

RESEARCH ARTICLE

Using digital organisms to study the evolutionary consequences of whole genome duplication and polyploidy

Yao Yao^{1,2}, Lorenzo Carretero-Paulet^{1,2}, Yves Van de Peer^{1,2,3*}

1 Center for Plant Systems Biology, VIB, Ghent, Belgium, **2** Department of Plant Biotechnology and Bioinformatics, Ghent University, Ghent, Belgium, **3** Center for Microbial Ecology and Genomics, Department of Biochemistry, Genetics and Microbiology, University of Pretoria, Pretoria, South Africa

* yves.vandeppeer@psb.vib-ugent.be

OPEN ACCESS

Citation: Yao Y, Carretero-Paulet L, Van de Peer Y (2019) Using digital organisms to study the evolutionary consequences of whole genome duplication and polyploidy. PLoS ONE 14(7): e0220257. <https://doi.org/10.1371/journal.pone.0220257>

Editor: Marc Robinson-Rechavi, University of Lausanne, SWITZERLAND

Received: March 7, 2019

Accepted: July 11, 2019

Published: July 31, 2019

Copyright: © 2019 Yao et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: Our framework is an adaptation of work previously described in (doi: [10.1371/journal.pone.0090695](https://doi.org/10.1371/journal.pone.0090695)) and (doi: [10.7717/peerj.2812](https://doi.org/10.7717/peerj.2812)). The source code of the model is available at <https://bitbucket.org/YaoYao11/artificial-evolution/src/master/>.

Funding: Y.V.d.P. acknowledges funding from the European Union Seventh Framework Programme (FP7/2007-2013) under European Research Council Advanced Grant Agreement 322739 – DOUBLEUP. The authors would also like to thank

Abstract

The potential role of whole genome duplication (WGD) in evolution is controversial. Whereas some view WGD mainly as detrimental and an evolutionary ‘dead end’, there is growing evidence that the long-term establishment of polyploidy might be linked to environmental change, stressful conditions, or periods of extinction. However, despite much research, the mechanistic underpinnings of why and how polyploids might be able to out-compete non-polyploids at times of environmental upheaval remain indefinable. Here, we improved our recently developed bio-inspired framework, combining an artificial genome with an agent-based system, to form a population of so-called Digital Organisms (DOs), to examine the impact of WGD on evolution under different environmental scenarios mimicking extinction events of varying strength and frequency. We found that, under stable environments, DOs with non-duplicated genomes formed the majority, if not all, of the population, whereas the numbers of DOs with duplicated genomes increased under dramatically challenging environments. After tracking the evolutionary trajectories of individual genomes in terms of sequence and encoded gene regulatory networks (GRNs), we propose that duplicated GRNs might provide polyploids with better chances to acquire the drastic changes necessary to adapt to challenging conditions, thus endowing DOs with increased adaptive potential under extinction events. In contrast, under stable environments, random mutations might easily render the GRN less well adapted to such environments, a phenomenon that is exacerbated in duplicated, more complex GRNs. We believe that our results provide some additional insights into how genome duplication and polyploidy might help organisms to compete for novel niches and survive ecological turmoil, and confirm the usefulness of our computational simulation in studying the role of WGD in evolution and adaptation, helping to overcome some of the traditional limitations of evolution experiments with model organisms.

Introduction

Polyploidy, or the duplication of entire genomes, is a common phenomenon in the evolutionary history of many eukaryotic organisms, especially plants. Whole genome duplications

Arthur Zwaenepoel and Eshchar Mizrachi for helpful discussions. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interests: The authors have declared that no competing interests exist.

(WGDs) have often been linked to speciation and species diversification, increase in biological complexity, alleviating the effects caused by extinction events, and an overall increased environmental and mutational robustness [1–3]. Nevertheless, although the prevalence of WGDs has now been firmly established, the effects and significance of WGDs for evolution remain vividly discussed. Whereas some regard WGD mainly as an evolutionary dead end, others see WGD primarily as an opportunity and source of evolutionary innovation. In support of the former, there is the observation that recent WGDs seem to outnumber established ancient polyploidy events by several orders of magnitude [4–6]. This relative paucity of anciently established WGDs seems to suggest that many WGDs have not survived on the long run, which may be due to the well-known detrimental effects of polyploidy, caused by minority cytotype exclusion, genomic instability, mitotic and meiotic abnormalities, or epigenetics changes [7].

On the other hand, we do observe ‘ancient’ organisms that underwent and survived WGDs, so that their descendants have outcompeted their diploid progenitors and bear traces of the duplication in their genomes. This is the case for several major eukaryote lineages such as vertebrates [8, 9], fishes [10], ciliate protozoans [11], hemiascomycetous yeasts [12, 13], and particularly plants [14–16]. This observation, together with earlier work on specific adaptations of polyploids, led some to conclude that WGDs might provide an adaptive advantage particularly under unstable, challenging and stressful environments or during periods of environmental upheaval [6, 17]. This hypothesis is supported, among others, by the wave of successful WGDs that seems to have occurred around the Cretaceous-Paleogene extinction event or K-Pg boundary [18–21].

Although there is increasing support for a correlation between polyploid establishment and environmental challenge [6, 17], the molecular evolutionary mechanisms behind this remain vague. Ideally, to comprehensively investigate the effects of WGDs on evolution and study the link between adaptation and polyploidy, we need to collect detailed evolutionary trajectory data of organisms with and without WGDs (e.g., the complete mutational landscape, dynamic gene expression profiling data, real-time adaptation data, etc.) under different environmental contexts. However, although there have been some very interesting and successful attempts, collecting these data from real biological evolution experiments is difficult and bound by experimental limitations [22]. Computational approaches, notwithstanding many other limitations, have the advantage that they can collect at least some of such evolutionary trajectory data and have already shown great potential in investigating the potential adaptive role of WGD. For example, computational evolutionary modeling of populations of so-called ‘virtual cells’ has been used to study the adaptive potential of polyploidy in times of drastic and enduring environmental change [23]. Although WGD was established only in a minority of lineages, polyploids were significantly more successful at adaptation (and readaptation), and therefore WGD seemed, at least in some cases, a powerful mechanism to cope with environmental challenges. However, as far as we understand, the model of Cuypers and Hogeweg (23) does not allow modeling the connection of the gene regulatory network (GRN) of these virtual cells with a particular environment.

We have recently published a computational framework aimed at mimicking biological evolution [24–26]. Our framework is based on so-called ‘digital organisms’ (DOs), which have their own genomes replicating under specific mutation and WGD rates. Furthermore, the expression of every gene in the genome is regulated by a set of environmental signals, or sensors, and the orchestrated activity of regulatory genes defined in the genome. The overall expression pattern of all genes encoded in the genome defines the GRN of the DO, which in turn determines its behavior, defined by a set of actuators. The strengths and advantages of our bio-inspired agent-based modeling approach have already been tested under a relatively

simple setup, such as emergent swarm behavior and overall enhanced adaptation in a changing environment [24, 25].

Here, we expand our previously developed framework to study the effects of polyploidy in a challenging environment. By considering the adaptation of each individual DO as an interactive and developmental process in evolution, we have a better chance to observe when and under which circumstances polyploidy can pose a potential selective advantage to the organism or population as a whole. Because the effects and consequences of polyploidy are complex and probably also often lineage-specific, we would like to stress that in the present study, we only focus on how polyploidy or WGDs might affect the evolution of GRNs and what might be some of the major constraints on the evolution of GRNs, and hence adaptation of their hosts, subsequent to WGDs. We do not study (the probability of) short-term establishment of polyploidy, nor longer-term phenomena such as gene loss or the functional divergence of genes (sub- or neofunctionalization). Also, reproductive effects (e.g. sexual reproduction and recombination) and population-specific features, such as allele frequency changes following WGDs, are not considered here. Although we are of course aware that these are all very important to fully understand the significance of polyploidy and WGD, we nevertheless hope that, by focusing on certain molecular properties, some aspects of WGDs can be better understood.

By analyzing multiple simulations of populations of clonal DOs under various environmental contexts, we found that DOs with non-duplicated genomes conformed most, if not all, of the population under stable environments. In contrast, environmental stress or change usually resulted in higher rates of extinction in the population, whereas WGDs seemed to increase the possibility of survival under dramatically challenging environments. Furthermore, tracking mutational changes in populations led us to observe that the GRNs in duplicated genomes tend to show greater flexibility with fewer mutations compared to those in non-duplicated genomes. Our results provide some initial insights into how genome duplication might help organisms to diversify, compete for novel niches, and survive ecological turmoil.

Methods

Our framework is an adaptation of work previously described in detail elsewhere [24, 25]. In short, to improve the adaptability and robustness of digital organisms under changing environments, we have developed a swarm evolutionary DO framework which is inspired by the concept of complex adaptive systems in biological evolution [24]. In this framework, we use the behavior of a swarm of virtual organisms or individuals and agents (i.e. individual ‘entities’—programs or functions—that can interact with each other and with the environment) to resemble gene regulation and adaptation of biological organisms. Complex adaptive systems originate through self-organization at different levels (the genome, GRN, organism, population, . . .) leading to swarm DOs that can more efficiently cope with the challenge of changing environments. Such a model did not only demonstrate the advantage of testing the adaptability of DOs in certain environments or under certain conditions [25, 27], but also provides a way to simulate, at least to some extent, the dynamic evolutionary trajectories that occur in biological evolution. By tracking the behavior of both DOs and agents (representing, amongst other things, expressed genes or gene products, i.e. proteins) during simulation, we can recreate certain evolutionary contexts and replay evolutionary scenarios over and over again and study the emergence of particular adaptations. Specific extensions made to our framework include an increase of simulation platform (a grid of 1000 by 1000 cells), allowing a larger population to evolve (up to 20k DOs), and a much larger genome size (starting from 100 kb, but which can grow—through WGDs—to a size that is only limited by the amount of memory).

The source code of the model is available at <https://bitbucket.org/YaoYao11/artificial-evolution/src/master/>.

Description of the simulation framework

Below, we provide a summarized description of the different components of our simulation framework, as well as their interactions, focusing on the ones relevant for the current study. A more detailed description of the framework can be found elsewhere [24, 25].

Digital Organism (DO). Every DO is defined by a genome and a set of i) virtual sensors, which allow DOs to measure different ecological and physiological parameters and ii) actuators, which allow them to interact with each other and with the environment and define the overall behavior of the DO [25] (Fig 1). These actuators confer DOs the ability to replicate, move, attack, defend or search for food [24, 25]. For the purposes of this work, we will focus on the actuators replication and movement.

Genes and genomes. In our framework, the genome of a DO has been inspired by the model previously proposed by Reil [28], and consists of a randomly created string of four digits representing each of the four nucleotides with a total initial size of 100 kb. Genes in the genome are not pre-specified, but identified in the randomly built genome, and typically a genome contains about 150–200 genes (Fig 2). Up to 1000 different gene products can be encoded by a single genome, each defined by a unique pre-specified cis-sequence of four digits, and associated with a specific identity number and decay rate. Compared to the initial model of Reil, our model makes an explicit distinction among three kinds of genes; i) structural genes, which control every particular behavior of the DO or actuator (e.g., replication or

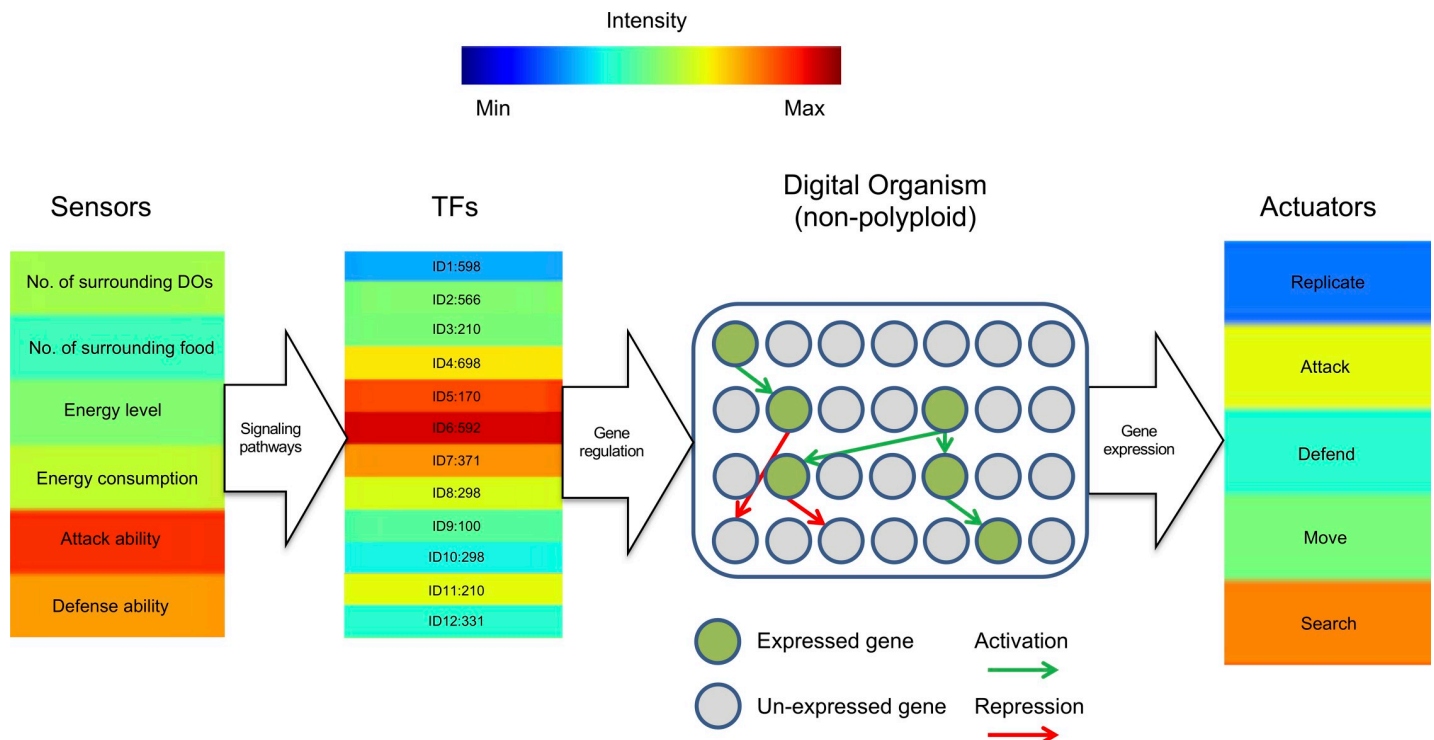


Fig 1. Schematic representation of the different components of our evolution simulation computational framework. At every time step in the simulation, the values of the sensors are converted by means of linear equations into a number (see Fig 2D). When this number matches the specific identity number of a particular gene product, there is an increase of its corresponding concentration level. The concentration level of all gene products determines the overall behavior of a DO, which is in turn defined by the set of possible actuators, each one encoded by its corresponding structural genes.

<https://doi.org/10.1371/journal.pone.0220257.g001>

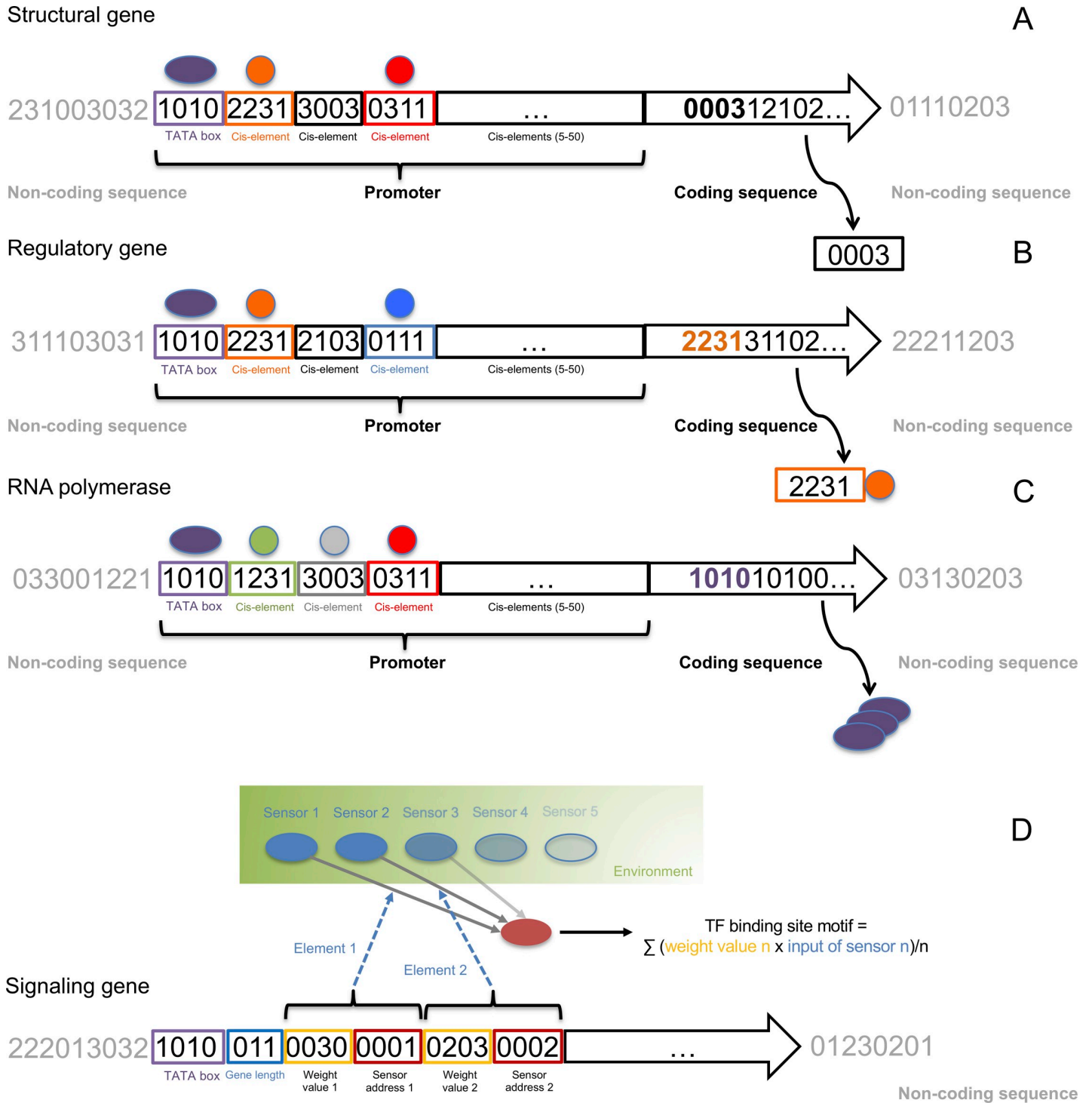


Fig 2. Every gene is defined by a promoter and a coding region. The promoter must start with a TATA box and be followed by a set of four-digit cis-elements numbering 5–50, each of which is bound by a specific kind of TF. The promoter region is followed by the coding region, whose first four digits define the specific kind of gene to be encoded. No interspersed sequence is allowed between any of the components of the gene. (A) Structural gene, which defines a specific kind of actuator. (B) Regulatory gene, which encodes a specific kind of TF defined by its cis-binding element (i.e., 2231). (C) RNA polymerase gene, which specifically binds to a TATA box (i.e., 1010). (D) Structure of a signaling gene, defining signaling pathways, which are defined as linear equations of sensor inputs. At each time step during the simulation, the value of each kind of sensor input (see Fig 1) is estimated and used as a variable in the equation. When the result of the equation corresponds to the identity number of a particular gene, its expression is activated. Linear equations are encoded in the sequence of the genome as shown. A special part of the gene sequence is specified for representing the equation and is composed of a group of ‘elements’, with each element having a fixed eight bases slot. The first four bases hold

the address of the corresponding sensor input data while the last four bases hold the weight for that particular sensor input. As such, we have a sensor input value and weight value for a particular sensor input. If we consider the particular sensor input value as a variable, a linear equation is obtained providing an output variable. A signaling pathway is then a combination of several such linear equations and their output values. The signaling gene also contains a field 'gene length' which will affect the number of 'elements' that can be present. Each signaling gene can have a minimum of three and at most 20 elements. These elements can be partially redundant. For example, two different elements may point to the same sensor inputs with different weight values.

<https://doi.org/10.1371/journal.pone.0220257.g002>

movement), either repressing or promoting it (Fig 2A); ii) regulatory genes, encoding transcription factors (TFs), which can specifically activate or repress (50% of each class) the expression of other genes by binding to specific cis-elements in their promoters (Fig 2B); and iii) a gene encoding 'RNA polymerase', which specifically binds to a 'TATA' box, which in turn must be present in the promoter of the target gene for proper expression (Fig 2C).

Gene expression. The expression level of a particular gene is calculated based on the combination of TF binding cis-elements found in the promoter of that gene (Fig 2). When the gene expression level reaches a minimum threshold, the gene is translated into its corresponding gene product. The total amount of gene product (i.e. dosage) corresponds to the expression level of that gene. At every time step during the simulation, the total amount of gene product decays gradually, mimicking protein degradation, under a specific decay rate that is randomly initialized for each gene at the beginning of the simulation. In turn, the decay of gene product is counteracted by the activity of activating TFs on that particular gene. The expression level of all gene products defines the global GRN of the DO, which is translated into a specific set of actuators [29], which in turn determines the overall behavior of the DO (Fig 3) [25].

Signaling pathways. Signaling pathways are defined as linear equations of sensor inputs (Fig 2D). At each time step during the simulation, the value of each kind of sensor input was estimated and used as a variable in the equation. Every DO has the ability to integrate different sensor input variables into the equation, including its energy level, the total number of surrounding food sources and/or DOs, and the ability to interact with other DOs through defense

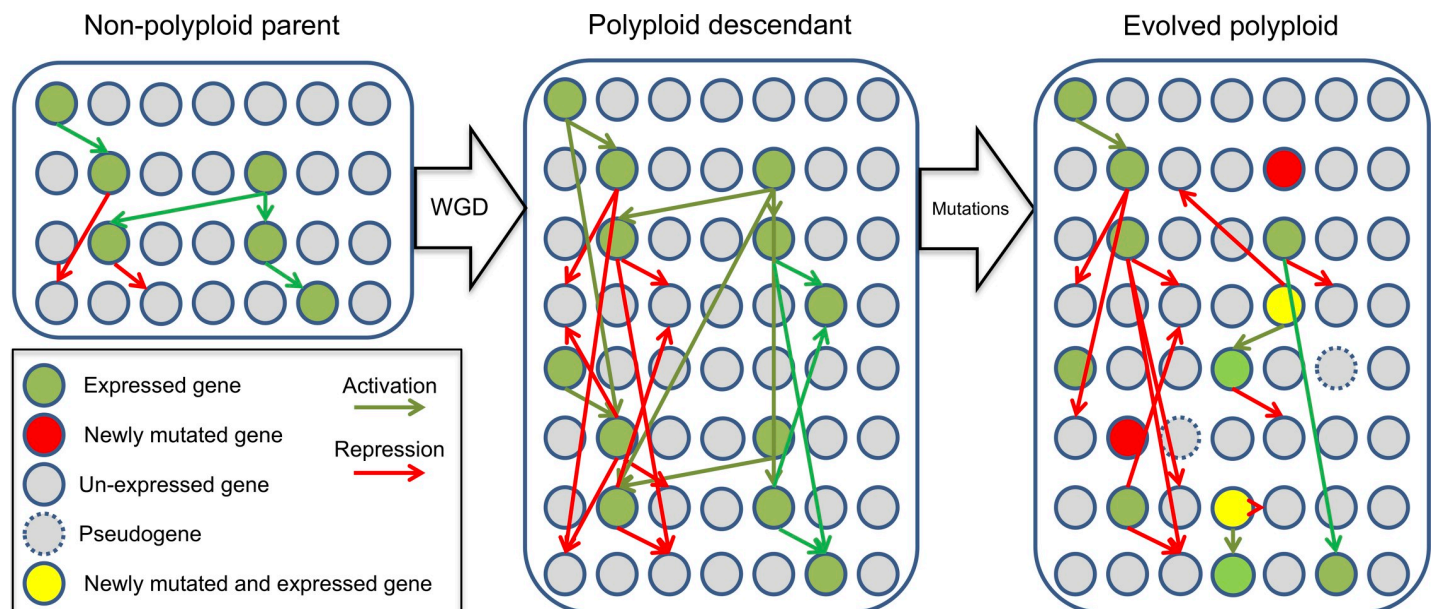


Fig 3. Evolution of a DO under specific WGD and substitution mutation rates. At the beginning of the simulation, the WGD rate is fixed at 40%. Once non-polyploid and polyploid populations sizes have reached equilibrium, the WGD operator is removed from the simulation. Similarly, at every replication step, a fixed substitution rate (10^{-4} per base) is set.

<https://doi.org/10.1371/journal.pone.0220257.g003>

or attack (Fig 1). When the result of the equation corresponds to the identity number of a particular gene, its expression is activated. Linear equations are encoded in the genetic sequence of the genome of a DO (Fig 2D), and can change in number and sequence through WGDs and substitutions, respectively. This way, upon initialization of a new DO, a set of 12 TFs is activated [24, 25] (Fig 1). Subsequently, at every time step during the simulation, the linear equations compute a new set of identity numbers, increasing the concentration of the corresponding gene product by one unit.

Food. Food represents a source of energy. Each food source can store from 10,000 up to 20,000 energy units. The energy of a food source increases 300 units at every time step until it reaches 20,000. At that point, the food source divides into two food sources, with the newly created food source being randomly placed around the old one in a neighboring cell. Food sources can thus, for instance, be compared to growing patches of grass, with new patches sprouting from older ones. When a DO finds a food source, it assimilates its energy level, and the food source is removed. In our framework, environmental challenges were introduced through the random removal of food sources at specific rates and time points.

Running simulations

In our evolutionary framework, populations of clonal DOs evolve under fixed WGD and substitution mutation rates, resulting in the evolutionary diversification of the population. Rather than introducing all polyploids at once in the simulation, we decided to gradually introduce them by setting a fixed WGD rate, which is a situation closer to the real evolutionary process [6, 30]. At the beginning of the simulation, the WGD rate was arbitrarily fixed at 40%, meaning that when a new DO is created (after replication, inheriting the genome and energy level of its 'parent'), there is a 40% chance it will have its genome duplicated. When the population size of polyploids equals that of the non-polyploids, which usually happens in about 100 to 200 time steps, the WGD operator is permanently removed. This high rate of WGD, as compared with what would be expected in nature, was chosen because lower WGD rates result in extremely long computation times until polyploids and non-polyploids reached similarly sized populations. Because our simulations focused less on the initial establishment of polyploids, but more on the potential longer-term consequences of doubling genome content and structure, we decided to start with populations of similar size as the initial condition of the competition. Similarly, at every replication step, a fixed substitution rate (10^{-4} per base) was applied to the offspring genome.

Digital organisms move in a matrix composed of 1000 x 1000 cells, looking for food. Every time a DO moves into a neighboring cell in the matrix, its energy level drops by 70 units. DOs were removed from the population when their energy levels dropped below 0. When the energy level of a particular DO reached a threshold of 20,000, the DO was considered to have reached a level of adaption, and was therefore allowed to replicate (a sign of fitness). However, the final decision of a DO to replicate depends on the relative concentration of the structural gene products involved in promoting and repressing replication. When both concentration levels are equal, the replication has a 50% possibility to happen at that time step.

In order to simulate different evolutionary scenarios, we applied different food removal patterns. In order to mimic environmental turmoil or radical environmental challenges (for instance as caused by an extinction event), we introduced drastic and/or repeated drops of food during our simulations. To investigate the interaction between adaptive potential and WGD, we examined the mutational landscape of the genomes across these various evolutionary and environmental contexts. For every DO and at every time step in the simulation, we computed its ploidy and energy level, as well as i) the short-term genetic distance (STGD),

defined as the average number of substitutions (per 10 kb) between the individual DO and its direct parent; ii) the long-term genetic distance (LTGD), defined as the average number of substitutions per 10 kb between the individual DO and the initial ancestor at time 0 in the simulation; and iii) the expression distance (ED), defined as the average of the concentrations of all gene products in a given DO compared to the average of the concentrations of all gene products in the previous time step. When the ED differs by more than 30% (arbitrarily defined) between two time steps, we considered the corresponding DO as having an unstable expression pattern. In our simulation, a stable expression pattern (considered to reflect homeostasis) is necessary to maintain a certain 'phenotype' and is considered a sign that the DO has adapted to its environment. Therefore, a sequence of low ED values helped us to identify those DOs that have evolved effective strategies for adaptation.

Results

Studying the impact of WGD on the evolution of DOs under different environmental conditions

To study the putative effect of WGDs on the evolution of DOs under challenging environmental conditions, we have simulated the evolution of clonal populations of DOs under two different scenarios of food availability (environmental challenges) and specific rates of WGD and nucleotide substitution. At every time step and for every simulation, the population size of polyploids and non-polyploids was computed. Although, after food removal, a drop in population size was generally observed, when and how and to what extent the population size drops can greatly vary from simulation to simulation [25] (see further). It is inherent to the system that DOs with different genomic makeups and encoded GRNs can have very different responses to the same amount of food removal, which makes it hard to summarize the results, except for evaluating overall success when putting polyploids and non-polyploids in competition. In the first scenario, food was dynamically removed every 60 time steps after the total population (polyploids plus non-polyploids) reached 1000 individuals. Starting from a population of 200 individuals at t_0 , having reached a population size of 1000 indicates that at least part of the population has adapted to the current environment and started to replicate rapidly [24, 25]. Under this first scenario, we performed 30 simulations under each of the following four food reduction levels (thus 120 simulations in total): no food reduction, except for what is found and consumed by the DOs, 50% food reduction, 70% food reduction, and 85% food reduction. After 2000 time steps, or until the subpopulation of polyploids or non-polyploids became extinct, we compared the population sizes of polyploids and non-polyploids (Table 1). As can be observed, under no or low food reduction scenarios, at the end of the experiment most of the population consisted of non-polyploid DOs. The population size of polyploids increased with the percentage of food removed, becoming bigger than that of non-polyploids after 70% of the food was removed repeatedly, i.e. every time the total population reached 1000. Taking away 85% of the food resulted in the extinction of the whole population of both polyploids and non-polyploids in all simulations.

Table 1. Summary of results from evolution simulation experiments under a scenario of dynamic food reduction. For every food reduction group, 30 simulations were run for up to 2000 time steps, or until either the polyploid or non-polyploid subpopulations become extinct.

	No food reduction	50% food reduction	70% food reduction	85% food reduction
Extinction (for all DOs)	0%	10%	16.7%	100%
Polyploid population bigger than non-polyploid population	10%	30%	63.3%	0%
Non-polyploid population bigger than polyploid population	90%	60%	20%	0%

<https://doi.org/10.1371/journal.pone.0220257.t001>

In the second scenario, ‘food removal’ was introduced at a fixed time step, namely at time step 300. Because the growth of food throughout the experiment is exponential, introducing food reduction at time step 300 represented an appropriate balance between earlier reduction before polyploid and non-polyploid populations have reached an equilibrium in the total population, and later reduction, which may not remove enough food for re-adaptation or efficient selection to occur. Under this scenario, 10 independent simulations were run under four different levels of fixed food reduction (0%, 30%, 60% and 90%, thus 40 simulations in total), each with populations starting with 200 clonal DOs each. Simulations were run for up to 1000 time steps, or until the subpopulation of polyploids or non-polyploids became extinct or started growing too big to be computationally tractable (due to exponential growth). At the end of every simulation, we checked whether the population size of polyploids was bigger than that of non-polyploids or the other way around (Table 2). Similarly to what we observed under the previous scenario, at 0 and 30% of food reduction, no polyploid DO survived at the end of the experiment, with some polyploid genomes surviving only under a 60% or more of food reduction. Polyploid populations were able to overtake non-polyploid ones only when significant levels of food (90%) were removed (Table 2). In Fig 4, we show the detailed evolution of population sizes of polyploids and non-polyploids with respect to food energy availability for two independent simulations under 30% or under 90% of fixed food reduction, illustrating the opposite evolutionary trajectories of polyploids and non-polyploids in terms of population sizes under less and more challenging conditions, respectively.

In summary, both simulation scenarios show that, while non-polyploids clearly outcompete polyploids when environmental conditions are stable, when major environmental changes are introduced in the form of drastic reductions of food available, polyploids seem to confer an advantage over non-polyploids. This effect is more noticeable under the first scenario, where the environmental changes are repeated multiple times through recurrent food removal, creating stronger environmental constraints and hence greater needs for multiple re-adaptations (see further).

Evolution of the mutational landscape of DO genomes following WGD

In order to gain further insight on why and how polyploid DO populations have, in most simulations, a greater chance of surviving extreme environmental changes (measured by the number of surviving DOs at the end of the simulation), we examined the evolution of the average long-term genetic distance (LTGD) and short-term genetic distance (STGD) of ‘adapted’ populations by running an additional set of 30 simulations under a 70% dynamic food reduction scenario and computing STGD and LTGD separately for non-polyploids and polyploids. When referring to an ‘adapted population’, we mean that each DO in the population has an energy level of at least 20,000 units at that particular time step. This ensures that we only reconstruct the evolutionary trajectories from DOs that have had a chance for replication in the population. Fig 5 summarizes the evolution of STGD (A) and LTGD (B) across all simulations, revealing non-polyploids showing higher STGD than polyploids during most time steps, while

Table 2. Summary of results from evolution simulation experiments under a scenario of fixed food reduction. For every food reduction group, 10 simulations were run for up to 1000 time steps, or until either the polyploid or non-polyploid subpopulations become extinct, or started growing exponentially.

	No food reduction	30% food reduction	60% food reduction	90% food reduction
Extinction (for all DOs)	0%	0%	0%	20%
Polyploid population bigger than non-polyploid population	0%	0%	20%	60%
Non-polyploid population bigger than polyploid population	100%	100%	80%	20%

<https://doi.org/10.1371/journal.pone.0220257.t002>

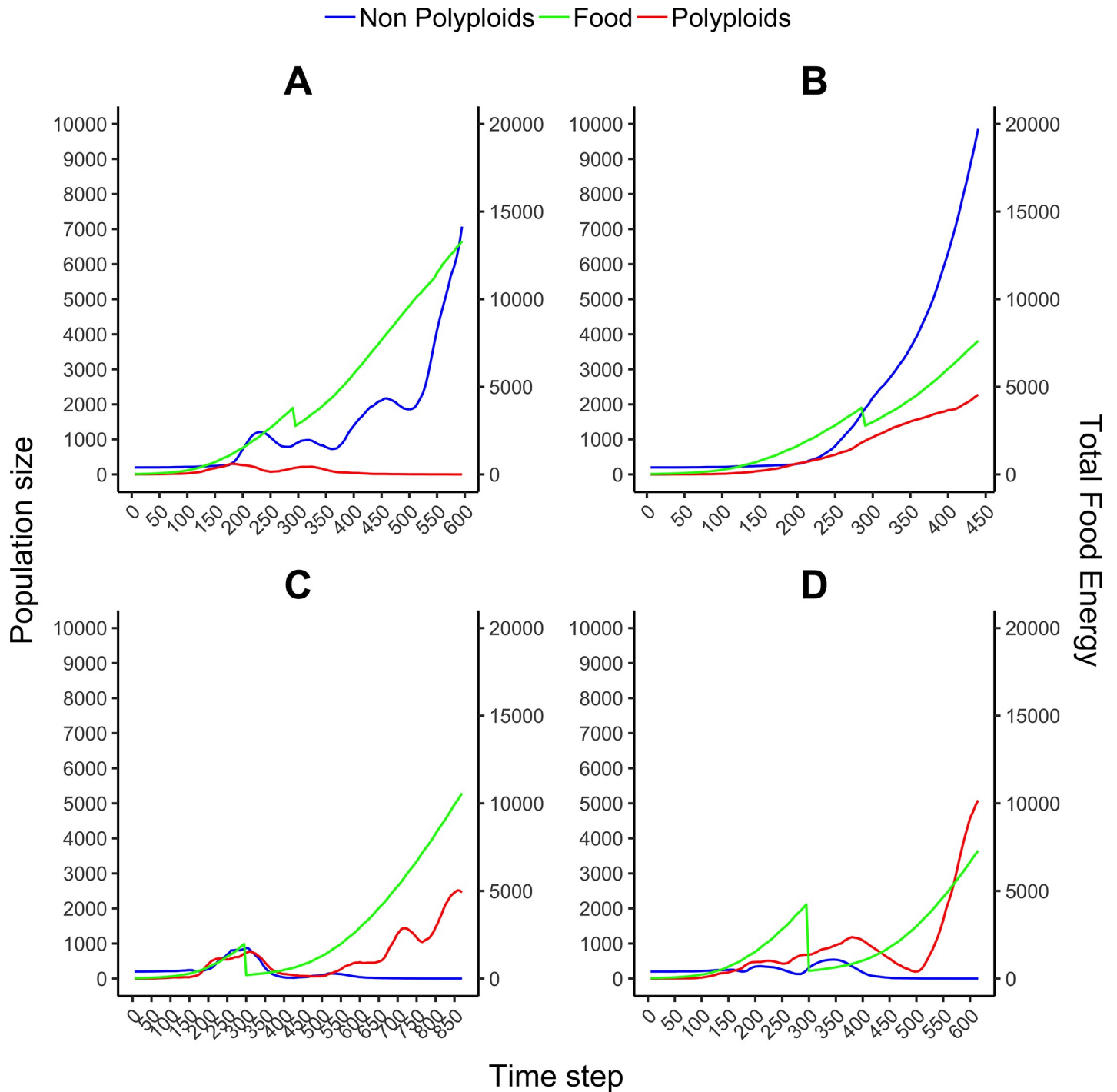


Fig 4. Evolution of population size of polyplods and non-polyplods with respect to food energy availability in four individual experiments under a scenario of fixed food reduction. 30% (A and B) or 90% (C and D), respectively, of the food was removed at time step 300. Simulations start with a population of 200 non-polyplod DOs, which evolve under a WGD rate set at 40%. The WGD rate is removed when the number of polyplods in the population equals the number of non-polyplods (usually at time step 150–200). Simulations were run for up to 1000 time steps, or until either the polyplod or non-polyplod subpopulations become extinct or start growing exponentially (imposing computational limitations).

<https://doi.org/10.1371/journal.pone.0220257.g004>

the opposite pattern was generally observed for LTGD, especially at later time steps. Overall differences in STGD and LTGD between non-polyplods and polyplods were further

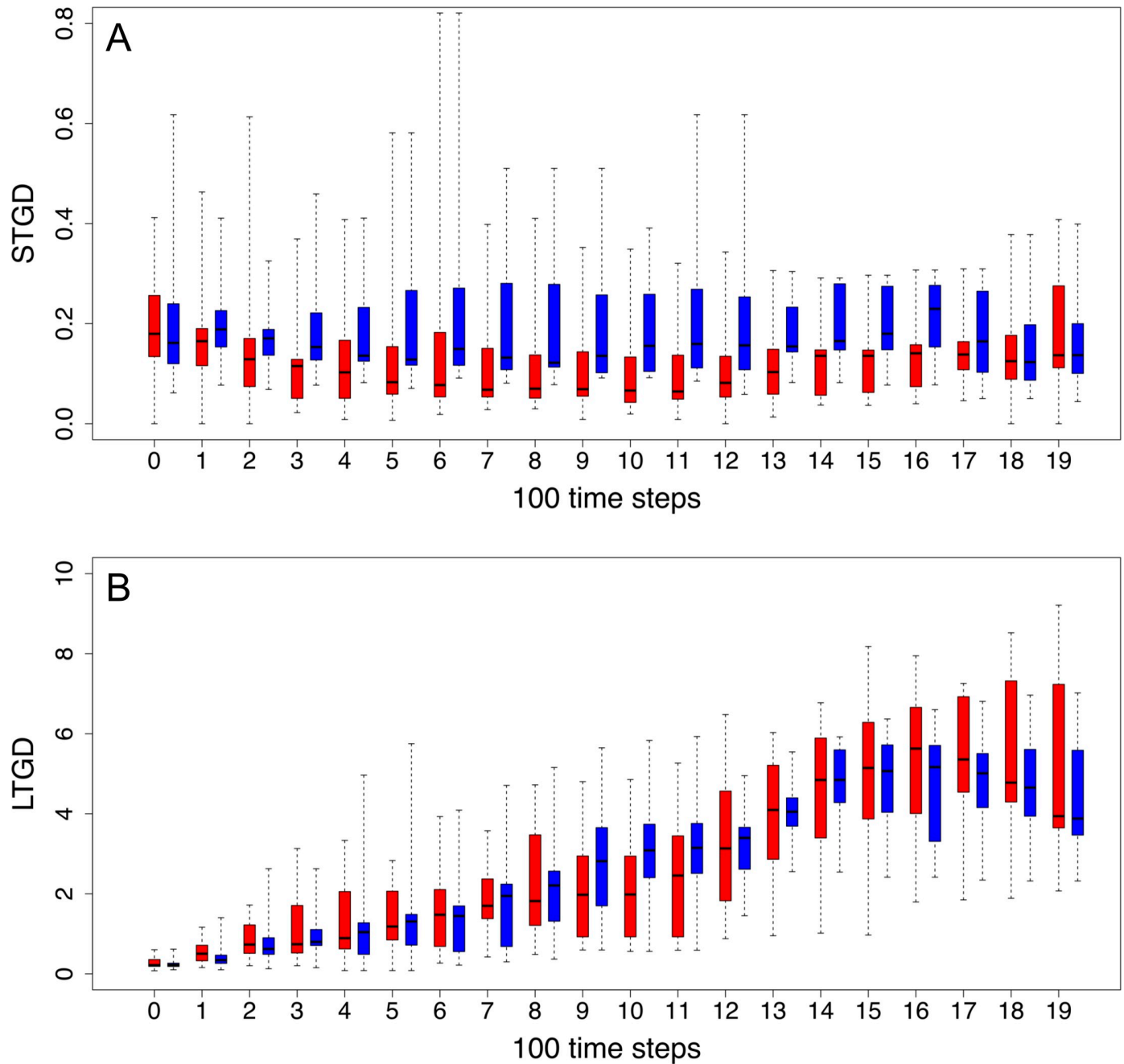


Fig 5. Evolution of short-term genetic distance (A, STGD) and long-term genetic distance (B, LTGD) across 30 experiments (simulations) under a dynamically changing environment (70% food reduction). To represent the evolution of STGD (A) and LTGD (B) across the simulation experiment, average values for all genomes and simulations were computed, ascribed to bins of 100 time steps, and plotted separately for ‘adapted’ polyloid and non-polyloid populations as red and blue box plots, respectively. A two sample *t*-test was based on the same values, comparing polyloid and non-polyloid populations. Solid lines in box blots represent the median values of the STGD and LTGD of all simulations, while the box borders correspond to the first and third quartile and the extreme values correspond to the lowest and highest values observed in any of the simulations. See text for details.

<https://doi.org/10.1371/journal.pone.0220257.g005>

examined by comparing their averages for all genomes and across all time steps and simulations by means of a two sample *t*-test. While non-polyloids showed significantly higher STGD than polyloids ($0.1876360 > 0.1325995$, $p\text{-value} = 2.2 \times 10^{-14}$), the opposite is true for average LTGD values ($2.646769 > 2.215343$, $p\text{-value} = 1.6 \times 10^{-10}$).

It should be noted that the substitution rate for all genomes is the same and hence, under the assumption that WGD has no effect on mutational patterns, the average LTGD and STGD should also be the same for both populations. Consequently, the lower values for the STGD in polyploids have to be the result of selection and thus selection on the short-term seems more stringent for polyploids than for non-polyploids. However, once fixed and inherited, mutations might have more chance to be retained for longer times in polyploid populations, likely because of redundancy in the genome—although this needs further investigation—, making polyploids less vulnerable for single detrimental mutations.

Evolution of GRNs following a WGD

We also investigated the effect of WGDs on the encoded GRNs by measuring the overall extent of changes in the concentration levels of its different components (gene products) upon adaptation. In stable environments, we assume that ‘adapted’ DOs maintain a relatively stable GRN most of the time. In other words, we consider the stability of GRN expression patterns—represented by expression distances (ED) between any DO and its descendants smaller than 30%, see [Methods](#) - as an indicator for adaptation. Abrupt food removal may cause rapid changes in gene expression patterns, ultimately altering the underlying GRN of the DO’s genome. Indeed, differences in food availability might lead to abrupt changes in the ‘sensed’ environment, in turn potentially resulting in domino effects on the concentration levels of particular TFs (see [Fig 1](#)). To study the link between adaptation and changes in the GRN, we ran 40 additional simulations under a scenario of 70% dynamic food reduction and extracted the subpopulations of polyploid and non-polyploid DOs that had likely established a comparatively stable GRN (ED values lower than 30%, reflecting homeostasis) and reached an at least temporary adaptive state (energy levels above 20,000). The evolutionary trajectories in terms of LTGD were then compared for each subpopulation to that of the total population of DOs in each simulation ([Fig 6](#)). Despite the overall higher LTGDs observed among polyploids, we observed that adapted and stable non-polyploids usually need to accumulate considerably more mutations to establish novel GRNs that have adapted to the new environment, while for the polyploid subpopulation, the difference between the adapted subpopulation and the total (including likely non or not-yet adapted DOs) population was less outspoken ([Fig 6](#)). In the current implementation of our framework, we do not know whether this is due to positive selection or selective sweeps in one group of DOs or due to purifying selection in the other, but the fact is that different regimes of selection are ongoing in polyploids and non-polyploids. This contrasting pattern between polyploids and non-polyploids is more obvious at later steps of the simulation ([Fig 6](#)), suggesting that polyploids and non-polyploids have indeed different ways to respond to environmental changes. In any case, based on these observations, we can infer that, for re-adapting under changed or changing environments, non-polyploid genomes need to accumulate more mutations than polyploid ones ([Fig 6](#)).

Discussion

With the purpose of investigating whether WGDs may confer an adaptive advantage in the face of drastic environmental changes or extinction events, such as abrupt reductions in available food resources, we have implemented here an improved version of our previously designed bio-inspired and agent-based computational framework for simulating biological evolution [24, 25]. Two different environmental scenarios of food reduction were set: either food was removed dynamically every time the total population reached 1000 individuals, or at a fixed time step during the simulation. From our simulations, we learned that under stable environments, non-polyploids in general do better than polyploids, as reflected by the low or

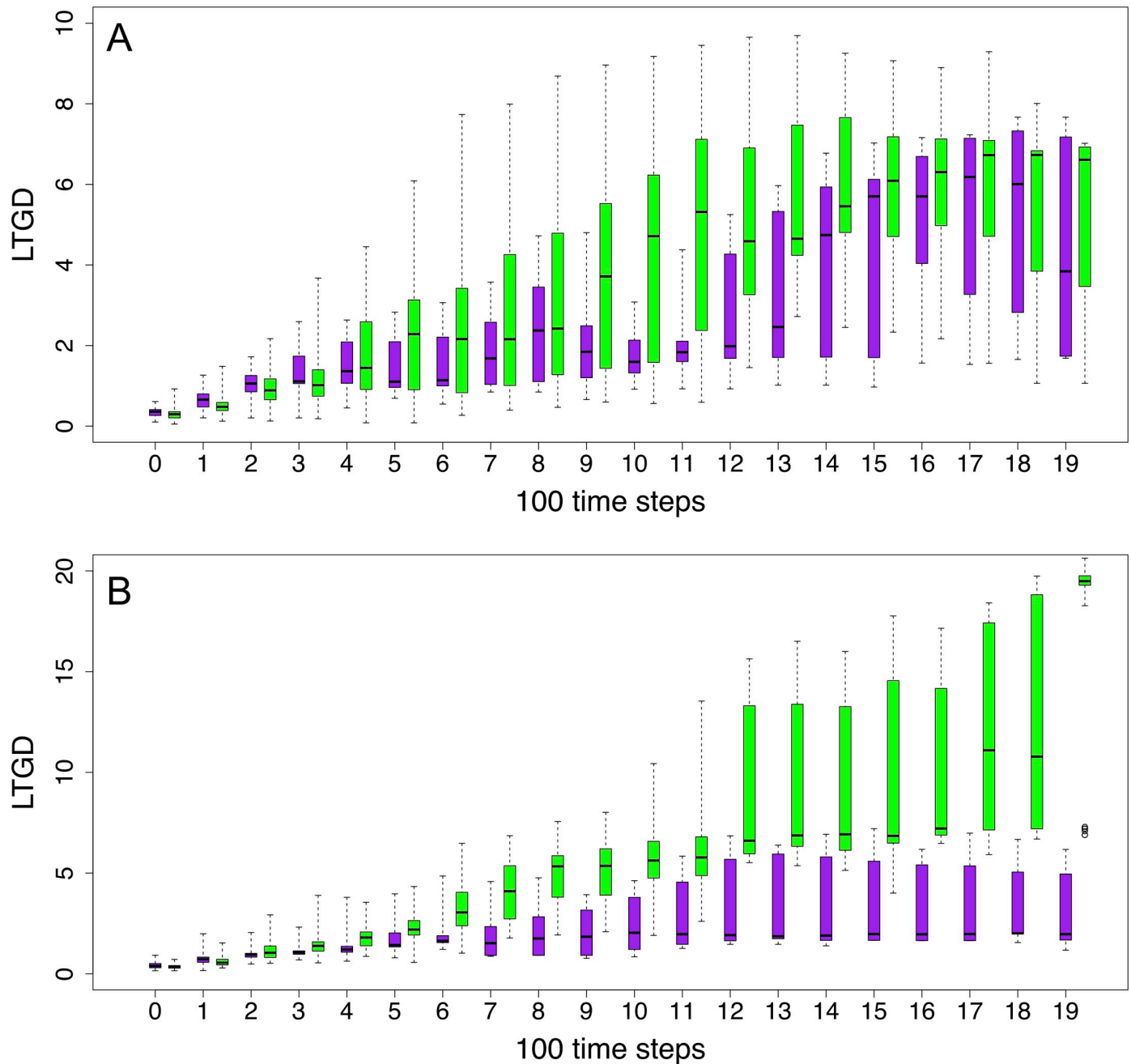


Fig 6. Evolution of the average LTGD of a subpopulation of DOs with stable and adapted GRNs versus all DOs from 40 experiments (simulations) under a dynamically changing environment (70% food reduction). Purple boxes represent the overall population and green boxes represent the population with stable GRNs (i.e. considered adapted). (A) polyloid populations; (B) non-polyloid populations. Interpretation is as in Fig 5.

<https://doi.org/10.1371/journal.pone.0220257.g006>

null fraction of polyloids surviving at the end of the different experiments performed under low or null food reduction. Under both scenarios, the relative fraction of polyloid DOs surviving at the end of the experiment increased with increasing levels of food reduction, overcoming the population of non-polyloid DOs when the environmental challenge introduced was strong enough. Taken as whole, these results suggest that WGDs are usually maladaptive under stable environments, but may actually confer an adaptive advantage under certain

environmental constraints, which seems in agreement with recent ‘real-life’ observations [6, 31–33].

Although the substitution rate was the same for all genomes and is not supposed to change after WGD, contrary to our initial expectations, we observed lower STGD among polyploid DOs, suggesting that duplicated genomes are under stronger purifying selection, at least on the short-term. Once fixed and inherited though, mutations might have more chance to be retained for longer times in duplicated genomes, as suggested by the higher LTGD values for polyploids, possibly because of redundancy in the genome, although this needs to be confirmed and further investigated.

We have also shown that DOs with non-duplicated genomes need to accumulate many more mutations than polyploid genomes to reach adaptation (using homeostasis as measured by ED and the energy level as proxies). We interpret this as follows: under stable environments, and in ‘well-adapted’ organisms with (relatively) stable GRNs, there is always a certain chance that a random mutation renders the GRN (and consequently, its host) less well adapted to that environment. This phenomenon is exacerbated in more complex GRNs. With an increasing number of edges and connections in the network, the probability increases that a random mutation in a regulator affects a larger number of genes, in turn increasing the chance of causing major disturbances to the underlying GRN [34–37]. In this respect, WGDs increase the number of edges in the network exponentially (Fig 7). For instance, in our simulations, while the average GRN for the total population consisted of 272 instances or agents at the start, this increased to about 1,700 for DOs with a duplicated genome at the end of the simulations.

From previous theoretical and experimental work, it has been shown that gene duplication and divergence may be at the origin of specific properties of biological networks such as GRNs, including increased complexity and, also, modularity, i.e. the occurrence of multiple functional subnetworks or modules formed by highly connected genes that are co-expressed and/or co-regulated by the same set of key regulators, while establishing sparser connections with nodes from different modules [1, 38–46]. As a consequence, single substitutions in duplicated genomes have the potential to affect many more key regulators, in turn having a greater impact on the underlying GRN and the resulting phenotypes. Again, whereas this may be disadvantageous in stable environments, under drastic environmental changes, DOs with duplicated genomes may actually have a better chance to acquire the drastic genomic—and consequently phenotypic—changes necessary to survive major changes in the fitness landscape (Fig 8). In addition, the duplication of genomes and their encoded GRNs may result in redundancy, contributing to the genetic or mutational robustness at the GRN level. This facilitates the rewiring of novel functional modules while maintaining the ancestral one, a feature that is expected to be especially advantageous under unstable, recurring, and/or challenging environments [23, 41]. In this respect, our observations are compatible with previous claims that increased complexity and modularity results in increased evolvability [36, 47–52]. For a population that is already well adapted to a certain environment, the increased complexity and modularity of the duplicated GRNs could be maladaptive or detrimental because the GRNs are more vulnerable to random mutations and the potential for increased evolvability may have little value or direct benefits. On the other hand, the complex and modular structure of duplicated GRNs might allow polyploids to explore a wider evolutionary landscape (Fig 8), providing short-term increased opportunity to adapt to novel, different, or rapidly changing environments, while redundancy facilitates the gradual evolution of duplicated networks for adapting in the longer term. In partial summary, increased complexity, modularity and functional redundancy following WGD would help to explain the different behavior of polyploids and non-polyploids under stable or challenging environments [6].

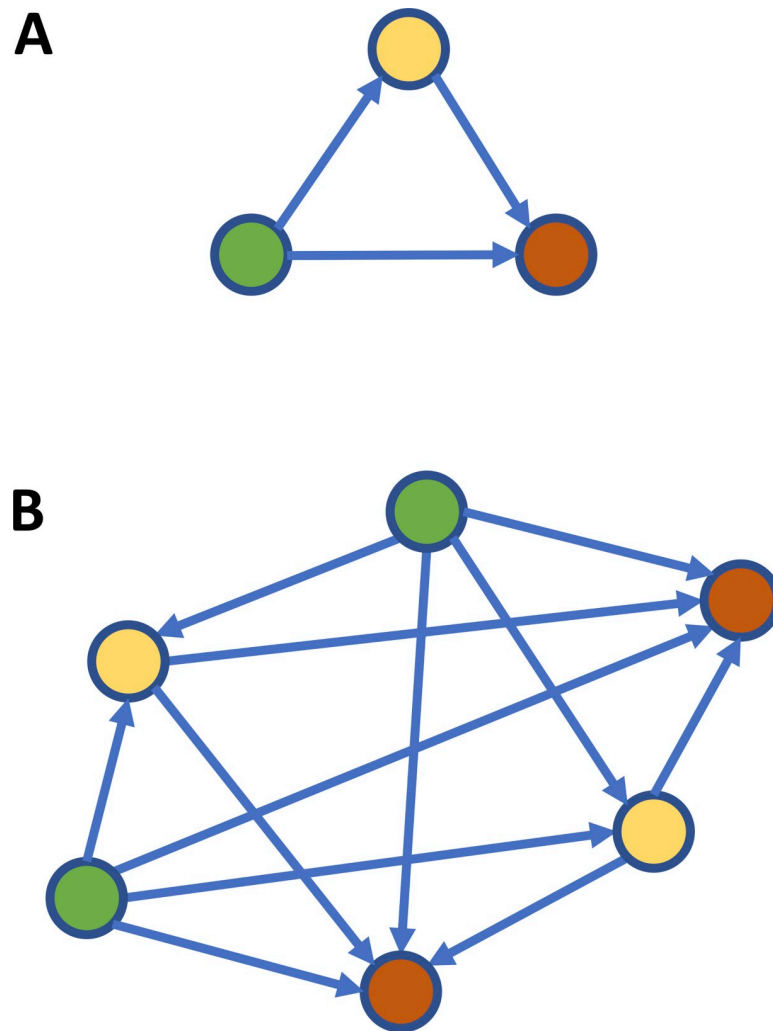


Fig 7. Simple network motif (feed forward loop) before (A) and after (B) genome wide duplication. After duplication, network complexity has increased and random mutations may affect (many) more genes, in turn increasing the chance of causing major disturbances to the GRN and underlying phenotype.

<https://doi.org/10.1371/journal.pone.0220257.g007>

Although many other factors may be at play as well, which, for simplicity, are not considered here (but will be addressed elsewhere), the effect of WGD on network complexity and redundancy could be part of the explanation why also in natural environments, floras in stable environments might have lower proportions of polyploids than highly, recently disturbed floras. In a recent study, Oberlander et al. [53] suggested that the hyper diverse South African Cape flora, which has anomalously low proportions of polyploids compared to global levels, might be due to long-term climatic and geological stability, supporting the hypothesis that WGD may be rare in stable environments. On the contrary, there is considerable evidence that polyploidy is much more common in disturbed habitats, or habitats that (have) show(n) environmental turmoil [6, 31, 54]. Previously, we have also wondered about the discrepancy between the existence of many polyploids of fairly recent origin and the scant evidence of ancient polyploidy events, certainly within the same evolutionary lineage [6]. The paucity of polyploidy events that 'survive' and are established in the long term would suggest that polyploidy is usually an evolutionary dead end. Again, this fits with our observation that the

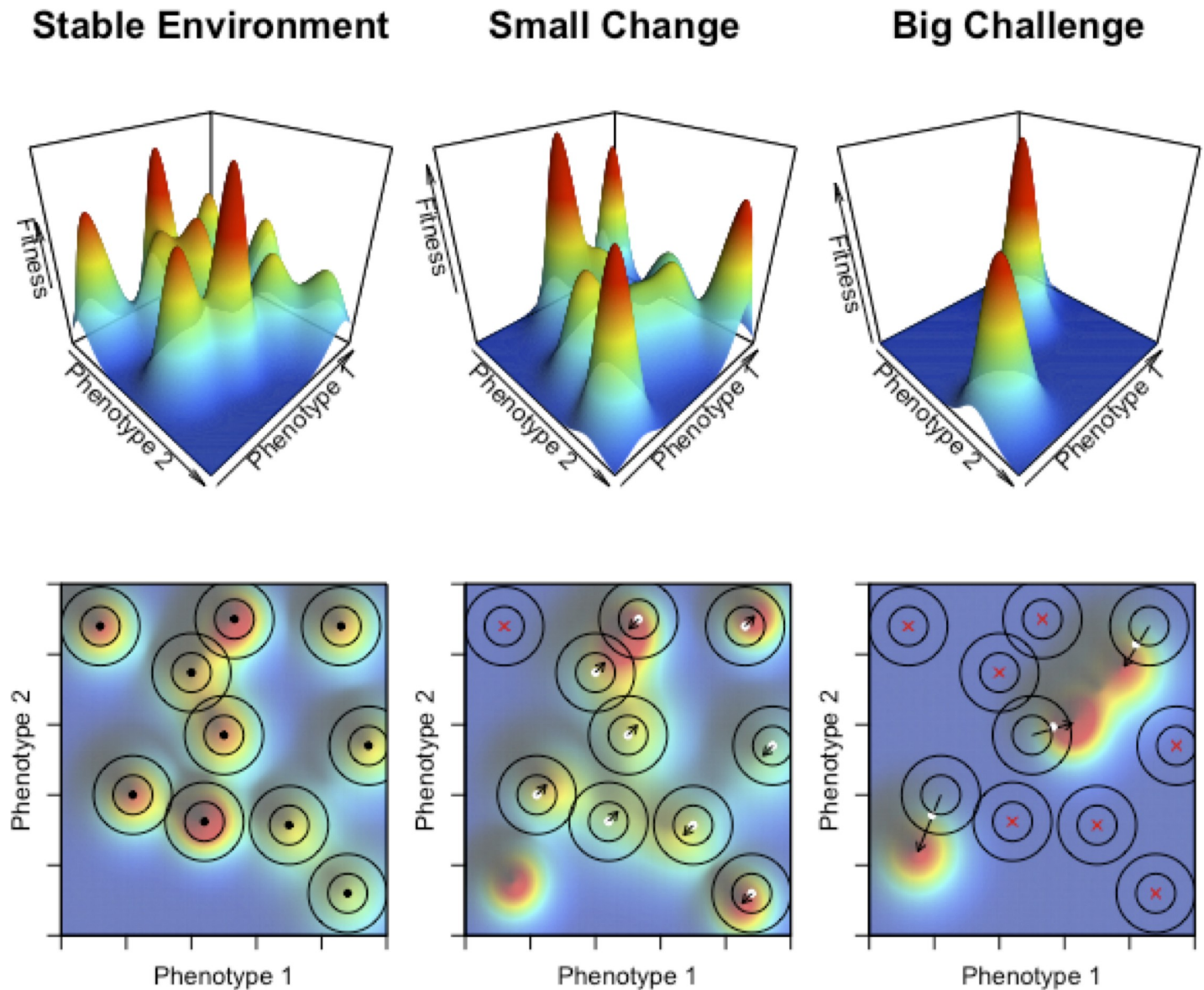


Fig 8. The role of polyploidy in exploring the fitness landscape under different environmental scenarios. Top panels, 3D surface representation of the phenotype space accessible for every combination of two hypothetical phenotypic traits, which ultimately correspond to an organism's specific genotypic configuration. Each coordinate in the XY plane conformed by phenotype 1 and phenotype 2 has a relative fitness value associated on the Z axis. The resulting fitness landscape is formed by a range of hills, each corresponding to the total phenotype space accessible to a specific organism or group of organisms, i.e. their ecological niche, with a local maximum or adaptive peak located at the top. Hills are surrounded by valleys or depressions, colored in dark blue, corresponding to regions of the phenotype space inaccessible to any organism. Bottom panels show the 2D colored image plots corresponding to the fitness landscapes above. Black dots, white dots and red crosses represent organisms occupying the fitness landscape. Concentric circles around organisms represent the areas of the phenotype space accessible by non-polyplod organisms (inner circle) or their polyplod relatives (outer circle). Arrows indicate the trajectories followed by every organism to reach their adaptive peak. Black dots represent organisms that have reached their local maxima and are thus located at their adaptive peaks; white dots represent organisms that survived after a change in the environmental conditions and subsequent shifts of their adaptive peaks; red crosses represent organisms that perish because of big shifts or extinction of their niches. From left to right, three different environmental scenarios are shown. In the left two panels, organisms evolving under a stable environment for a certain time are expected to have reached their local adaptive peaks. In the middle two panels, a small change in environmental conditions (e.g., a small drop in the total food available) may result in small shifts of the peaks or their relative fitness contribution, or eventually in the extinction or emergence of specific ecological niches. In most cases, the new location of the peaks fall within the phenotype area (still) accessible to non-polyplod organisms (inner circles). Under this scenario, non-polyplod organisms are expected to quickly reach the new peaks, rapidly outcompeting their polyplod relatives, which are more likely to fall farther away (outer circles) and have well-known detrimental effects. Finally, in the right panels, a cataclysmic or extinction event (e.g., a big drop in the total food available) is represented, resulting in a reduction in the number of available ecological niches, and big shifts in their relative locations. Under these conditions, although most organisms are expected to perish, polyplod organisms, featured by wider accessible phenotype space (as a result of the potential higher impact of mutations in more complex genomes, see text for details), have better chances to fall near the peak of a newly formed adaptive hill and to develop the necessary evolutionary innovations to colonize empty niches.

<https://doi.org/10.1371/journal.pone.0220257.g008>

increased complexity and modularity of the underlying networks may be detrimental (or at least sub-optimal) for species that are well-adapted to their environment, when accumulating (too) many changes (Fig 8). In this respect, it is also noteworthy to mention that during our simulations, rarely two or three WGDs were observed, and never more. Also in real-life organisms, very few examples are known of consecutively established WGDs that have occurred in a short period of time. Some of these presumed exceptions are *Musa acuminata*, *Spirodela polyrhiza*, and *Arabidopsis thaliana* [6]. However, at least for *Arabidopsis* [55] and *Spirodela* [56], the number of genes is low compared to what would be expected following several rounds of WGD, suggesting huge gene loss and fractionation following WGD. In line with our observations made in the current study, quickly losing a large amount of the extra genetic material created through a WGD, thereby potentially reducing the complexity of the active GRNs, might have facilitated surviving these major events.

Finally, our observations might also have some broader implications and might help to understand the origin of complex phenotypes and discontinuities in biological evolution. Traditionally, evolution is regarded as progressing at a more or less constant rate, which is linked to the gradual accumulation of small genetic changes over time. This model successfully explains, for instance, the variation between (closely) related species but often fails to explain bigger 'leaps' in evolution. However, there is ample recording of 'novel' or highly divergent species that seem to have emerged in a relatively short evolutionary time frame and sometimes even representing important evolutionary transitions. Such observations often seem difficult to reconcile with an evolutionary model that is based on the gradual accumulation of small changes, and a huge body of literature has been devoted to this seemingly contradictory phenomenon [57, 58].

As supported by our simulations, and in previous research [24, 25], in a stable environment, gradual evolution can successfully explain the divergence and optimization of evolutionary processes. In contrast, when the environment drastically changes, in a short geological time frame, gradual evolution has difficulty to keep up, while new species or life forms, if different enough from the original ones, can successfully occupy the 'new' niches that have become available. When the environment is rapidly changing, existing species may not have enough time to adapt and will disappear. WGD might provide a way out and might be one way reconciling a model of gradual evolution with adaptation to a rapidly changing environment [3]. Gene and genome duplication has been suggested before as a way to explain 'saltational' jumps in evolution [59–61]. Indeed, the duplication of genes and particularly the duplication of entire genomes immediately creates redundant informational 'entities' or 'modules' in the genome, offering possibilities for a more drastic change and a wider exploration of genotypic and phenotypic space (Fig 8). While in a constant environment, in which organisms are already fairly well-adapted, we only could observe the further slow optimization of biological systems, a WGD creates more drastic changes in certain individual species. These individuals often remain hidden in the population—and usually have a hard time competing with their non-polyploid progenitors, for reasons explained higher and elsewhere [6, 7, 31], but could 'grab their chance' upon different conditions or under different contexts. Increased gene pleiotropy evoked by WGD thus forms a fertile substrate allowing evolution to explore more diverse options possibly explaining those bigger jumps in evolution observed. Again, such evolvability might only be successful when the existing environmental conditions have been (seriously) disrupted (Fig 8). When a new environmental 'equilibrium' has been reached and the new environmental conditions have stabilized again, the increased evolvability through WGD will become less important in further optimizing the system and on the contrary selection on a more complex system (due to WGD, see higher) might be constrained again. Considering polyploidy in the evolutionary history of organisms might thus indeed be one additional way

of possibly explaining bigger leaps or major transitions in evolution, as already suggested by Susumu Ohno in 1970 [62], but certainly needs further investigation.

Conclusions

Our observations seem to support the two-bladed sword role claimed for WGD in evolution, constituting on the one hand an evolutionary dead end in stable environments, but on the other hand offering opportunities to avoid or reduce the risk of extinction in unstable and/or drastically changing environments [1, 6]. Furthermore, we hope that this work shows the usefulness of our DO-based evolution simulation computational framework in studying the role of WGD in evolution and adaptation, by tracking the evolutionary trajectories of individual artificial genomes in terms of sequence and expression diversification, helping to overcome some of the traditional limitations of evolution experiments with model organisms.

Acknowledgments

Y.V.d.P. acknowledges funding from the European Union Seventh Framework Programme (FP7/2007-2013) under European Research Council Advanced Grant Agreement 322739 – DOUBLEUP. The authors would also like to thank Arthur Zwaenepoel and Eshchar Mizrahi for helpful discussions, Steven Maere for help with Fig 8, and two anonymous reviewers for constructive comments that greatly helped improving the paper. The funder had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The authors have declared that no competing interests exist.

Author Contributions

Conceptualization: Yao Yao, Lorenzo Carretero-Paulet, Yves Van de Peer.

Data curation: Yao Yao, Lorenzo Carretero-Paulet.

Formal analysis: Yao Yao, Lorenzo Carretero-Paulet.

Funding acquisition: Yves Van de Peer.

Writing – original draft: Yao Yao, Lorenzo Carretero-Paulet.

Writing – review & editing: Yves Van de Peer.

References

1. Crow KD, Wagner GP. What is the role of genome duplication in the evolution of complexity and diversity? *Molecular Biology and Evolution*. 2006; 23(5):887–92. <https://doi.org/10.1093/molbev/msj083> WOS:000237320800007. PMID: 16368775
2. Semon M, Wolfe KH. Consequences of genome duplication. *Curr Opin Genet Dev*. 2007; 17(6):505–12. <https://doi.org/10.1016/j.gde.2007.09.007> PMID: 18006297.
3. Van de Peer Y, Maere S, Meyer A. The evolutionary significance of ancient genome duplications. *Nat Rev Genet*. 2009; 10(10):725–32. <https://doi.org/10.1038/nrg2600> PMID: 19652647.
4. Arrigo N, Barker MS. Rarely successful polyploids and their legacy in plant genomes. *Curr Opin Plant Biol*. 2012; 15(2):140–6. <https://doi.org/10.1016/j.pbi.2012.03.010> PMID: 22480430.
5. Soltis PS, Marchant DB, Van de Peer Y, Soltis DE. Polyploidy and genome evolution in plants. *Curr Opin Genet Dev*. 2015; 35:119–25. <https://doi.org/10.1016/j.gde.2015.11.003> PMID: 26656231.
6. Van de Peer Y, Mizrahi E, Marchal K. The evolutionary significance of polyploidy. *Nat Rev Genet*. 2017. <https://doi.org/10.1038/nrg.2017.26> PMID: 28502977.
7. Comai L. The advantages and disadvantages of being polyploid. *Nat Rev Genet*. 2005; 6(11):836–46. <https://doi.org/10.1038/nrg1711> PMID: 16304599.

8. Jaillon O, Aury JM, Brunet F, Petit JL, Stange-Thomann N, Mauceli E, et al. Genome duplication in the teleost fish *Tetraodon nigroviridis* reveals the early vertebrate proto-karyotype. *Nature*. 2004; 431(7011):946–57. <https://doi.org/10.1038/nature03025> PMID: 15496914.
9. Dehal P, Boore JL. Two rounds of whole genome duplication in the ancestral vertebrate. *PLoS Biol*. 2005; 3(10):e314. <https://doi.org/10.1371/journal.pbio.0030314> PMID: 16128622; PubMed Central PMCID: PMC1197285.
10. Meyer A, Van de Peer Y. From 2R to 3R: evidence for a fish-specific genome duplication (FSGD). *Bioessays*. 2005; 27(9):937–45. <https://doi.org/10.1002/bies.20293> PMID: 16108068.
11. Aury JM, Jaillon O, Duret L, Noel B, Jubin C, Porcel BM, et al. Global trends of whole-genome duplications revealed by the ciliate *Paramecium tetraurelia*. *Nature*. 2006; 444(7116):171–8. <https://doi.org/10.1038/nature05230> WOS:000241867200033. PMID: 17086204
12. Albertin W, Marullo P. Polyploidy in fungi: evolution after whole-genome duplication. *Proc Biol Sci*. 2012; 279(1738):2497–509. <https://doi.org/10.1098/rspb.2012.0434> PMID: 22492065; PubMed Central PMCID: PMC3350714.
13. Kellis M, Birren BW, Lander ES. Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*. *Nature*. 2004; 428(6983):617–24. <https://doi.org/10.1038/nature02424> PMID: 15004568.
14. Jiao Y, Wickett NJ, Ayyampalayam S, Chanderbali AS, Landherr L, Ralph PE, et al. Ancestral polyploidy in seed plants and angiosperms. *Nature*. 2011; 473(7345):97–100. <https://doi.org/10.1038/nature09916> PMID: 21478875.
15. Wendel JF. The wondrous cycles of polyploidy in plants. *Am J Bot*. 2015; 102(11):1753–6. <https://doi.org/10.3732/ajb.1500320> PMID: 26451037.
16. Van de Peer Y, Fawcett JA, Proost S, Sterck L, Vandepoele K. The flowering world: a tale of duplications. *Trends Plant Sci*. 2009; 14(12):680–8. <https://doi.org/10.1016/j.tplants.2009.09.001> PMID: 19818673.
17. Sessa E. Polyploidy as a mechanism for surviving global change. *New Phytol*. 2019; 221:5–6. <https://doi.org/10.1111/nph.15513> PMID: 30488604
18. Vanneste K, Baele G, Maere S, Van de Peer Y. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. *Genome Res*. 2014; 24(8):1334–47. <https://doi.org/10.1101/gr.168997.113> PMID: 24835588; PubMed Central PMCID: PMC4120086.
19. Vanneste K, Maere S, Van de Peer Y. Tangled up in two: a burst of genome duplications at the end of the Cretaceous and the consequences for plant evolution. *Philos Trans R Soc Lond B Biol Sci*. 2014; 369(1648). <https://doi.org/10.1098/rstb.2013.0353> PMID: 24958926; PubMed Central PMCID: PMC4071526.
20. Lohaus R, Van de Peer Y. Of dups and dinos: evolution at the K/Pg boundary. *Curr Opin Plant Biol*. 2016; 30:62–9. <https://doi.org/10.1016/j.pbi.2016.01.006> PMID: 26894611.
21. Cai L, Xi Z, Amorim AM, Sugumaran M, Rest JS, Liu L, et al. Widespread ancient whole-genome duplications in Malpighiales coincide with Eocene global climatic upheaval. *New Phytol*. 2019; 221(1):565–76. Epub 2018/07/22. <https://doi.org/10.1111/nph.15357> PMID: 30030969.
22. Selmecki AM, Maruvka YE, Richmond PA, Guillet M, Shores N, Sorenson AL, et al. Polyploidy can drive rapid adaptation in yeast. *Nature*. 2015; 519(7543):349–52. <https://doi.org/10.1038/nature14187> PMID: 25731168; PubMed Central PMCID: PMC4497379.
23. Cuyppers TD, Hogeweg P. A synergism between adaptive effects and evolvability drives whole genome duplication to fixation. *PLoS Comput Biol*. 2014; 10(4):e1003547. <https://doi.org/10.1371/journal.pcbi.1003547> PMID: 24743268; PubMed Central PMCID: PMC3990473.
24. Yao Y, Marchal K, Van de Peer Y. Improving the adaptability of simulated evolutionary swarm robots in dynamically changing environments. *PLoS One*. 2014; 9(3):e90695. <https://doi.org/10.1371/journal.pone.0090695> PMID: 24599485; PubMed Central PMCID: PMC3944896.
25. Yao Y, Storme V, Marchal K, Van de Peer Y. Emergent adaptive behaviour of GRN-controlled simulated robots in a changing environment. *PeerJ*. 2016; 4:e2812. <https://doi.org/10.7717/peerj.2812> PMID: 28028477; PubMed Central PMCID: PMC5180581.
26. Yao Y, Van de Peer Y, editors. Simulating biological complexity through artificial evolution. *Cybernetics (CYBCONF)*, 2017 3rd IEEE International Conference on; 2017: IEEE.
27. Yao Y, Marchal K, Van de Peer Y. Adaptive self-organizing organisms using a bio-inspired gene regulatory network controller: for the aggregation of evolutionary robots under a changing environment. Hershey PA, USA: ISI Global; 2016 2016. 68–82 p.
28. Reil T. Dynamics of gene expression in an artificial genome—Implications for biological and artificial ontogeny. *Advances in Artificial Life, Proceedings*. 1999; 1674:457–66. WOS:000170202200063.

29. Grozinger CM, Richards J, Mattila HR. From molecules to societies: mechanisms regulating swarming behavior in honey bees (*Apis* spp.). *Apidologie*. 2014; 45(3):327–46. WOS:000334412100004.
30. Meyers LA, Levin DA. On the abundance of polyploids in flowering plants. *Evolution*. 2006; 60(6):1198–206. PMID: [16892970](#).
31. te Beest M, Le Roux JJ, Richardson DM, Brysting AK, Suda J, Kubesova M, et al. The more the better? The role of polyploidy in facilitating plant invasions. *Ann Bot*. 2012; 109(1):19–45. <https://doi.org/10.1093/aob/mcr277> PMID: [22040744](#); PubMed Central PMCID: PMC3241594.
32. Diallo AM, Nielsen LR, Kjaer ED, Petersen KK, Raebild A. Polyploidy can Confer Superiority to West African *Acacia senegal* (L.) Willd. Trees. *Frontiers in Plant Science*. 2016; 7. ARTN 821 10.3389/fpls.2016.00821. WOS:000377658100001.
33. Godfree RC, Marshall DJ, Young AG, Miller CH, Mathews S. Empirical evidence of fixed and homeostatic patterns of polyploid advantage in a keystone grass exposed to drought and heat stress. *R Soc Open Sci*. 2017; 4(11):170934. <https://doi.org/10.1098/rsos.170934> PMID: [29291088](#)
34. Moore MJ. The Cost of Complexity. *Proceedings of the Institution of Mechanical Engineers Part a-Journal of Power and Energy*. 1995; 209(4):251–. WOS:A1995TF22900001.
35. Orr HA. Adaptation and the cost of complexity. *Evolution*. 2000; 54(1):13–20. WOS:000086315700002. PMID: [10937178](#)
36. Welch JJ, Waxman D. Modularity and the cost of complexity. *Evolution*. 2003; 57(8):1723–34. <https://doi.org/10.1554/02-673> WOS:000185073500002. PMID: [14503615](#)
37. Gorochowski TE, Grierson CS, di Bernardo M. Organization of feed-forward loop motifs reveals architectural principles in natural and engineered networks. *Sci Adv*. 2018; 4(3):eaap9751. <https://doi.org/10.1126/sciadv.aap9751> PMID: [29670941](#); PubMed Central PMCID: PMC5903899.
38. West GB, Brown JH, Enquist BJ. The fourth dimension of life: fractal geometry and allometric scaling of organisms. *Science*. 1999; 284(5420):1677–9. <https://doi.org/10.1126/science.284.5420.1677> PMID: [10356399](#).
39. Freeling M, Thomas BC. Gene-balanced duplications, like tetraploidy, provide predictable drive to increase morphological complexity. *Genome Research*. 2006; 16(7):805–14. WOS:000238712400001. <https://doi.org/10.1101/gr.3681406> PMID: [16818725](#)
40. Rowland MA, Abdelzaher A, Ghosh P, Mayo ML. Crosstalk and the Dynamical Modularity of Feed-Forward Loops in Transcriptional Regulatory Networks. *Biophys J*. 2017; 112(8):1539–50. <https://doi.org/10.1016/j.bpj.2017.02.044> PMID: [28445746](#); PubMed Central PMCID: PMC5406374.
41. Espinosa-Soto C. On the role of sparseness in the evolution of modularity in gene regulatory networks. *PLoS Comput Biol*. 2018; 14(5):e1006172. <https://doi.org/10.1371/journal.pcbi.1006172> PMID: [29775459](#); PubMed Central PMCID: PMC5979046.
42. Fusco D, Grassi L, Bassetti B, Caselle M, Cosentino Lagomarsino M. Ordered structure of the transcription network inherited from the yeast whole-genome duplication. *BMC Syst Biol*. 2010; 4:77. <https://doi.org/10.1186/1752-0509-4-77> PMID: [20525287](#)
43. Carbone I, Ramirez-Prado JH, Jakobek JL, Horn BW. Gene duplication, modularity and adaptation in the evolution of the aflatoxin gene cluster. *BMC Evol Biol*. 2007; 7:111. <https://doi.org/10.1186/1471-2148-7-111> PMID: [17620135](#)
44. Hallinan J. Gene duplication and hierarchical modularity in intracellular interaction networks. *Biosystems*. 2004; 74(1–3):51–62. WOS:000221560400005. <https://doi.org/10.1016/j.biosystems.2004.02.004> PMID: [15125992](#)
45. Hintze A, Adami C. Evolution of complex modular biological networks. *Plos Computational Biology*. 2008; 4(2). ARTN e23 <https://doi.org/10.1371/journal.pcbi.0040023> WOS:000255407800029. PMID: [18266463](#)
46. Altenberg L. *Genome growth and the evolution of the genotype-phenotype map*. Berlin: Springer-Verlag; 1995.
47. Wagner GP, Altenberg L. Perspective: Complex adaptations and the evolution of evolvability. *Evolution*. 1996; 50(3):967–76. WOS:A1996UY55500001. <https://doi.org/10.1111/j.1558-5646.1996.tb02339.x> PMID: [28565291](#)
48. Wagner GP, Pavlicev M, Cheverud JM. The road to modularity. *Nature Reviews Genetics*. 2007; 8(12):921–31. <https://doi.org/10.1038/nrg2267> WOS:000251551100011. PMID: [18007649](#)
49. Crombach A, Hogeweg P. Evolution of Evolvability in Gene Regulatory Networks. *Plos Computational Biology*. 2008; 4(7). ARTN e1000112 <https://doi.org/10.1371/journal.pcbi.1000112> WOS:000260039300021. PMID: [18617989](#)
50. Ten Tusscher KH, Hogeweg P. Evolution of networks for body plan patterning; interplay of modularity, robustness and evolvability. *PLoS Comput Biol*. 2011; 7(10):e1002208. <https://doi.org/10.1371/journal.pcbi.1002208> PMID: [21998573](#); PubMed Central PMCID: PMC3188509.

51. Calabretta R, Nolfi S, Parisi D, Wagner GP. A case study of the evolution of modularity: Towards a bridge between evolutionary biology, artificial life, neuro- and cognitive science. *Artificial Life Vi*. 1998;275–84. WOS:000075924900031.
52. Bolker JA. Modularity in development and why it matters to evo-devo. *American Zoologist*. 2000; 40(5):770–6. [https://doi.org/10.1668/0003-1569\(2000\)040\[0770:Midawij2.0.Co;2](https://doi.org/10.1668/0003-1569(2000)040[0770:Midawij2.0.Co;2) WOS:000166493300007.
53. Oberlander KC, Dreyer LL, Goldblatt P, Suda J, Linder HP. Species-rich and polyploid-poor: Insights into the evolutionary role of whole-genome duplication from the Cape flora biodiversity hotspot. *Am J Bot*. 2016; 103(7):1336–47. <https://doi.org/10.3732/ajb.1500474> PMID: 27352831.
54. Pandit MK, Pockock MJO, Kunin WE. Ploidy influences rarity and invasiveness in plants. *J Ecol*. 2011; 99:1108–15.
55. Arabidopsis Genome I. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature*. 2000; 408(6814):796–815. <https://doi.org/10.1038/35048692> PMID: 11130711.
56. Wang W, Haberer G, Gundlach H, Glasser C, Nussbaumer T, Luo MC, et al. The *Spirodela polyrhiza* genome reveals insights into its neotenuous reduction fast growth and aquatic lifestyle. *Nat Commun*. 2014; 5:3311. <https://doi.org/10.1038/ncomms4311> PMID: 24548928; PubMed Central PMCID: PMC3948053.
57. Theißen G. Saltational evolution: hopeful monsters are here to stay. *Theory Biosci*. 2009; 128(1):43–51. /. <https://doi.org/10.1007/s12064-009-0058-z> PMID: 19224263
58. Theißen G. The proper place of hopeful monsters in evolutionary biology. *Theory Biosci*. 2006; 124(3–4):349–69. <https://doi.org/10.1016/j.thbio.2005.11.002> PMID: 17046365
59. Soltis PS, Liu X, Marchant DB, Visger CJ, Soltis DE. Polyploidy and novelty: Gottlieb's legacy. *Philos Trans R Soc Lond B Biol Sci*. 2014; 369(1648). <https://doi.org/10.1098/rstb.2013.0351> PMID: 24958924; PubMed Central PMCID: PMC4071524.
60. Soltis PS, Soltis DE. Ancient WGD events as drivers of key innovations in angiosperms. *Curr Opin Plant Biol*. 2016; 30:159–65. <https://doi.org/10.1016/j.pbi.2016.03.015> PMID: 27064530.
61. Soltis DE, Visger CJ, Soltis PS. The polyploidy revolution then . . . and now: Stebbins revisited. *Am J Bot*. 2014; 101(7):1057–78. <https://doi.org/10.3732/ajb.1400178> PMID: 25049267.
62. Ohno S. *Evolution by gene duplication*: Springer-Verlag; 1970.