; Pierce, A.J. Genomic architecture and inheritance of human ribosomal RNA gene clusters. *Genome Res.* **2007**, *18*, 13–18. [CrossRef] [PubMed]

77. Stults, D.M.; Killen, M.W.; Williamson, E.P.; Hourigan, J.S.; Vargas, H.D.; Arnold, S.M.; Moscow, J.A.; Pierce, A.J. Human rRNA Gene Clusters Are Recombinational Hotspots in Cancer. *Cancer Res.* **2009**, *69*, 9096–9104. [CrossRef]

78. Paredes, S.; Branco, A.T.; Hartl, D.L.; Maggert, K.A.; Lemos, B. Ribosomal DNA deletions modulate genome-wide gene expression: "rDNA-sensitive" genes and natural variation. *PLoS Genet.* **2011**, *7*, e1001376. [CrossRef]

79. Paredes, S.; Maggert, K.A. Ribosomal DNA contributes to global chromatin regulation. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 17829–17834. [CrossRef]

80. Prokopowich, C.D.; Gregory, T.R.; Crease, T.J. The correlation between rDNA copy number and genome size in eukaryotes. *Genome* **2003**, *46*, 48–50. [CrossRef] [PubMed]

81. Banditt, M.; Koller, T.; Sogo, J.M. Transcriptional activity and chromatin structure of enhancer-deleted rRNA genes in Saccharomyces cerevisiae. *Mol. Cell. Biol.* **1999**, *19*, 4953–4960. [CrossRef] [PubMed]

82. Pederson, T. The Nucleolus. *Cold Spring Harb. Perspect. Biol.* **2011**, *3*, a000638. [CrossRef]

83. Yu, S.; Lemos, B. A Portrait of Ribosomal DNA Contacts with Hi-C Reveals 5S and 45S rDNA Anchoring Points in the Folded Human Genome. *Genome Biol. Evol.* **2016**, *8*, 3545–3558. [CrossRef] [PubMed]

84. Yu, S.; Lemos, B. The long-range interaction map of ribosomal DNA arrays. *PLoS Genet.* **2018**, *14*, e1007258. [CrossRef] [PubMed]

85. Bierhoff, H.; Postepska-Igielska, A.; Grummt, I. Noisy silence: Non-coding RNA and heterochromatin formation at repetitive elements. *Epigenetics* **2014**, *9*, 53–61. [CrossRef] [PubMed]

86. McStay, B.; Grummt, I. The Epigenetics of rRNA Genes: From Molecular to Chromosome Biology. *Annu. Rev. Cell Dev. Biol.* **2008**, *24*, 131–157. [CrossRef]

87. Schöfer, C.; Weipoltshammer, K. Nucleolus and chromatin. *Histochem. Cell Biol.* **2018**, *150*, 209–225. [CrossRef] [PubMed]

88. Eickbush, T.H.; Eickbush, D.G. Finely Orchestrated Movements: Evolution of the Ribosomal RNA Genes. *Genetics* **2007**, *175*, 477–485. [CrossRef] [PubMed]

89. Robicheau, B.M.; Susko, E.; Harrigan, A.M.; Snyder, M. Ribosomal RNA Genes Contribute to the Formation of Pseudogenes and Junk DNA in the Human Genome. *Genome Biol. Evol.* **2017**, *9*, 380–397. [CrossRef]

90. Bernardi, G. *Structural and Evolutionary Genomics: Natural Selection in Genome Evolution*; Elsevier: Amsterdam, The Netherlands, 2005; ISBN 978-0-08-046187-8.

91. Galtier, N.; Piganeau, G.; Mouchiroud, D.; Duret, L. GC-content evolution in mammalian genomes: The biased gene conversion hypothesis. *Genetics* **2001**, *159*, 907–911.

92. Varriale, A.; Torelli, G.; Bernardi, G. Compositional properties and thermal adaptation of 18S rRNA in vertebrates. *RNA* **2008**, *14*, 1492–1500. [CrossRef]

93. Wang, H.-C.; Xia, X.; Hickey, D. Thermal adaptation of the small subunit ribosomal RNA gene: A comparative study. *J. Mol. Evol.* **2006**, *63*, 120–126. [CrossRef]

94. Parker, M.S.; Balasubramaniam, A.; Sallee, F.R.; Parker, S.L. The Expansion Segments of 28S Ribosomal RNA Extensively Match Human Messenger RNAs. *Front. Genet.* **2018**, *9*. [CrossRef]

95. Parker, M.S.; Sallee, F.R.; Park, E.A.; Parker, S.L. Homoiterons and expansion in ribosomal RNAs. *FEBS Open Bio* **2015**, *5*, 864–876. [CrossRef]

96. Demeshkina, N.; Repkova, M.; Ven'Yaminova, A.; Graifer, D.; Karpova, G. Nucleotides of 18S rRNA surrounding mRNA codons at the human ribosomal A, P, and E sites: A crosslinking study with mRNA analogs carrying an aryl azide group at either the uracil or the guanine residue. *RNA* **2000**, *6*, 1727–1736. [CrossRef]

97. Barendt, P.A.; Shah, N.A.; Barendt, G.A.; Kothari, P.A.; Sarkar, C.A. Evidence for Context-Dependent Complementarity of Non-Shine-Dalgarno Ribosome Binding Sites to *Escherichia coli* rRNA. *ACS Chem. Biol.* **2013**, *8*, 958–966. [CrossRef]

98. Forsdyke, D.R. *Evolutionary Bioinformatics*, 3rd ed.; Springer: Cham, Switzerland, 2016; ISBN 978-3-319-28755-3.

99. Coleman, A.W. ITS2 is a double-edged tool for eukaryote evolutionary comparisons. *Trends Genet.* **2003**, *19*, 370–375. [CrossRef]

100. Coleman, A.W. Pan-eukaryote ITS2 homologies revealed by RNA secondary structure. *Nucleic Acids Res.* **2007**, *35*, 3322–3329. [CrossRef]

101. Coleman, A.W. Nuclear rRNA transcript processing versus internal transcribed spacer secondary structure. *Trends Genet.* **2015**, *31*, 157–163. [CrossRef]

102. Ruhl, M.W.; Wolf, M.; Jenkins, T.M. Compensatory base changes illuminate morphologically difficult taxonomy. *Mol. Phylogenet. Evol.* **2010**, *54*, 664–669. [CrossRef] [PubMed]

103. Muller, T.; Philippi, N.; Dandekar, T.; Schultz, J.; Wolf, M. Distinguishing species. *RNA* **2007**, *13*, 1469–1472. [CrossRef] [PubMed]

104. Song, J.; Shi, L.; Li, D.; Sun, Y.; Niu, Y.; Chen, Z.; Luo, H.; Pang, X.; Sun, Z.; Liu, C.; et al. Extensive Pyrosequencing Reveals Frequent Intra-Genomic Variations of Internal Transcribed Spacer Regions of Nuclear Ribosomal DNA. *PLoS ONE* **2012**, *7*, e43971. [CrossRef] [PubMed]

105. Schultz, J.; Wolf, M. ITS2 sequence–structure analysis in phylogenetics: A how-to manual for molecular systematics. *Mol. Phylogenet. Evol.* **2009**, *52*, 520–523. [CrossRef] [PubMed]

106. Ruhl, W.M. Compensatory Base Changes Illuminate Morphologically Difficult Taxonomy. Master Thesis, University of Georgia in Athens, Athens, Greece, 2009.

107. Nakhoul, H.; Ke, J.; Zhou, X.; Liao, W.; Zeng, S.X.; Lu, H. Ribosomopathies: mechanisms of disease. *Clin. Med. Insights Blood Disord.* **2014**, *7*, 7–16. [CrossRef] [PubMed]

108. Hetman, M. Role of the nucleolus in human diseases. *Biochim. Biophys. Acta BBA Mol. Basis Dis.* **2014**, *1842*, 757. [CrossRef] [PubMed]

109. Quin, J.E.; Devlin, J.R.; Cameron, D.; Hannan, K.M.; Pearson, R.B.; Hannan, R.D. Targeting the nucleolus for cancer intervention. *Biochim. Biophys. Acta BBA Mol. Basis Dis.* **2014**, *1842*, 802–816. [CrossRef] [PubMed]

110. Woods, S.J.; Hannan, K.M.; Pearson, R.B.; Hannan, R.D. The nucleolus as a fundamental regulator of the p53 response and a new target for cancer therapy. *Biochim. Biophys. Acta BBA Gene Regul. Mech.* **2015**, *1849*, 821–829. [CrossRef] [PubMed]