Article

Phylostratic Shift of Whole-Genome Duplications in Normal Mammalian Tissues towards Unicellularity Is Driven by Developmental Bivalent Genes and Reveals a Link to Cancer

Olga V. Anatskaya [1,*,†] [iD]**, Alexander E. Vinogradov** [1,*,†] [iD]**, Ninel M. Vainshelbaum** [2,3]**, Alessandro Giuliani** [4] [iD] **and Jekaterina Erenpreisa** [2,*,†] [iD]

[1] Department of Bioinformatics and Functional Genomics, Institute of Cytology, Russian Academy of sciences, 194064 St. Petersburg, Russia
[2] Department of Oncology, Latvian Biomedical Research and Study Centre, Cancer Research Division, LV-1067 Riga, Latvia; ninela.vainselbauma@biomed.lu.lv
[3] Faculty of Biology, University of Latvia, LV-1586 Riga, Latvia
[4] Istituto Superiore di Sanità, 00161 Rome, Italy; alessandro.giuliani@iss.it
* Correspondence: olga.anatskaya@gmail.com (O.V.A.); aevin@incras.ru (A.E.V.); katrina@biomed.lu.lv (J.E.)
† These authors equally contributed to this work.

check for updates

Abstract: Tumours were recently revealed to undergo a phylostratic and phenotypic shift to unicellularity. As well, aggressive tumours are characterized by an increased proportion of polyploid cells. In order to investigate a possible shared causation of these two features, we performed a comparative phylostratigraphic analysis of ploidy-related genes, obtained from transcriptomic data for polyploid and diploid human and mouse tissues using pairwise cross-species transcriptome comparison and principal component analysis. Our results indicate that polyploidy shifts the evolutionary age balance of the expressed genes from the late metazoan phylostrata towards the upregulation of unicellular and early metazoan phylostrata. The up-regulation of unicellular metabolic and drug-resistance pathways and the downregulation of pathways related to circadian clock were identified. This evolutionary shift was associated with the enrichment of ploidy with bivalent genes ($p < 10^{-16}$). The protein interactome of activated bivalent genes revealed the increase of the connectivity of unicellulars and (early) multicellulars, while circadian regulators were depressed. The mutual polyploidy-*c-MYC*-bivalent genes-associated protein network was organized by gene-hubs engaged in both embryonic development and metastatic cancer including driver (proto)-oncogenes of viral origin. Our data suggest that, in cancer, the atavistic shift goes hand-in-hand with polyploidy and is driven by epigenetic mechanisms impinging on development-related bivalent genes.

Keywords: polyploidy; unicellularity; early multicellularity; embryonality; cancer; bivalent genes; viral-origin oncogenes

1. Introduction

Whole-genome duplications (WGD) and recurrent polyploidization, providing a source for gene divergence and adaptability to environmental changes, are central in the evolution of biodiversity [1–3]. Polyploid cells are also present in the tissues of vertebrates: in humans and other mammalians, cells with multiplied genomes appear during normal organogenesis of definitive and provisional organs (placenta, heart, brain, liver, skin, blood) [4–10]. Polyploidy also arises in response to stress, wounding, and in cancer [11–18].

The relationship between polyploidy and stemness, both found as typical features of aggressive tumours [19–22], may be associated with the re-activation of evolutionarily ancient programs. The observations show that malignant cells often acquire the phenotypes and reproductive behaviour of unicellular organisms through transient polyploidy "life-cycle", which is reciprocally linked with a cell cycle [23–27]. This behaviour is characteristic for cancer cells resistant to therapeutic treatments [28].

The atavistic theory of oncogenesis suggests that cancer is a reversal from a multicellular to a unicellular state [7,29–33]. The two giant clusters were revealed in the human coexpression gene network (also based on bulk tissue analysis): a widely expressed cluster enriched in genes of unicellular origin enriching the cancer tissues and the other giant cluster of multicellular genes [34,35]. Phylostratigraphic tracking of the genes involved in cancer [36] suggests their link to the emergence of multicellularity in metazoan, many cancer genes have a viral origin and even the typical somatic mutations of the cancer genes can be traced as ancestral [37].

In turn, the studies of Trigos and colleagues [38,39] also showed that the genes of unicellular origin are overexpressed in human cancers as compared to their normal counterpart tissues, whereas the genes appearing at multicellular stages of evolution were downregulated. Moreover, in tumours, the interaction between the unicllularity and multicellularity gene networks is weakened [38,40].

The upregulation of the unicellular giant cluster in cancer cells was shown in the single-cells transcriptomes of various cancer types and in invasive as compared with non-invasive cancer [35,41]. These two clusters can be defined as "echoing" a proper dynamical attractor given, strictly speaking, they are not mutually exclusive (both in cancer and normal cell genes from both clusters are activated). The prevalence of ancient unicellular state of cancer cells with respect to normal tissue corresponds to a partial "going back" to the more ancient unicellular pattern preceding the rise of multicellularity. This implies we do not expect that cancer cells become 'identical to unicellular organism' but that lose some phenotypic characters essential for the existence of an organized tissue like contact inhibition, the crucial role played by shape changes in cancer development is consistent with this view [42,43].

The epigenetic mechanisms favouring transition to unicellularity and cancerogenesis and, thus, the proposed evolution reversal of human cancer cells remain unclear. Since oncogenesis is frequently associated with polyploidy [2,26,44–47], we suggested the involvement of polyploidy in this process and addressed here the phylostratigraphy and protein networks of the genes differentially expressed in polyploid versus diploid tissues.

We choose this approach, because mammalian homologous tissues differ by cell ploidy levels [1,2,5,8,12]. Some species have predominantly polyploid heart and diploid liver (pig and primates), whereas others possess mainly polyploid liver and predominantly diploid heart (rodents) [1,5,8]. Here, we rely on this fact to generate a balanced factorial design with a criss-cross distribution of ploidy/tissue across human and mouse that allows to separate ploidy effect from both tissue and species influences (see Methods).

To promote understanding of the epigenetic nature of ploidy-associated gene regulation, we also investigated how polyploidy influences the expression of bivalent genes. Bivalent genes are characterized by epigenetic ambiguity, bearing in their promoters or enhancers two opposite epigenetic modifications of the histone H3, the repressing H3K27me3 and the activating H3K4me3, poised for transcription but capable for its rapid activation. These poised genes were established in embryonic stem cells [48,49] and are the main players in the early and post-implantation development and cell fate change. Changes of bivalent chromatin coincide with the increased expression of developmental genes in cancer and are also likely involved there in the epigenetic changes [50–55]. On a more general ground the notion of cancer as 'development gone awry' supports the focus on such genes [56].

The obtained results indicate that polyploidy activates unicellular metabolic pathways and ancient programs of development related to carcinogenesis. This evolutionary retour is accompanied by the activation of bivalent genes and deregulation of circadian rhythms. Altogether, these results provide evidence that polyploidy is an important driver of epigenetic changes linked to cancer that may reorganize gene regulatory networks.

2. Results

2.1. Polyploidy Causes Transition of Gene Phylostratic Balance towards Evolutionary Old Unicellular Phylostrata

To find out whether polyploidy affects the expression of genes of various evolutionary ages, we first applied pairwise cross-species transcriptome comparison to polyploid vs. diploid organs in human and mouse (human heart vs. mouse heart and mouse liver vs. human liver). Then we evaluated the effects of polyploidy at various thresholds of the expression difference. This evaluation gave 584 upregulated and 711 downregulated genes at two-fold difference and 1028 upregulated and 887 downregulated genes at a 1.3-fold difference (Tables S1 and S2).

Phylostratic distribution of ploidy-associated genes was analysed by grouping genes according to their age of evolutionary origin taken from [38]. Our data indicated that polyploidy is associated with the upregulation of ancient genes originating in unicellular ancestors (1-3 phylostrata) and the downregulation of genes starting with phylostratum 6 (Bilateria) and onwards, while the early multicellularity strata 4–5 (Metazoa and Eumetazoa) did not reveal a clear difference (Figure 1A,B).
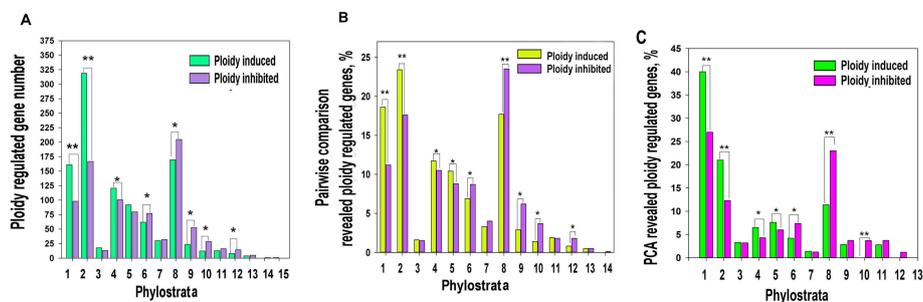

Figure 1. Ploidy associated gene distribution across evolutionary phylostrata. Phylostratigraphy of ploidy associated genes revealed by pairwise cross-species (criss-cross) comparison (**A**,**B**) and by PCA (**C**). This figure illustrates that polyploidy shifts the gene age balance of expressed genes from metazoan phylostrata (6–16, Bilateria and later) towards unicellular phylostrata (1–3) via transitional early metazoan phylostrata (4–5) * $p < 0.05$ for the difference, ** $p < 0.01$ for the difference. Gene expression difference above two-fold. The phylostrata are as follows: 1—cellular organisms, 2—Eukaryota, 3—Opisthokonta, 4—Metazoa, 5—Eumetazoa, 6—Bilateria, 7—Chordata, 8—Euteleostomi, 9—Amniota, 10—Mammalia, 11—Ttheria, 12—Eutheria, 13—Euarchontoglires, 14—Catarrhini, 15—Homininae.

To verify this result with a well-known traditional approach, we applied principal component analysis (PCA). PCA identified 103 upregulated and 96 downregulated genes having above two standard deviations. Of these genes, 78 and 87 genes demonstrated a more than two-fold expression difference (Tables S1 and S2). With excluding the tissue-specific component, the phylostratic distribution of PCA-revealed genes also indicated the activation by polyploidy of evolutionarily conserved programs (genes from phylostrata 1–3) and the suppression of young programs maintaining multicellularity (phylostrata 6–13, Figure 1C). The phylostrata 4–5 were up-regulated.

2.2. Polyploidy Activates Recapitulation of Evolutionary Developmental Programs Associated with Carcinogenesis

To investigate the functional consequences of ploidy-associated phylostrata rearrangement at the gene module level, we performed gene module enrichment analysis for genes with more than two-fold expression difference between polyploid vs. diploid organs. The results are presented in Table S3. For unicellular phylostrata (1–3) the main effects of polyploidy were the induction of modules related to drug ABC pump and drug metabolism (hsa02010;GO:0017144), protein and ribosome synthesis (hsa01230;hsa03010), oxidation-reduction (GO:0055114) and energy metabolism

(both aerobic respiration and carbohydrate metabolism (GO:0019752; GO:0008152; GO:0044262; hsa00051). This picture is seen from the higher significance of enrichment and the larger gene number for upregulated genes than for downregulated ones. For the up-regulated by polyploidy early metazoan phylostrata (4–5) our data reveal the upregulation of modules related to embryonic development (GO:0009790), stem cell commitment (GO:0045165), pluripotency (hsa04550) and Kegg pathways of carcinogenesis (hsa05200; hsa05205). The phylostrata 6–8 were down-regulated by polyploidy. They are involved in the development of the multicellular organism complexity. The downregulated genes of the 6–8 strata were mostly enriched in biological processes related to immunity (GO:0002682), inflammation (GO:0050727) and communication (GO:0010646) that are the specific biological features of multi-cellularity. Consistently, multicellular phylostrata (10–15, i.e., Mammalia and later) were not enriched for gene modules depending on polyploidy.

The results of functional enrichment of PCA-ploidy-revealed genes for old phylostrata were consistent with the results of pair-wise cross-species comparison (Table S4). Thus, the carcinogenesis pathways favoured by polyploidy appear in the early metazoan (strata 4–5) in one pack with embryogenesis and cell fate change (stemness commitment), associated with asexual reproduction.

Similar changes in the general phylostratigraphic landscape shifted towards unicellularity found here for polyploidy were previously shown for tumorigenesis as such [38]. The authors also paid attention to the enrichment of the intermediate between unicellularity and multicellularity phylostrata with cancer driver genes [40]. In comparison with our data, this similarity suggests that polyploidy may favour the epigenetic evolutionary shift of normal cells to cancer. The interactome of *c-MYC* proto-oncogene provides a useful link to verify this suggestion because *c-MYC* in an important regulator of both ploidy and cancer.

2.3. c-MYC Induction Drives Polyploidy-Associated Transcriptomic Changes towards Unicellularity

c-MYC is a regulator of the normal cell cycle. However, when over-expressed it disjoins replication from cell division and favours polyploidy [57,58]. Moreover, while the overexpressed non-mutant (often amplified) *c-MYC* is a powerful reprogramming factor and oncogene, its downregulation in transgenic mice causes tumour regression [59–62].

The pleiotropic *c-MYC* is capable to drive the opposite processes, proliferation versus apoptosis and is acting as a master activator of the bivalent genes involved in the epigenetic regulation of development [60,63–65].

Here we evaluated the phylostratic distribution of ploidy-dependent *c-MYC* -interacting genes at two-fold expression difference (Tables S5 and S6) and explored its link through polyploidy with bivalent genes (Figure 2). The data on ploidy -Myc-ploidy-related gene phylostratigraphic changes repeat the ploidy-dependent relationships (compare with Figure 1)
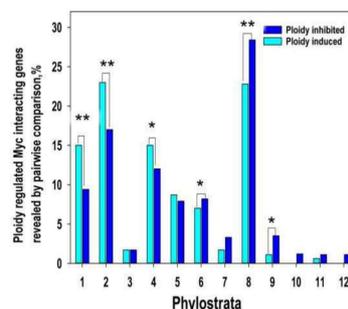

Figure 2. Ploidy associated *c-MYC* interacting gene distribution across evolutionary phylostrata. This figure illustrates that *c-MYC* -interacting genes repeat pattern of all ploidy related gene age distribution. Gene age is shifted from young multicellular phylostrata (6–16) towards unicellular phylostrata (1–3 phylostrata) via transitional metazoan phylostrata (4–5). * $p < 0.05$ for the difference, ** $p < 0.01$. Gene expression difference is above two-fold.

It is clear that the non-mutant *c-MYC* should elaborate these effects in the epigenetic modus. Therefore, we paid attention to the fact that *c-MYC* controls gene transcription by activating bivalent genes.

2.4. The Phylostratigraphic Polyploidy-Associated Effect of c-MYC Is Associated with the Regulation of Bivalent Genes

The genome-wide studies indicate that the overexpressed *c-MYC* demethylates the H3K27me3 repressive domain of bivalent genes and thereby switches these genes to the active state [63,64], it also activates Polymerase II paused in bivalent genes [53]. To find out whether overexpressed Myc also increases expression of ploidy-related bivalent genes, we first identified these genes among ploidy-regulated *c-MYC* interacting genes with expression difference above two-fold. Genes of interest were identified using the list of bivalent genes in human embryonic stem cells [54]. Our data identified 60 bivalent genes of 161 ploidy-upregulated Myc interactants and 27 bivalent genes among 101 ploidy-downregulated *c-MYC* interactants (Figure 3A,B, Tables S7 and S8).
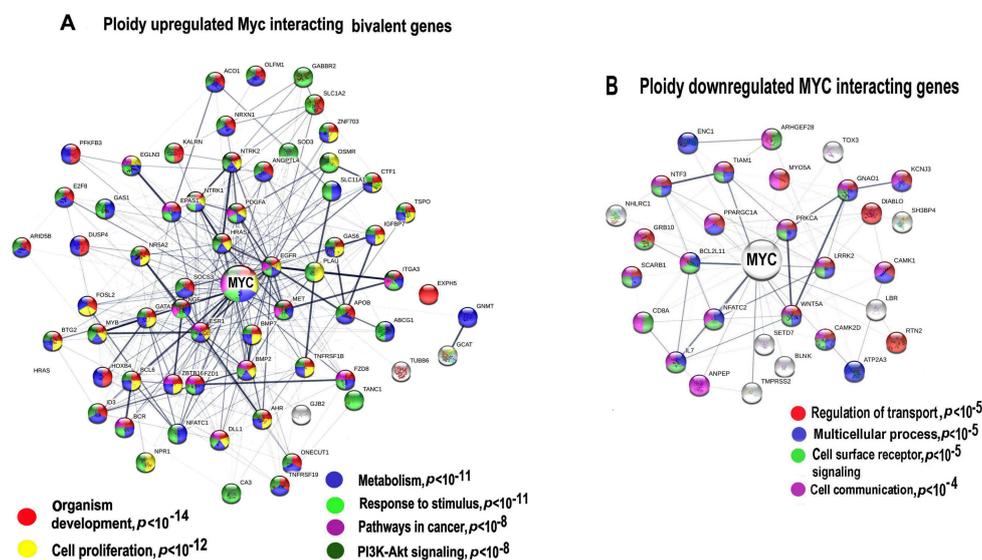

Figure 3. Protein interaction networks (PPIs) for ploidy associated *c-MYC* -interacting bivalent genes. PPI for upregulated (**A**) and downregulated (**B**) genes. Gene expression difference is above two-fold. The networks are constructed with server String. Gene pathway enrichment was found using the same server.

Protein interaction network presented in Figure 3A indicates the Myc interacting and ploidy upregulated bivalent genes. Gene module functional enrichment analysis indicated that these genes are enriched in pathways of development, proliferation, stress response, and pathways in cancer with high significance. Figure 3B that presents *c-MYC* interacting ploidy downregulated genes shows that the cell surface-mediated functions (transport, reception, communication) are downregulated.

It is important to note that the network for the upregulated bivalent genes shows higher connectivity (13.5 connection per a node) compared to the downregulated genes (only 5.13 connections per a node), which implies that *c-MYC* and polyploidy- induced activity of bivalent genes considerably increases epigenetic plasticity of protein interaction network. It should be noted that the overexpressed functional pack again included a tandem of development and carcinogenesis modules, which we have found in the ploidy-related gene phylostratigraphic analysis. Therefore, we decided to extend the investigation of bivalent genes in relation to ploidy and gene phylostrata through the entire transcriptome.

2.5. Ploidy-Associated Genes Are Enriched in Bivalent Genes of the Entire Genome, Prevailing among Unicellular and Eumetazoan Phylostrata.

We compared the genes with above two-fold expression difference dependent on polyploidy (vs. diploidy) with bivalent genes in human embryonic stem cells [54]. Among the 584 up and 711 downregulated by ploidy genes (Tables S1 and S2) we found 267 (45.7%) and 222 (31.6%) of up-and down-regulated bivalent genes, respectively (Tables S9 and S10). The results of the binomial test indicate that bivalent genes enrich ploidy-upregulated genes compared to the entire transcriptome (3024 vs. 14093) with high significance ($p < 10^{-16}$ binomial test). For the downregulated genes the significance was $p < 10^{-8}$ (for 21.3 vs. 31.2%), respectively. These results indicate that polyploidy mostly induced bivalent genes.

The percentage of phylostratigraphic distribution of ploidy associated bivalent and common genes is shown in Figure 4. The bivalent genes, comprising ~21% of the human genome, originated in evolution in all phylostrata but in the larger proportion with development of multicellularity, from phylostratum 4 (Metazoa) to 8 (Euteleostomi). The distribution of the expression of the ploidy upregulated bivalent genes by phylostrata is different from that for all bivalent genes. It is clearly seen that these genes repeat the general effect of all ploidy-regulated genes and *c-MYC*-ploidy regulated genes. Among the ploidy upregulated bivalent genes prevail the proportions of phylostrata 1 (Prokaryota) to 5 (Eumetazoa). Proportions of phylostrata 6 and 7 (Bilateria and Chordata) are ambivalent. Proportions of phylostrata 8 (Euteleostomi) to 10 (Mammalia) show clear decrease (Figure 4). The ploidy downregulated bivalent genes generally repeat the pattern for all bivalent genes.
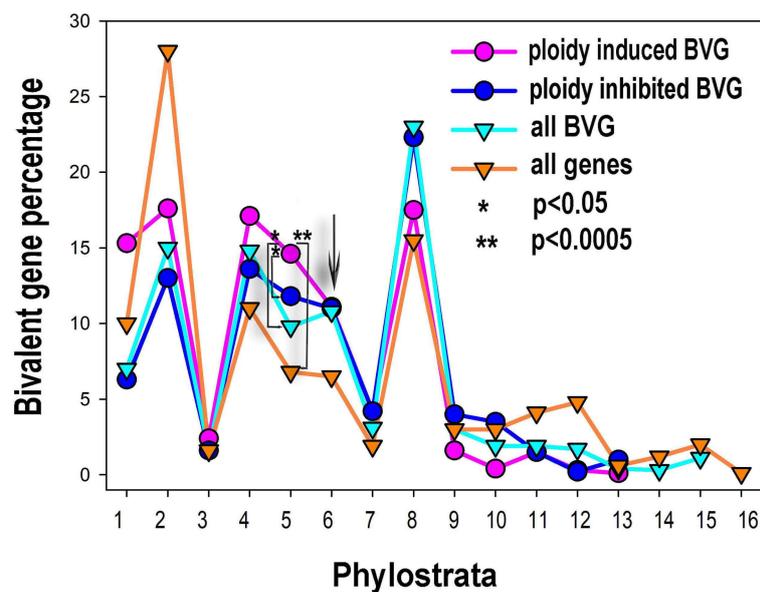

Figure 4. The percentual proportions of gene origins and distribution of bivalent genes (BVG) in the phylostratic tree of life (strata 1–16) and the effect of polyploidy on it. The upregulation of bi-valent genes by polyploidy includes strata 1, 2 (unicellularians), stratum 4 (metazoa) and, prominently, stratum 5 (eumetazoa—the appearance of embryo, germ layer, and gastrulation). The arrowed cross-point starting down-regulation of bivalent genes by polyploidy in late metazoa falls upon stratum 6 (bilateria).

2.6. Protein Interaction Networks for Ploidy-Associated Bivalent Genes Are Involved in the Upregulation of the Developmental and Carcinogenesis Genes and the Downregulation of the Networks Related to Differentiation Biological Quality and Circadian Clock

We constructed protein interaction networks (PPIs) for ploidy up- and downregulated bivalent genes marked with H3K27me3 and H3K4me3 in human ESC using the String server. We also included *c-MYC* that is not a bivalent gene, but it is of particular interest because this important oncogene is an inducer of bivalent genes [64]. The networks were constructed for genes with above two-fold expression

difference. Then we extracted the whole connected component from the up- and downregulated networks and clustered them using the same server. For the upregulated genes the network contains 165 bivalent of 267 ploidy upregulated bivalent genes (62%). For the ploidy downregulated bivalent genes it contains 63 genes of 222 (24%). As in the case of *c-MYC*, the striking difference in gene numbers of the connected components of up- and downregulated networks (163 vs. 63) by bivalent genes in the whole transcriptome, despite approximately the same numbers of up- and downregulated genes, shows that the upregulated by ploidy bivalent genes create essentially more functional connections and master regulators compared to the downregulated genes. Below we provide a functional description of the up and downregulated PPIs (Figure 5A,B). It is clearly seen that the upregulated network contains more hub regulators than the downregulated network (29 vs. 13).
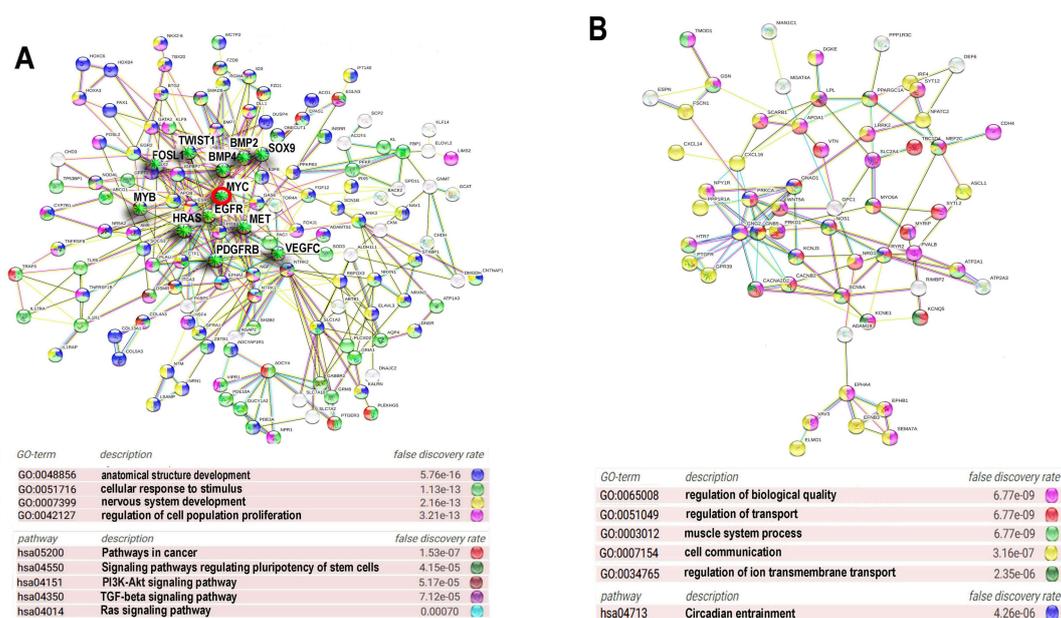

Figure 5. Protein interaction networks for ploidy associated bivalent genes. The most connected components of protein interaction networks of upregulated (**A**) and downregulated (**B**) bivalent genes constructed using STRING server. Gene expression difference is above two-fold. Gene module analysis and node degree evaluation was done using the same server. The titles of GO biological processes are given with for false discovery rate. The important drivers of carcinogenesis are marked with green asterisks. Myc oncogene is marked with a red circle.

The PPI for the ploidy upregulated bivalent genes (Figure 5A) is enriched in the GO biological processes related to the development of anatomical structures (GO:0048856) and the nervous system (GO:007399), proliferation (GO:0042127), response to stress and stimulus (GO:0051716), and in KEGG pathways involved in the regulation of pluripotency (hsa04550), pathway of carcinogenesis (hsa05200) and several pro-carcinogenic and development related pathways, including *PI3K-AKT* (hsa04151) *TGFB* (hsa04350) and *RAS* (hsa04014) signalling pathways (Figure 5A). It is important to note that this PIN unifies proteins involved both in development and metastatic cancer that comprise dense functional nucleus of the induced network and are the hub proteins with high degree of connectivity originating from 1–6 phylostrata genes (Table 1). The PPI for the downregulated genes is enriched for GO biological processes participating in transport (GO:0051049, GO:0034765), cell communication (GO:0007154), regulation of biological quality (GO:0065008), muscle system processes (GO:003012) and circadian entrainment (hsa04713). Thus, the functional picture provided by PPI for the upregulated sgenes indicates that ploidy-related bivalent genes are associated with the induction of embryonic developmental programs including drivers of carcinogenesis (Figure 5A, Table 1). PPI for the downregulated genes illustrates the depression of circadian clocks.

Table 1. Characteristics of the central hubs of Myc-related network of proteins encoded by bivalent genes upregulated by polyploidy (presented in Figure 5A).

Gene/ Protein	Folds Up-Regulation	Number of Connections	Phylo Stratum	Function Involvement in Cancer	References
BMP4	3.24	26	4	Development, cell motility, oncogene	[66]
BMP2	3.84	14	5	Development, cell motility, oncogene	[67]
c-MYC	3.2	29	3	Stemness, proliferation, apoptosis, polyploidy induction, driver oncogene	[68]
MET	3.7	15	6	Embryonic development, proliferation; when mutationally activated is involved in multiple cancers	[69]
MYB	3.8	8	2	Development of colon epithelial progenitors, high in gastric, colon, and breast cancer	[70]
HRAS	3.6	19	1	Proliferation, senescence, germ development, driver oncogene	[71,72]
EGFR	4.94	34	6	Epithelium development, driver oncogene	[73]
PDGFRB	2.23	15	6	Development, cancer-related angiogenesis	[74]
VEGFC (C)	2.88	12	5	Blood vessels development, cancer-related angiogenesis	[74]
SOX9	3.86	15	4	Germ (male) and skeleton development	[75]
TWIST1	2.53	11	2	Mesoderm development, motility, metastatic cancer	[76]
FOSL1	2.19	11	5	Early stress response, reinforces Myc, oncogenesis	[77]

2.7. The Relation of Cancer Driver Genes to Polyploidy, Bivalency and Their Phylostratigraphic Origin

To obtain more evidence concerning features of carcinogenesis, we evaluated the expression of tumour suppressors and oncogenes in relation to polyploidy based on the list from [78]. The obtained data presented in Figure 6, together with phylostratic origin of genes, indicates that polyploidy is associated with preferential downregulation of tumour suppressors and strong upregulation of many oncogenes. It is important to note that the upregulated oncogenes were more strongly enriched with bivalent genes than the downregulated oncogenes and tumour suppressors ($p < 0.02$ for all comparisons, binomial test). The notably down-regulated by polyploidy tumour suppressors belong to DNA damage response: *ATR*, (2nd phylostratum), *ATR* (2nd), *CHEK2* (2nd), *CHEK1* (3rd), *TP53* (5th)—however, its protein inhibitor *MDM2* (seen in the column for oncogenes) is also inhibited; apoptosis executor Fas, prominent tumour suppressor *PTEN* (2nd) and the E-cadherin-associated gene *CDH1* (8th phylostratum). Thus, cell communication is suppressed by polyploidy, while cells become more tolerant to DNA damage and inhibit apoptosis.
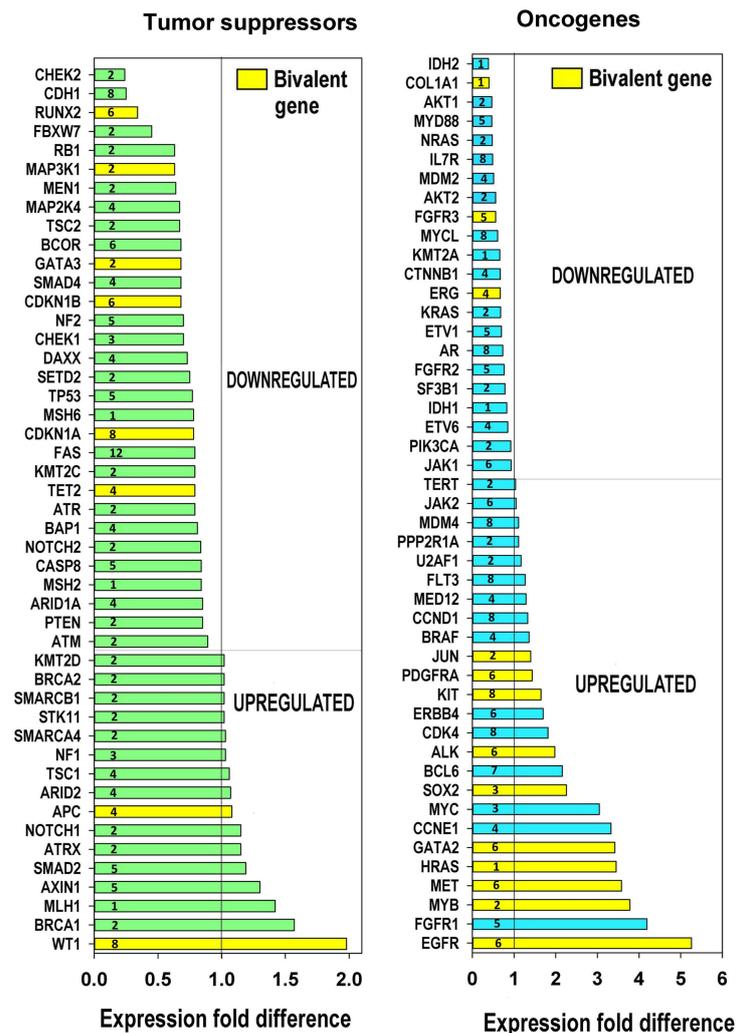
Figure 6. Ploidy associated tumor suppressors and oncogenes. This figure illustrates the downregulation of tumour suppressors, the upregulation of oncogenes and the enrichment of the induced oncogenes with bivalent genes $p < 0.05$, binomial test. The phylostrata numbers of gene evolutionary origins are indicated.

Among the prominent oncogenes upregulated by polyploidy, we find a master transcription factor (TF) of early stress response Jun (2nd phylostratum), *KIT*-tyrosine-protein-kinase (6th); among 3–5-fold upregulated oncogenes, we see *GATA* 2 TF (6th) responsible for embryonic development, *c-MYC* TF (3rd, viral origin); *MYB* TF (2nd, viral origin); *MET* Tyrosin-protein-kinase (6th)— often a target for avian leukemia retrovirus [79] *HA-RAS* G-protein (1st phylostratum, viral origin)—the most common oncogene in at least 30% of all tumours and, finally, its upstream regulator *EGFR* (6th); both are bivalent (Table S9).

Clearly, in this analysis, we have used the transcriptome data from polyploid tissues, where the activated protooncogenes were not mutated. In addition to their relation to bivalency and embryonic development, we see that the main tumour suppressor *TP53* and strong oncogene *EGFR* belong to the intermediate phylostrata 5–6, while many important driver oncogenes of this list activated by polyploidy are from the earliest evolutionary phylostrata. Furthermore, the most often involved in carcinogenesis prominent complementary pair of oncogenes *c-MYC* and *HRAS* (phylostrata 1–3) are of the most ancient origin.

3. Discussion

Despite extensive development of therapies, metastatic cancer disease remains incurable, resulting in high morbidity and mortality [80]. The association of aggressive incurable cancer with polyploid giant cells has been shown [28]. The extraordinary resistance to extinction therapies suggests that cancers recapitulate the phylogenetic endurance acquired in the long evolution of the lifeforms on Earth [81]. The cancer genome sequencing projects compromised, to some extent, a somatic mutation theory [82,83]. In recent times, the epigenetic aspects of the whole genome regulation in cancer stepped forward and are now intensively explored. One of these aspects is associated with typical for cancers whole genome duplications, or polyploidy (and aneuploidy inevitably linked to it), as a feature characteristically accompanying tumour growth and aggravating with tumour aggression [18]. The genetic sequences of polyploidy and aneuploidy in the microevolution of cancer were in the focus of studies for rather a long time [84,85]. However, the epigenetic aspects of aneuploidy became also attractive [86,87]. Considering a possible atavistic effect of cancer polyploidy, the phylostratigraphic analysis considering the evolutionary origin of genes and the change of their expression balance may be helpful.

Therefore, an interesting epigenetic approach may represent the study of the phylostratigraphic effect of polyploidy of not mutant normal mammalian tissues. In this paper we have investigated these issues by bioinformatical means using transcriptome data of polyploid versus diploid cells of normal human heart and mouse liver.

We have revealed that polyploidy changes the phylostratigraphic balance of the mammalian cellular network in favour of the enhanced expression of the evolutionary most ancient gene phylostrata—prokaryotic, early eukaryotic and early metazoan (1–5 phylostrata) up to bilateria, after which the trend is changed to the opposite. Although found here in normal polyploid tissues, this shift is surprisingly similar to the changes revealed for tumours when compared with normal counterparts [34–36,38]. With this revelation we arrive to the immediate conclusion: this epigenetic shift in tumours to the unicellular and early multicellular forms of life is associated or at least highly favoured by polyploidization.

The question arises—how it is regulated (or rather disregulated)? One of the mechanisms found in our study—it is by use of so-called bivalent genes, the epigenetically modified by histone H3 repressive and activating modifications of the same gene promoter or enhancer, namely by their prevailing activation. The particularity of bivalent genes comprising ~21% of human genes and employed for cell fate change in development is that they are poised (paused) but can be rapidly activated. That means that they favour cell fate change by the rules of non-equilibrium thermodynamics, in particular by critical state transition [88].

It is known now that the chromatin conformation at the supranucleosome and higher levels of the genome architecture is much a subject of the biophysical processes which assemble the active and inactive chromatin in two separate self-organizing high-order domains [89–93]. Self-organization is an important driving force of evolution [94]. On one hand, polyploidy is well known as driving gene diversification and speciation in evolution [1], which is a slow process. On the other hand, its epigenetic effect causing rapid reorganization of the transcriptional regulatory network after genome duplication was shown [95,96].

It can be reasoned that multiplication in polyploid cells of the identical chromosome alleles with the identical epigenetic status of bivalent genes should not just linearly increase their epigenetic effect proportionally to gene dosage but potentiate it by the thermodynamics of structural self-organization (phase transition). The role of the closely-juxtaposed DNA fibres' electrostatic forces yielding fine-tuned structure-specific recognition and pairing [97,98] should also contribute in the chromatin self-organization by polyploidy. A similar thought on the potentiating structural effect of additional DNA by the chromatin modifiers and transcription factors has been proposed earlier [99,100].

The Kegg "pathways in carcinogenesis", along with the genes known as being mutants in cancers [101], also include many developmental genes. Here we revealed that this developmental

component is largely introduced through the polyploidy-dependent enrichment of the expression networks with the activated bivalent genes upregulated by *c-MYC* master hub (stratum 3). While in general gene evolution the proportion of bivalent genes increases from phylostrata 3 to 8, the polyploidy is mostly affected by enrichment with expression of bivalency in the phylostrata of Prokaryote and Eumetazoa (1—5). Very likely, that it is achieved by Myc up-regulating both polyploidy and bivalency. Thus, the carcinogenic *H-RAS-c-MYC* feed-back loop (strata 1–3) linked to developmental genes of early metazoans (strata 4–5) enriched with developmental pathways becomes involved, while stratum 6 (Bilateria) becomes a cross-point of the ambivalence (with repulsion from vertebrates).

c-MYC–mtH-RAS complementary pair has been distinguished in the early cancer chemical carcinogenesis research in a two-hit theory of cancer with Myc determining immortality and cancer initiation and *H-RAS* mutant in the codon 12 or 61 accomplishing cancer promotion [71,102]. Afterwards, ample studies have extended this theory by demonstrating that a potent oncogenicity of Myc can be further enhanced by collaborations with not only *RAS* mutant but also many extracellular growth stimuli that activate *RAS*, such as epidermal growth factor and its receptor (*EGFR*) or transforming growth factors [103,104]. The found phylostratic effect of ploidy-associated *c-MYC* collaborating with bivalent genes, e.g., both *H-RAS* (stratum 1) and *EGFR* (stratum 6) may explain its critical role in cancer initiation when overexpressed, and cancer regression when locked [59,60,62]. This phylostratic effect induced by the oncogenic potency of *c-MYC* linked to induced polyploidy, metabolic stemness, and bivalency likely creates the evolutionary "cancer attractor" postulated by Kauffman [94] as situated close to the summit of metaphoric Waddington Hill of the development potential [105,106]. Moreover, the upregulated *c-MYC* cooperates with mutant ras through programming inflammation and suppressing the tumour immunity [61], and *c-MYC* acts by cell selection promoting proliferation versus apoptosis depending on the supply of growth factors [60]. Likely, this trigger can occur through asymmetric cell division of epigenetically diverged bi-nuclear cells [107,108]. In fact, the main actors of this complex "story" including the complementary *c-MYC* and *RAS* pair can be found in Table 1 describing tumour and stromal components (like angiogenesis) of the polyploidy-dependent "cancer attractor".

Interestingly, cnidaria (Hydra), an organism where the basic processes of the multicellularity were established represents phylogenetically the oldest organisms, where the tumours were revealed [109].

The early eukaryotic gene cluster in human genome is more stable than the cluster of complex metazoans (starting from vertebrates and peaking in acquired number of new genes at stratum 8) [38,40,41]. The unicellular gene network is responsible for basic cellular functions and also for resistance to extinction elaborated through series of earth catastrophes in billion years—we highlighted here that the genes responsible for DNA damage response, drug resistance, early stress response, and the proto-oncogenes of the viral origin, are from the 1–3 phylostrata. These likely act through feedback loops as the driving belts epigenetically attracting the cells from the basin of early metazoans into a unicellularity network.

On the other hand, the basic evolutionary processes of the early embryonal development of mammals are conserved from their origin in Eumetazoa, stratum 5 [110]. Therefore, aggressive cancer associated with polyploidy and multinuclearity often bears the features of the early embryo [111,112], including such inalienable components of morphogenesis as cell motility and angiogenesis. Moreover, through polyploidy, the *TP53* cancer mutants can acquire the phenotypical features of more ancient unicellular eukaryotes ("amoeboidisation") [24,27,33,108,113,114], in line with the atavistic theory of cancer [29,30,33,115,116], confirmed by gene phylostratigraphic analysis [38] and here.

The early metazoans are tolerant to polyploidy because use it reversibly in their life-cycles [117]; it is sufficient to mention *Amoeba vulgaris* and *Entamoeba histolytica* [33,118,119] and amoebal slime molds [120]. When associated with polyploidy, as we found, the bivalent genes increase the connectivity and basin of attraction and become embedded in the network with the oncogene hubs of the viral origin, such as *c-MYC*, *H-RAS*, MYB—originated and having strong established interactions in the older strata (1–2–3). Moreover, polyploidy permits DNA replication but stops cell division and induces suppression

of the bivalent genes regulating circadian rhythms, which entrain many cellular processes, including cell cycles [121]. The cell-autonomous circadian timers established already in cyanobacteria are composed of a transcription–translation-based auto-regulatory feedback loops [122]. Apparently, this detrainment is an important factor causing along with epigenetic bibivalency a metastable dynamic state, which can also favour a reshape and glide from the multi-cellular to unicellular gene expression network.

However, de-polyploidisation, in turn, can probably cause the back-shift of the transcription network: the two cycles, polyploidy (life-cycle-like) and diploidy (cell cycle) are reciprocally linked in tumours [23,87]; the recovery diploid fraction reciprocally downregulates the expression of amoebal genes in polyploidized human breast cancer treated with anticancer drugs; in addition, the amoebal giant cells can play a "nursing" function for the reproductive cell line [27]. Without any doubt, the constitutively activating mutations of oncogenes increase, push and likely stabilize the epigenetic shift to unicellularity. Such prominent tumour suppressors as *TP53* and *HYPPO/YAP* pathway, which also serve as guardians of the barrier to polyploidy [123], normally prevent the carcinogenic epigenetic shift to unicellularity caused by overcoming this barrier.

4. Methods

4.1. Data Sources and Comparative Criss-Cross Analysis

In this study we applied the method of reciprocal pair-wise cross-species transcriptome comparison that we used previously [12,57,124]. We choose this approach, because mammalian homologous tissues differ by cell ploidy levels [1,2,5,8,12]. Certain species have predominantly polyploid heart and diploid liver (pig and primates), whereas others possess mainly polyploid liver and predominantly diploid heart (rodents) [1,5,8]. Humans have highly polyploid hearts where cardiomyocytes contain nuclei with 4–16 genomes, whereas mouse hearts, on the contrary, consist of cardiomyocytes with mostly diploid nuclei [12]. At the same time, human hepatocytes are mostly diploid, whereas mouse hepatocytes contain nuclei with 4–8 genomes [5,12]. This inverse pattern of polyploidization enables identifying pure ploidy specific effects since it removes species-specific and tissue-specific noise [12]. In this study we applied pairwise cross species transcriptome comparison and principal component analysis (PCA) to comparisons of polyploid vs. diploid tissues: human heart vs. mouse heart and mouse liver vs. human liver. In the present work there are three main advancements over previous studies. Firstly, here we made the functional analysis of all ploidy-associated genes based on the whole-transcriptome data. Previously, the analysis was limited to the *c-MYC* interactome only [58,124]. Secondly, we studied the gene evolutionary origin (phylostratic distribution of all polyploidy-associated genes), which is important for cancer research in the context of the main contradiction arising between the unicellular and multicellular level features in tissue organization [34,38,41]. Thirdly, we addressed possible epigenetic mechanisms underlying ploidy-related changes, focusing on the bivalent gene expression and interactome. The work builds upon the idea of reciprocal cross-species comparison. The cross-species approach is informative because evolutionary distance enhances the contrast allowing the polyploidy-specific signature to appear [125].

The signal-to-noise filtration, where signal corresponds to polyploidy and noise to species- and tissue-specific signatures, is especially important for investigation of polyploidy because polyploidy preserves gene-dosage balance and thus may exert only weak and idiosyncratic effects on gene expression [126]. We performed the reciprocal pair-wise cross-species comparison using the transcriptomic data for polyploid vs. diploid organs. Specifically, we compared human heart (polyploid) vs. mouse heart (diploid) and mouse liver (polyploid) vs. human liver (diploid). The data were taken from the database obtained using RNA-seq [127].

The processed data were taken from the Supplement at the journal site [127]. We matched human-mouse orthologous genes and normalized the data in pairwise comparisons (human-mouse heart, human-mouse liver) using "quantile" normalization implemented in the "limma" package [128], as was done previously [124,129]. Finally, there were 13,913 orthologous gene pairs for which we found

stratigraphic data in Trigos et al. [38]. The transcriptomic data were uniformly obtained for all tissues and species [127]. This database was developed specially for interspecies comparisons (in particular, the samples were taken from the same parts of organs).

We detected the genes whose expression was changed in the same direction with regard to ploidy both in the heart and liver. In these comparisons the genes should have higher (or lower) expression in both polyploid tissues (human heart and mouse liver) compared with corresponding diploid tissues (mouse heart and human liver). Since we compared two different tissues in opposite directions in different species (human vs. mouse in the case of heart, and mouse vs. human in the case of liver), the tissue-specific and species-specific effects were presumably removed.

The one-to-one human-mouse orthologous genes were obtained from the NCBI database [130]. The expression levels of orthologous genes were analysed using the "limma" package specially developed for revealing differentially expressed genes in whole-transcriptome analyses [128]. A comparison of different software packages showed that limma is the method of choice for goals similar to those pursued in our work [131]. The data were normalized with "quantile" normalization implemented in "limma". The differential gene expression (with statistical significance) was determined using the "voom" limma procedure. Then, we selected the genes with different expression contrasts between polyploid and diploid tissues as indicated in Results.

4.2. Principal Component Analysis

To determine whether the results of cross-species comparison can be confirmed by another approach, we applied principal component analysis (PCA) to the raw data matrix having samples as variables and genes as statistical units. The idea is to confirm the gene-by-gene a priori approach with a data-driven strategy, letting a ploidy-specific principal component to emerge from the data. The principal components are orthogonal to each other by construction; the data-driven emergence of a "ploidy" component distinct from tissue and "species" components is equivalent to an unsupervised normalization for tissue and species effects [132]. The genes endowed with extreme scores on such a "ploidy" component are the "image in light" of tissue and species independent ploidy effect on transcription pattern.

For this propose we used the same transcriptomic data for human and mouse heart and liver. This approach enabled us to evaluate the impact of shared variation (driven by housekeeping genes), species-specific, tissue-specific and ploidy-specific variables as separated mutually independent components [133].

As a result, we obtained two lists of genes demonstrating statistically significant ploidy-associated variation (above two standard deviations) for human and mouse heart and liver. Then these gene lists were subject to gene module enrichment analysis and the modules that are regulated by ploidy were identified. PCA is a purely geometrical non probabilistic procedure [132]. This means that each gene is simply projected into a rotated space (in this case having the same dimension of the original one, thus, with no loss of information) spanned by mutually orthogonal axes (components) extracted in decreasing order of explained variance. This unsupervised purely geometrical approach gave rise to a loading pattern mirroring the batch, tissue, species, ploidy a priori classification so demonstrating the tenability of the existence of "ploidy specific" genes. At this point the identification of genes having a score > |2| as markers of the ploidy component corresponds to the usual 95% confidence intervals (component scores are equivalent to z scores). The actual *p*-values for phylostratigraphic analysis were based on relative enrichment of genes pertaining to different stratigraphic level based on the entire gene ontology. The samples having same tissue and species origin are practically coincident [see 132] for the presence of a common gene expression ideal profile for same tissue independent samples). RNA-seq analysis was based on the shared gene products across different samples; these shared genes were the pivots for coupling the relative vectors.

The species-specific and tissue-specific effects were automatically removed by PCA [124]. As a matter of fact the principal component analysis applied in an unsupervised way to the gene expression

profiles, generated a four-component solution correspondent to batch effect (PC1, 45.6% of variance explained), tissue-effect (PC2, 30.5% of variance); species-specific effect (PC3, 15% of variance) and ploidy effect (PC4, 9% of variance). We concentrated on genes having relevant score (higher than |2| corresponding to two standard deviations from mean) on PC4. Principal components are each other independent by construction, therefore batch, tissue and species confounding are eliminated by the spectral decomposition of the dataset.

4.3. Analysis of Gene Modules

To determine which biological modules were over-represented among the ploidy-associated genes, we applied a double control. We tested the genes from all three datasets with higher and lower expression in polyploid vs. diploid tissues for enrichment of Gene Ontology (GO) categories and molecular pathways with regard to all human-mouse orthologous genes. The enriched GO categories and molecular pathways were found using the hypergeometric distribution of probability (implemented in R package) as previously [134,135]. GO categories were taken from GO database [136]. For each GO category, we collected all its subcategories using Gene Ontology acyclic directed graphs, and a gene was regarded as belonging to a given category if it was mapped to any of its subcategories. As a source of molecular pathways, the NCBI BioSystems was used, which is a most complete compendium of molecular pathways from different databases [130]. Redundancy of this compendium was removed by uniting entries with identical gene sets. The adjustment for multiple comparisons was done according to the method by [137]. This procedure gives a q-value, which can be considered a *p*-value corrected for multiple tests.

4.4. Protein-Protein Interaction Network

The protein-protein interaction networks (PPI) were constructed and visualized using the STRING server [138]. The analysis of network connectivity and the identification of causal regulators in modular and network organization were done with the same server.

4.5. Phylostratigraphic Analysis of Ploidy Associated Genes

The attribution of genes to phylostrata was taken from [38]. The phylostrata are as follows: 1—cellular organisms, 2—Eukaryota, 3—Opisthokonta, 4—Metazoa, 5—Eumetazoa, 6—Bilateria, 7—Chordata, 8—Euteleostomi, 9—Amniota, 10—Mammalia, 11—Theria, 12—Eutheria, 13—Euarchontoglires, 14—Catarrhini, 15—Homininae, 16—*Homo sapiens*.

4.6. Identification of Ploidy Associated Bivalent Gene

To find out how polyploidy affects bivalent gene expression landscape, we evaluated the expression of ploidy-associated genes that are known to be marked by bivalent chromatin in stem cells. The list of these genes was taken from [54].

5. Conclusions

The obtained results indicate that polyploidy activates the unicellular pathways of resistance to extinction fusing with ancient programs of ontogenesis related to carcinogenesis which originated in the early metazoa while suppressing the expression of genes from the late metazoa. This evolutionary retour is favoured by the activation of bivalent genes and deregulation of circadian rhythms. The data highlight the paramount role of polyploidy in the atavistic origin of cancer, the reason for the incurability of metastatic cancer, and indicate to targeting polyploidy for overcoming resistance to therapies.

References

1. Comai, L. The advantages and disadvantages of being polyploid. *Nat. Rev. Genet.* **2005**, *6*, 836–846. [CrossRef] [PubMed]
2. Fox, D.T.; Soltis, D.E.; Soltis, P.S.; Ashman, T.-L.; Van de Peer, Y. Polyploidy: A Biological Force from Cells to Ecosystems. *Trends Cell Biol.* **2020**, *30*, 688–694. [CrossRef] [PubMed]
3. Avdeyev, P.; Alexeev, N.; Rong, Y.; Alekseyev, M.A. A unified ILP framework for core ancestral genome reconstruction problems. *Bioinformatics* **2020**, *36*, 2993–3003. [CrossRef] [PubMed]
4. Anatskaya, O.V.; Vinogradov, A.E. Myocyte ploidy in heart chambers of birds with different locomotor activity. *J. Exp. Zool.* **2002**, *293*, 427–441. [CrossRef] [PubMed]
5. Anatskaya, O.V.; Vinogradov, A.E. Heart and liver as developmental bottlenecks of mammal design: Evidence from cell polyploidization. *Biol. J. Linn. Soc.* **2004**, *83*, 175–186. [CrossRef]
6. Sher, N.; Von Stetina, J.R.; Bell, G.W.; Matsuura, S.; Ravid, K.; Orr-Weaver, T.L. Fundamental differences in endoreplication in mammals and *Drosophila* revealed by analysis of endocycling and endomitotic cells. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 9368–9373. [CrossRef]
7. Orr-Weaver, T.L. When bigger is better: The role of polyploidy in organogenesis. *Trends Genet.* **2015**, *31*, 307–315. [CrossRef]
8. Lazzeri, E.; Angelotti, M.L.; Conte, C.; Anders, H.-J.; Romagnani, P. Surviving Acute Organ Failure: Cell Polyploidization and Progenitor Proliferation. *Trends Mol. Med.* **2019**, *25*, 366–381. [CrossRef]
9. Derks, W.; Bergmann, O. Polyploidy in Cardiomyocytes: Roadblock to Heart Regeneration? *Circ. Res.* **2020**, *126*, 552–565. [CrossRef]
10. Gan, P.; Patterson, M.; Sucov, H.M. Cardiomyocyte Polyploidy and Implications for Heart Regeneration. *Annu. Rev. Physiol.* **2020**, *82*, 45–61. [CrossRef]
11. Anatskaya, O.V.; Vinogradov, A.E.; Kudryavtsev, B.N. Cardiomyocyte ploidy levels in birds with different growth rates. *J. Exp. Zool.* **2001**, *289*, 48–58. [CrossRef]
12. Anatskaya, O.V.; Vinogradov, A.E. Genome multiplication as adaptation to tissue survival: Evidence from gene expression in mammalian heart and liver. *Genomics* **2007**, *89*, 70–80. [CrossRef] [PubMed]
13. Silva, I.S.; Ghiraldini, F.G.; Veronezi, G.M.B.; Mello, M.L.S. Polyploidy and nuclear phenotype characteristics of cardiomyocytes from diabetic adult and normoglycemic aged mice. *Acta Histochem.* **2018**, *120*, 84–94. [CrossRef] [PubMed]
14. Bensley, J.G.; Stacy, V.K.; De Matteo, R.; Harding, R.; Black, M.J. Cardiac remodelling as a result of pre-term birth: Implications for future cardiovascular disease. *Eur. Heart J.* **2010**, *31*, 2058–2066. [CrossRef]
15. Gjelsvik, K.J.; Besen-McNally, R.; Losick, V.P. Solving the Polyploid Mystery in Health and Disease. *Trends Genet.* **2019**, *35*, 6–14. [CrossRef]
16. Donne, R.; Saroul-Aïnama, M.; Cordier, P.; Celton-Morizur, S.; Desdouets, C. Polyploidy in liver development, homeostasis and disease. *Nat. Rev. Gastroenterol. Hepatol.* **2020**, *17*, 391–405. [CrossRef]
17. Storchova, Z.; Pellman, D. From polyploidy to aneuploidy, genome instability and cancer. *Nat. Rev. Mol. Cell Biol.* **2004**, *5*, 45–54. [CrossRef]
18. Coward, J.; Harding, A. Size Does Matter: Why Polyploid Tumor Cells are Critical Drug Targets in the War on Cancer. *Front. Oncol.* **2014**, *4*, 123. [CrossRef]
19. Ben-Porath, I.; Thomson, M.W.; Carey, V.J.; Ge, R.; Bell, G.W.; Regev, A.; Weinberg, R.A. An embryonic stem cell-like gene expression signature in poorly differentiated aggressive human tumors. *Nat. Genet.* **2008**, *40*, 499–507. [CrossRef]

20. Salmina, K.; Jankevics, E.; Huna, A.; Perminov, D.; Radovica, I.; Klymenko, T.; Ivanov, A.; Jascenko, E.; Scherthan, H.; Cragg, M.; et al. Up-regulation of the embryonic self-renewal network through reversible polyploidy in irradiated p53-mutant tumour cells. *Exp. Cell Res.* **2010**, *316*, 2099–2112. [CrossRef]

21. Nguyen, H.G.; Ravid, K. Polyploidy: Mechanisms and Cancer Promotion in Hematopoietic and Other Cells. In *Polyploidization and Cancer*; Poon, R.Y.C., Ed.; Advances in Experimental Medicine and Biology; Springer: New York, NY, USA, 2010; Volume 676, pp. 105–122. ISBN 978-1-4419-6198-3.

22. Lagadec, C.; Vlashi, E.; Della Donna, L.; Dekmezian, C.; Pajonk, F. Radiation-Induced Reprogramming of Breast Cancer Cells: Radiation-Induced Cancer Stem Cells. *Stem Cells* **2012**, *30*, 833–844. [CrossRef] [PubMed]

23. Erenpreisa, J.; Cragg, M.S. Cancer: A matter of life cycle? *Cell Biol. Int.* **2007**, *31*, 1507–1510. [CrossRef] [PubMed]

24. Erenpreisa, J.; Cragg, M.S. Life-Cycle Features of Tumour Cells. In *Evolutionary Biology from Concept to Application*; Pontarotti, P., Ed.; Springer: Berlin/Heidelberg, Germany, 2008; pp. 61–71. ISBN 978-3-540-78992-5.

25. Mirzayans, R.; Andrais, B.; Scott, A.; Wang, Y.; Kumar, P.; Murray, D. Multinucleated Giant Cancer Cells Produced in Response to Ionizing Radiation Retain Viability and Replicate Their Genome. *Int. J. Mol. Sci.* **2017**, *18*, 360. [CrossRef] [PubMed]

26. Liu, J. The dualistic origin of human tumors. *Semin. Cancer Biol.* **2018**, *53*, 1–16. [CrossRef] [PubMed]

27. Salmina, K.; Bojko, A.; Inashkina, I.; Staniak, K.; Dudkowska, M.; Podlesniy, P.; Rumnieks, F.; Vainshelbaum, N.M.; Pjanova, D.; Sikora, E.; et al. "Mitotic Slippage" and Extranuclear DNA in Cancer Chemoresistance: A Focus on Telomeres. *Int. J. Mol. Sci.* **2020**, *21*, 2779. [CrossRef] [PubMed]

28. Mirzayans, R.; Andrais, B.; Murray, D. Roles of Polyploid/Multinucleated Giant Cancer Cells in Metastasis and Disease Relapse Following Anticancer Treatment. *Cancers* **2018**, *10*, 118. [CrossRef] [PubMed]

29. Vincent, M.D. Cancer: Beyond Speciation. In *Advances in Cancer Research*; Elsevier: Amsterdam, The Netherlands, 2011; Volume 112, pp. 283–350. ISBN 978-0-12-387688-1.

30. Davies, P.C.W.; Lineweaver, C.H. Cancer tumors as Metazoa 1.0: Tapping genes of ancient ancestors. *Phys. Biol.* **2011**, *8*, 015001. [CrossRef] [PubMed]

31. Davies, P. Exposing cancer's deep evolutionary roots. *Phys. Cancer Phys. World* **2013**, *26*, 37–40. [CrossRef]

32. Aktipis, C.A.; Boddy, A.M.; Jansen, G.; Hibner, U.; Hochberg, M.E.; Maley, C.C.; Wilkinson, G.S. Cancer across the tree of life: Cooperation and cheating in multicellularity. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **2015**, *370*. [CrossRef]

33. Niculescu, V.F. Developmental and Non Developmental Polyploidy in Xenic and Axenic Cultured Stem Cell Lines of *Entamoeba invadens* and *E. histolytica*. *Insights Stem Cells* **2016**, *2*, 1–9.

34. Vinogradov, A.E. Human transcriptome nexuses: Basic-eukaryotic and metazoan. *Genomics* **2010**, *95*, 345–354. [CrossRef] [PubMed]

35. Vinogradov, A.E.; Anatskaya, O.V. Evolutionary framework of the human interactome: Unicellular and multicellular giant clusters. *Biosystems* **2019**, *181*, 82–87. [CrossRef] [PubMed]

36. Domazet-Lošo, T.; Tautz, D. Phylostratigraphic tracking of cancer genes suggests a link to the emergence of multicellularity in metazoa. *BMC Biol.* **2010**, *8*, 66. [CrossRef] [PubMed]

37. Gu, X.; Zou, Z.; Yang, J. Tracing Evolutionary Ages of Cancer-Driving Sites by Cancer Somatic Mutations. *bioRxiv* **2020**. [CrossRef]

38. Trigos, A.S.; Pearson, R.B.; Papenfuss, A.T.; Goode, D.L. Altered interactions between unicellular and multicellular genes drive hallmarks of transformation in a diverse range of solid tumors. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6406–6411. [CrossRef] [PubMed]

39. Bussey, K.J.; Cisneros, L.H.; Lineweaver, C.H.; Davies, P.C.W. Ancestral gene regulatory networks drive cancer. *Proc. Natl. Acad. Sci. USA* **2017**, *114*, 6160–6162. [CrossRef] [PubMed]

40. Trigos, A.S.; Pearson, R.B.; Papenfuss, A.T.; Goode, D.L. How the evolution of multicellularity set the stage for cancer. *Br. J. Cancer* **2018**, *118*, 145–152. [CrossRef] [PubMed]

41. Vinogradov, A.E.; Anatskaya, O.V. Cell-cycle dependence of transcriptome gene modules: Comparison of regression lines. *FEBS J.* **2020**, *287*, 4427–4439. [CrossRef]

42. Bizzarri, M.; Giuliani, A.; Cucina, A.; D'Anselmi, F.; Soto, A.M.; Sonnenschein, C. Fractal analysis in a systems biology approach to cancer. *Semin. Cancer Biol.* **2011**, *21*, 175–182. [CrossRef]

43. Bizzarri, M.; Giuliani, A.; Minini, M.; Monti, N.; Cucina, A. Constraints Shape Cell Function and Morphology by Canalizing the Developmental Path along the Waddington's Landscape. *Bioessays* **2020**, *42*, e1900108. [CrossRef]

44. Erenpreisa, J.; Cragg, M.S. Three steps to the immortality of cancer cells: Senescence, polyploidy and self-renewal. *Cancer Cell Int.* **2013**, *13*, 92. [CrossRef] [PubMed]

45. Van de Peer, Y.; Mizrachi, E.; Marchal, K. The evolutionary significance of polyploidy. *Nat. Rev. Genet.* **2017**, *18*, 411–424. [CrossRef] [PubMed]

46. Moein, S.; Adibi, R.; da Silva Meirelles, L.; Nardi, N.B.; Gheisari, Y. Cancer regeneration: Polyploid cells are the key drivers of tumor progression. *Biochim. Biophys. Acta (BBA) Rev. Cancer* **2020**, 188408. [CrossRef] [PubMed]

47. Mirzayans, R.; Murray, D. Intratumor Heterogeneity and Therapy Resistance: Contributions of Dormancy, Apoptosis Reversal (Anastasis) and Cell Fusion to Disease Recurrence. *Int. J. Mol. Sci.* **2020**, *21*, 1308. [CrossRef] [PubMed]

48. Bernstein, B.E.; Mikkelsen, T.S.; Xie, X.; Kamal, M.; Huebert, D.J.; Cuff, J.; Fry, B.; Meissner, A.; Wernig, M.; Plath, K.; et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* **2006**, *125*, 315–326. [CrossRef] [PubMed]

49. Mikkelsen, T.S.; Ku, M.; Jaffe, D.B.; Issac, B.; Lieberman, E.; Giannoukos, G.; Alvarez, P.; Brockman, W.; Kim, T.-K.; Koche, R.P.; et al. Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* **2007**, *448*, 553–560. [CrossRef]

50. Widschwendter, M.; Fiegl, H.; Egle, D.; Mueller-Holzner, E.; Spizzo, G.; Marth, C.; Weisenberger, D.J.; Campan, M.; Young, J.; Jacobs, I.; et al. Epigenetic stem cell signature in cancer. *Nat. Genet.* **2007**, *39*, 157–158. [CrossRef]

51. Balch, C.; Nephew, K.P.; Huang, T.H.-M.; Bapat, S.A. Epigenetic "bivalently marked" process of cancer stem cell-driven tumorigenesis. *Bioessays* **2007**, *29*, 842–845. [CrossRef]

52. Blanco, E.; González-Ramírez, M.; Alcaine-Colet, A.; Aranda, S.; Di Croce, L. The Bivalent Genome: Characterization, Structure, and Regulation. *Trends Genet.* **2020**, *36*, 118–131. [CrossRef]

53. Bernhart, S.H.; Kretzmer, H.; Holdt, L.M.; Jühling, F.; Ammerpohl, O.; Bergmann, A.K.; Northoff, B.H.; Doose, G.; Siebert, R.; Stadler, P.F.; et al. Changes of bivalent chromatin coincide with increased expression of developmental genes in cancer. *Sci. Rep.* **2016**, *6*, 37393. [CrossRef]

54. Court, F.; Arnaud, P. An annotated list of bivalent chromatin regions in human ES cells: A new tool for cancer epigenetic research. *Oncotarget* **2017**, *8*, 4110–4124. [CrossRef] [PubMed]

55. Kuroda, M.I.; Kang, H.; De, S.; Kassis, J.A. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. *Annu. Rev. Biochem.* **2020**, *89*, 235–253. [CrossRef] [PubMed]

56. Sonnenschein, C.; Soto, A.M. Cancer Metastases: So Close and So Far. *JNCI J. Natl. Cancer Inst.* **2015**, *107*. [CrossRef] [PubMed]

57. Li, Q.; Dang, C.V. Myc Overexpression Uncouples DNA Replication from Mitosis. *Mol. Cell. Biol.* **1999**, *19*, 5339–5351. [CrossRef] [PubMed]

58. Grendler, J.; Lowgren, S.; Mills, M.; Losick, V.P. Wound-induced polyploidization is driven by Myc and supports tissue repair in the presence of DNA damage. *Development* **2019**, *146*, dev173005. [CrossRef]

59. Soucek, L.; Whitfield, J.; Martins, C.P.; Finch, A.J.; Murphy, D.J.; Sodir, N.M.; Karnezis, A.N.; Swigart, L.B.; Nasi, S.; Evan, G.I. Modelling Myc inhibition as a cancer therapy. *Nature* **2008**, *455*, 679–683. [CrossRef]

60. Evan, G. Taking a Back Door to Target Myc. *Science* **2012**, *335*, 293–294. [CrossRef]

61. Kortlever, R.M.; Sodir, N.M.; Wilson, C.H.; Burkhart, D.L.; Pellegrinet, L.; Brown Swigart, L.; Littlewood, T.D.; Evan, G.I. Myc Cooperates with Ras by Programming Inflammation and Immune Suppression. *Cell* **2017**, *171*, 1301–1315. [CrossRef]

62. Sui, Y.; Gu, R.; Janknecht, R. Crucial Functions of the JMJD1/KDM3 Epigenetic Regulators in Cancer. *Mol. Cancer Res.* **2020**. [CrossRef]

63. Neri, F.; Zippo, A.; Krepelova, A.; Cherubini, A.; Rocchigiani, M.; Oliviero, S. Myc Regulates the Transcription of the PRC2 Gene to Control the Expression of Developmental Genes in Embryonic Stem Cells. *Mol. Cell. Biol.* **2012**, *32*, 840–851. [CrossRef]

64. Ullius, A.; Luscher-Firzlaff, J.; Costa, I.G.; Walsemann, G.; Forst, A.H.; Gusmao, E.G.; Kapelle, K.; Kleine, H.; Kremmer, E.; Vervoorts, J.; et al. The interaction of MYC with the trithorax protein ASH2L promotes gene transcription by regulating H3K27 modification. *Nucleic Acids Res.* **2014**, *42*, 6901–6920. [CrossRef] [PubMed]

65. Vinogradov, A.E.; Shilina, M.A.; Anatskaya, O.V.; Alekseenko, L.L.; Fridlyanskaya, I.I.; Krasnenko, A.; Kim, A.; Korostin, D.; Ilynsky, V.; Elmuratov, A.; et al. Molecular Genetic Analysis of Human Endometrial Mesenchymal Stem Cells That Survived Sublethal Heat Shock. *Stem Cells Int* **2017**, *17*, 2362630. [CrossRef]

66. Kallioniemi, A. Bone morphogenetic protein 4—A fascinating regulator of cancer cell behavior. *Cancer Genet.* **2012**, *205*, 267–277. [CrossRef] [PubMed]

67. Bach, D.-H.; Park, H.J.; Lee, S.K. The Dual Role of Bone Morphogenetic Proteins in Cancer. *Mol. Ther. Oncolytics* **2018**, *8*, 1–13. [CrossRef] [PubMed]

68. Chen, H.; Liu, H.; Qing, G. Targeting oncogenic Myc as a strategy for cancer treatment. *Signal Transduct. Target. Ther.* **2018**, *3*, 5. [CrossRef] [PubMed]

69. Tovar, E.A.; Graveel, C.R. MET in human cancer: Germline and somatic mutations. *Ann. Transl. Med.* **2017**, *5*, 205. [CrossRef]

70. Mitra, P. Transcription regulation of MYB: A potential and novel therapeutic target in cancer. *Ann. Transl. Med.* **2018**, *6*, 443. [CrossRef]

71. Erenpreiss, J.O. *Current Concepts of Malignant Growth*; Zinâtne Publ.: Riga, Latvia, 1993.

72. Moore, A.R.; Rosenberg, S.C.; McCormick, F.; Malek, S. RAS-targeted therapies: Is the undruggable drugged? *Nat. Rev. Drug Discov.* **2020**, *19*, 533–552. [CrossRef]

73. Sigismund, S.; Avanzato, D.; Lanzetti, L. Emerging functions of the EGFR in cancer. *Mol. Oncol.* **2018**, *12*, 3–20. [CrossRef]

74. Hosaka, K.; Yang, Y.; Seki, T.; Du, Q.; Jing, X.; He, X.; Wu, J.; Zhang, Y.; Morikawa, H.; Nakamura, M.; et al. Therapeutic paradigm of dual targeting VEGF and PDGF for effectively treating FGF-2 off-target tumors. *Nat. Commun.* **2020**, *11*, 3704. [CrossRef]

75. Khurana, N.; Sikka, S.C. Interplay between SOX9, Wnt/β-Catenin and Androgen Receptor Signaling in Castration-Resistant Prostate Cancer. *Int. J. Mol. Sci.* **2019**, *20*, 2066. [CrossRef] [PubMed]

76. Qin, Q.; Xu, Y.; He, T.; Qin, C.; Xu, J. Normal and disease-related biological functions of Twist1 and underlying molecular mechanisms. *Cell Res.* **2012**, *22*, 90–106. [CrossRef] [PubMed]

77. Petrizzo, A.; Caruso, F.P.; Tagliamonte, M.; Tornesello, M.L.; Ceccarelli, M.; Costa, V.; Aprile, M.; Esposito, R.; Ciliberto, G.; Buonaguro, F.M.; et al. Identification and Validation of HCC-specific Gene Transcriptional Signature for Tumor Antigen Discovery. *Sci. Rep.* **2016**, *6*, 29258. [CrossRef] [PubMed]

78. Davoli, T.; Xu, A.W.; Mengwasser, K.E.; Sack, L.M.; Yoon, J.C.; Park, P.J.; Elledge, S.J. Cumulative Haploinsufficiency and Triplosensitivity Drive Aneuploidy Patterns and Shape the Cancer Genome. *Cell* **2013**, *155*, 948–962. [CrossRef]

79. Justice, J.; Malhotra, S.; Ruano, M.; Li, Y.; Zavala, G.; Lee, N.; Morgan, R.; Beemon, K. The MET Gene Is a Common Integration Target in Avian Leukosis Virus Subgroup J-Induced Chicken Hemangiomas. *J. Virol.* **2015**, *89*, 4712–4719. [CrossRef] [PubMed]

80. Pienta, K.J.; Hammarlund, E.U.; Axelrod, R.; Amend, S.R.; Brown, J.S. Convergent Evolution, Evolving Evolvability, and the Origins of Lethal Cancer. *Mol. Cancer Res.* **2020**, *18*, 801–810. [CrossRef]

81. Walther, V.; Hiley, C.T.; Shibata, D.; Swanton, C.; Turner, P.E.; Maley, C.C. Can oncology recapitulate paleontology? Lessons from species extinctions. *Nat. Rev. Clin. Oncol.* **2015**, *12*, 273–285. [CrossRef]

82. Versteeg, R. Tumours outside the mutation box. *Nature* **2014**, *506*, 438–439. [CrossRef]

83. Gatenby, R.A. Is the Genetic Paradigm of Cancer Complete? *Radiology* **2017**, *284*, 1–3. [CrossRef]

84. Ganem, N.J.; Storchova, Z.; Pellman, D. Tetraploidy, aneuploidy and cancer. *Curr. Opin. Genet. Dev.* **2007**, *17*, 157–162. [CrossRef]

85. López, S.; Lim, E.L.; Horswell, S.; Haase, K.; Huebner, A.; Dietzen, M.; Mourikis, T.P.; Watkins, T.B.K.; Rowan, A.; Dewhurst, S.M.; et al. Interplay between whole-genome doubling and the accumulation of deleterious alterations in cancer evolution. *Nat. Genet.* **2020**, *52*, 283–293. [CrossRef] [PubMed]

86. Beach, R.R.; Ricci-Tam, C.; Brennan, C.M.; Moomau, C.A.; Hsu, P.-H.; Hua, B.; Silberman, R.E.; Springer, M.; Amon, A. Aneuploidy Causes Non-genetic Individuality. *Cell* **2017**, *169*, 229–242.e21. [CrossRef] [PubMed]

87. Salmina, K.; Huna, A.; Kalejs, M.; Pjanova, D.; Scherthan, H.; Cragg, M.S.; Erenpreisa, J. The Cancer Aneuploidy Paradox: In the Light of Evolution. *Genes* **2019**, *10*, 83. [CrossRef] [PubMed]

88. Mojtahedi, M.; Skupin, A.; Zhou, J.; Castaño, I.G.; Leong-Quong, R.Y.Y.; Chang, H.; Trachana, K.; Giuliani, A.; Huang, S. Cell Fate Decision as High-Dimensional Critical State Transition. *PLoS Biol.* **2016**, *14*, e2000640. [CrossRef]

89. Cavalli, G. Chromosome kissing. *Curr. Opin. Genet. Dev.* **2007**, *17*, 443–450. [CrossRef]

90. Lieberman-Aiden, E.; van Berkum, N.L.; Williams, L.; Imakaev, M.; Ragoczy, T.; Telling, A.; Amit, I.; Lajoie, B.R.; Sabo, P.J.; Dorschner, M.O.; et al. Comprehensive Mapping of Long-Range Interactions Reveals Folding Principles of the Human Genome. *Science* **2009**, *326*, 289–293. [CrossRef]

91. Solovei, I.; Thanisch, K.; Feodorova, Y. How to rule the nucleus: Divide et impera. *Curr. Opin. Cell Biol.* **2016**, *40*, 47–59. [CrossRef]

92. Bakhmet, E.I.; Nazarov, I.B.; Gazizova, A.R.; Vorobyeva, N.E.; Kuzmin, A.A.; Gordeev, M.N.; Sinenko, S.A.; Aksenov, N.D.; Artamonova, T.O.; Khodorkovskii, M.A.; et al. hnRNP-K Targets Open Chromatin in Mouse Embryonic Stem Cells in Concert with Multiple Regulators: hnRNP-K Targets Open Chromatin in Mouse ESCs. *Stem Cells* **2019**, *37*, 1018–1029. [CrossRef]

93. Nazarov, I.B.; Bakhmet, E.I.; Tomilin, A.N. KH-Domain Poly(C)-Binding Proteins as Versatile Regulators of Multiple Biological Processes. *Biochem. Mosc.* **2019**, *84*, 205–219. [CrossRef]

94. Kauffman, S.A. *The Origins of Order: Self-Organization and Selection in Evolution*; Series on Directions in Condensed Matter Physics; World Scientific: Singapore, 1992; Volume 6, pp. 61–100. ISBN 978-9971-5-0537-0.

95. Conant, G.C. Rapid reorganization of the transcriptional regulatory network after genome duplication in yeast. *Proc. Biol. Sci.* **2010**, *277*, 869–876. [CrossRef]

96. Conant, G.C. The lasting after-effects of an ancient polyploidy on the genomes of teleosts. *PLoS ONE* **2020**, *15*, e0231356. [CrossRef] [PubMed]

97. Cherstvy, A.G.; Teif, V.B. Structure-driven homology pairing of chromatin fibers: The role of electrostatics and protein-induced bridging. *J. Biol. Phys.* **2013**, *39*, 363–385. [CrossRef] [PubMed]

98. Jerković, I.; Szabo, Q.; Bantignies, F.; Cavalli, G. Higher-Order Chromosomal Structures Mediate Genome Function. *J. Mol. Biol.* **2020**, *432*, 676–681. [CrossRef]

99. Vinogradov, A.E. Buffering: A Possible Passive-homeostasis Role for Redundant DNA. *J. Theor. Biol.* **1998**, *193*, 197–199. [CrossRef] [PubMed]

100. Vinogradov, A.E. Genome size and chromatin condensation in vertebrates. *Chromosoma* **2005**, *113*, 362–369. [CrossRef] [PubMed]

101. Kanehisa, M.; Sato, Y.; Furumichi, M.; Morishima, K.; Tanabe, M. New approach for understanding genome variations in KEGG. *Nucleic Acids Res.* **2019**, *47*, D590–D595. [CrossRef]

102. Land, H.; Parada, L.F.; Weinberg, R.A. Tumorigenic conversion of primary embryo fibroblasts requires at least two cooperating oncogenes. *Nature* **1983**, *304*, 596–602. [CrossRef]

103. Knudson, A.G. Two genetic hits (more or less) to cancer. *Nat. Rev. Cancer* **2001**, *1*, 157–162. [CrossRef]

104. Wang, C.; Lisanti, M.P.; Liao, D.J. Reviewing once more the c-myc and Ras collaboration: Converging at the cyclin D1-CDK4 complex and challenging basic concepts of cancer biology. *Cell Cycle* **2011**, *10*, 57–67. [CrossRef]

105. Huang, S.; Ernberg, I.; Kauffman, S. Cancer attractors: A systems view of tumors from a gene network dynamics and developmental perspective. *Semin. Cell Dev. Biol.* **2009**, *20*, 869–876. [CrossRef]

106. Huang, S.; Kauffman, S. How to escape the cancer attractor: Rationale and limitations of multi-target drugs. *Semin. Cancer Biol.* **2013**, *23*, 270–278. [CrossRef] [PubMed]

107. Erenpreisa, J.; Roach, H.I. Epigenetic selection as a possible component of transdifferentiation. Further study of the commitment of hypertrophic chondrocytes to become osteocytes. *Mech. Ageing Dev.* **1996**, *87*, 165–182. [CrossRef]

108. Erenpreisa, J.; Salmiņa, K.; Belyayev, A.; Inashkina, I.; Cragg, M.S. Survival at the Brink. In *Autophagy: Cancer, Other Pathologies, Inflammation, Immunity, Infection, and Aging*; Elsevier: Amsterdam, The Netherlands, 2017; pp. 275–294. ISBN 978-0-12-812146-7.

109. Domazet-Lošo, T.; Klimovich, A.; Anokhin, B.; Anton-Erxleben, F.; Hamm, M.J.; Lange, C.; Bosch, T.C.G. Naturally occurring tumours in the basal metazoan Hydra. *Nat. Commun.* **2014**, *5*, 4222. [CrossRef] [PubMed]

110. Arthur, W. *Evolution: A Developmental Approach*; Wiley-Blackwell: Chichester, UK, 2011; ISBN 978-1-4051-8658-2.

111. Erenpreisa, J.; Salmina, K.; Huna, A.; Jackson, T.R.; Vazquez-Martin, A.; Cragg, M.S. The "virgin birth", polyploidy, and the origin of cancer. *Oncoscience* **2014**, *2*, 3–14. [CrossRef]

112. Niu, N.; Mercado-Uribe, I.; Liu, J. Dedifferentiation into blastomere-like cancer stem cells via formation of polyploid giant cancer cells. *Oncogene* **2017**, *36*, 4887–4900. [CrossRef] [PubMed]

113. Fais, S.; Overholtzer, M. Cell-in-cell phenomena, cannibalism, and autophagy: Is there a relationship? *Cell Death Dis.* **2018**, *9*, 95. [CrossRef] [PubMed]

114. Goodkov, A.V.; Berdieva, M.A.; Podlipaeva, Y.I.; Demin, S.Y. The Chromatin Extrusion Phenomenon in *Amoeba proteus* Cell Cycle. *J. Eukaryot. Microbiol.* **2020**, *67*, 203–208. [CrossRef]

115. Arguello, F. Atavistic Metamorphosis: A New and Logical Explanation for the Origin and Biological Nature of Cancer: With a Discussion on a Novel Approach to Treat Cancer. CreateSpace: Scotts Valley, CA, USA, 2011; ISBN 978-1-4609-6899-4.

116. Lineweaver, C.H.; Davies, P.C.W.; Vincent, M.D. Targeting cancer's weaknesses (not its strengths): Therapeutic strategies suggested by the atavistic model: Insights & Perspectives. *BioEssays* **2014**, *36*, 827–835. [CrossRef] [PubMed]

117. Raikov, I.B. Nuclear apparatus of Mesokaryotic Protozoa. In *The Protozoan Nucleus, Morphology and Evolution*; Cell Biology Monographs; Springer: New York, NY, USA, 1982; pp. 1–174.

118. Demin, S.Y.; Berdieva, M.A.; Podlipaeva, Y.I.; Goodkov, A.V. Karyotypic instability of endoprophase and mitotic cells of *Amoeba* sp. strain Cont from the "proteus-type" group (Amoebozoa, Euamoebida, Amoebidae). *Eur. J. Protistol.* **2020**, *74*, 125691. [CrossRef]

119. Niculescu, V.F. The reproductive life cycle of cancer: Hypotheses of cell of origin, TP53 drivers and stem cell conversions in the light of the atavistic cancer cell theory. *Med Hypotheses* **2019**, *123*, 19–23. [CrossRef]

120. Nanjundiah, V. Cellular slime mold development as a paradigm for the transition from unicellular to multicellular life. In *Multicellularity: Origins and Evolution*; MIT Press: Cambridge, MA, USA, 2016.

121. Andreani, T.S.; Itoh, T.Q.; Yildirim, E.; Hwangbo, D.-S.; Allada, R. Genetics of Circadian Rhythms. *Sleep Med. Clin.* **2015**, *10*, 413–421. [CrossRef] [PubMed]

122. Camponeschi, I.; Damasco, A.; Uversky, V.N.; Giuliani, A.; Bianchi, M.M. Phenotypic suppression caused by resonance with light-dark cycles indicates the presence of a 24-hours oscillator in yeast and suggests a new role of intrinsically disordered protein regions as internal mediators. *J. Biomol. Struct. Dyn.* **2020**. [CrossRef] [PubMed]

123. Øvrebø, J.I.; Edgar, B.A. Polyploidy in tissue homeostasis and regeneration. *Development* **2018**, *145*. [CrossRef] [PubMed]

124. Vazquez-Martin, A.; Anatskaya, O.V.; Giuliani, A.; Erenpreisa, J.; Huang, S.; Salmina, K.; Inashkina, I.; Huna, A.; Nikolsky, N.N.; Vinogradov, A.E. Somatic polyploidy is associated with the upregulation of c-MYC interacting genes and EMT-like signature. *Oncotarget* **2016**, *7*, 75235–75260. [CrossRef] [PubMed]

125. Whitehead, A.; Crawford, D.L. Variation within and among species in gene expression: Raw material for evolution: Review of gene expression variation. *Mol. Ecol.* **2006**, *15*, 1197–1211. [CrossRef]

126. Otto, S.P. The Evolutionary Consequences of Polyploidy. *Cell* **2007**, *131*, 452–462. [CrossRef]

127. Brawand, D.; Soumillon, M.; Necsulea, A.; Julien, P.; Csárdi, G.; Harrigan, P.; Weier, M.; Liechti, A.; Aximu-Petri, A.; Kircher, M.; et al. The evolution of gene expression levels in mammalian organs. *Nature* **2011**, *478*, 343–348. [CrossRef]

128. Ritchie, M.E.; Phipson, B.; Wu, D.; Hu, Y.; Law, C.W.; Shi, W.; Smyth, G.K. limma powers differential expression analyses for RNA-sequencing and microarray studies. *Nucleic Acids Res.* **2015**, *43*, e47. [CrossRef]

129. Vinogradov, A.E.; Anatskaya, O.V. Systemic evolutionary changes in mammalian gene expression. *Biosystems* **2020**, *198*, 104256. [CrossRef]

130. Sayers, E.W.; Beck, J.; Brister, J.R.; Bolton, E.E.; Canese, K.; Comeau, D.C.; Funk, K.; Ketter, A.; Kim, S.; Kimchi, A.; et al. Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.* **2020**, *48*, D9–D16. [CrossRef]

131. Seyednasrollah, F.; Laiho, A.; Elo, L.L. Comparison of software packages for detecting differential expression in RNA-seq studies. *Brief. Bioinform.* **2015**, *16*, 59–70. [CrossRef] [PubMed]

132. Giuliani, A. The application of principal component analysis to drug discovery and biomedical data. *Drug Discov. Today* **2017**, *22*, 1069–1076. [CrossRef] [PubMed]

133. Roden, J.C.; King, B.W.; Trout, D.; Mortazavi, A.; Wold, B.J.; Hart, C.E. Mining gene expression data by interpreting principal components. *BMC Bioinform.* **2006**, *7*, 194. [CrossRef] [PubMed]

134. Vinogradov, A.E. 'Genome design' model and multicellular complexity: Golden middle. *Nucleic Acids Res.* **2006**, *34*, 5906–5914. [CrossRef]

135. Vinogradov, A.E.; Anatskaya, O.V. DNA helix: The importance of being AT-rich. *Mamm. Genome* **2017**, *28*, 455–464. [CrossRef]

136. The Gene Ontology Consortium. The Gene Ontology Resource: 20 years and still GOing strong. *Nucleic Acids Res.* **2019**, *47*, D330–D338. [CrossRef]

137. Storey, J.D.; Tibshirani, R. Statistical significance for genomewide studies. *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9440–9445. [CrossRef]

138. Szklarczyk, D.; Gable, A.L.; Lyon, D.; Junge, A.; Wyder, S.; Huerta-Cepas, J.; Simonovic, M.; Doncheva, N.T.; Morris, J.H.; Bork, P.; et al. STRING v11: Protein–protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **2019**, *47*, D607–D613. [CrossRef]