Report

Current Biology

Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*

Highlights

- Cells evolving relatively less genome for their size have higher fitness
- Adaptive evolution tends to minimize genome size relative to cell size
- An absolute lower limit for the evolution of smaller genome exists in this species
- We found a delayed response of genome size to the evolution of cell size

Authors

Martino E. Malerba, Giulia Ghedini, Dustin J. Marshall

Correspondence

dustin.marshall@monash.edu (D.J.M.), giulia.ghedini@monash.edu (G.G.), martino.malerba@gmail.com (M.E.M.)

In Brief

By using evolved lineages of the same eukaryotic species, Malerba et al. provide evidence for fitness benefits associated with reducing relative genome size and DNA content within a cell—as predicted by the selfish DNA hypothesis. We also see indications of countervailing forces imposing a minimum genome size in this species.

Malerba et al., 2020, Current Biology *30*, 1–7
September 7, 2020
https://doi.org/10.1016/j.cub.2020.06.033

CellPress

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

Current Biology

CellPress

Report

Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*

Martino E. Malerba,[1,2,*] Giulia Ghedini,[1,*] and Dustin J. Marshall[1,*]
[1]Centre of Geometric Biology, School of Biological Sciences, Monash University, Melbourne, VIC 3800, Australia
[2]Lead Contact
*Correspondence: martino.malerba@gmail.com (M.E.M.), giulia.ghedini@monash.edu (G.G.), dustin.marshall@monash.edu (D.J.M.)
https://doi.org/10.1016/j.cub.2020.06.033

SUMMARY

Genome size is tightly coupled to morphology, ecology, and evolution among species [1–5], with one of the best-known patterns being the relationship between cell size and genome size [6, 7]. Classic theories, such as the "selfish DNA hypothesis," posit that accumulating redundant DNA has fitness costs but that larger cells can tolerate larger genomes, leading to a positive relationship between cell size and genome size [8, 9]. Yet the evidence for fitness costs associated with relatively larger genomes remains circumstantial. Here, we estimated the relationships between genome size, cell size, energy fluxes, and fitness across 72 independent lineages in a eukaryotic phytoplankton. Lineages with relatively smaller genomes had higher fitness, in terms of both maximum growth rate and total biovolume reached at carrying capacity, but paradoxically, they also had lower energy fluxes than lineages with relative larger genomes. We then explored the evolutionary trajectories of absolute genome size over 100 generations and across a 10-fold change in cell size. Despite consistent directional selection across all lineages, genome size decreased by 11% in lineages with absolutely larger genomes but showed little evolution in lineages with absolutely smaller genomes, implying a lower absolute limit in genome size. Our results suggest that the positive relationship between cell size and genome size in nature may be the product of conflicting evolutionary pressures, on the one hand, to minimize redundant DNA and maximize performance—as theory predicts—but also to maintain a minimum level of essential function.

RESULTS AND DISCUSSION

The nuclear genome spans 200,000-fold size range across the Tree of Life, and this variation has fascinated scientists for decades [10]. There is no correlation between genome size and animal complexity [11, 12], but genome size matters. Species with larger genomes tend to have shorter dispersal [13], lower metabolic rates [1–3], and slower rates of development and growth [4, 5, 14–16]. The evolutionary forces that shape genome size remain remarkably poorly understood and are strongly contested [10, 12].

Cell size covaries with genome size. Indeed, most of the variance in genome size can be explained by cell size, with larger cells having larger genomes [6, 7]. The "selfish DNA hypothesis" (similar to the "junk DNA hypothesis") offers an explanation for this size dependency: it posits that natural selection at the genome level favors self-replicating genes with no phenotypic expression, which can act as a burden on the fitness of a cell [8, 9]. Hence, the DNA keeps expanding uacc7fntil the fitness costs on the host cell are no longer tolerated and thus countered by selection at the cell level. The proposed explanation is that producing excessive DNA results in an energetic burden to a cell, and larger cells with greater total energy budgets should better tolerate large DNA contents. However, there is limited empirical support for the maladaptive nature of genome expansion among eukaryotes. In fact, the only evidence comes from two meta-analyses showing that, within a family, species with relatively larger genomes are usually at greater risk of extinction [17, 18].

Although comparing cell size and genome size among species is an essential start, these approaches only offer indirect support. For example, most species with larger genomes also have slower rates of metabolism, development, and growth [1, 15]. So is the relationship between genome size and fitness causative or is it just an evolutionary by-product of correlated differences in the pace of life? Ultimately, if increasing genome size impacts fitness—as predicted by the selfish DNA hypothesis—individuals with more DNA relative to their cell size should have lower fitness. Yet within-species tests remain extremely rare [19].

We used 72 independent lineages of the eukaryotic microalga *Dunaliella tertiolecta* to test the relationships among genome size, cell size, energy use, and fitness. First, we tested the selfish DNA hypothesis by comparing genome size—while controlling for cell size—to estimate the effect on fitness (see Fitness and Relative Genome Size) and energy use (see Energy Rates and Relative Genome Size). Second, we used our artificial selection experiment to compare the evolutionary dynamics of absolute genome size across 100 generations (ca. 7 months) among lineages that differed dramatically in both cell size and genome size (see Evolutionary Trajectories of Absolute Genome Size). In this way, we could explore the limits of genome size evolution within

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

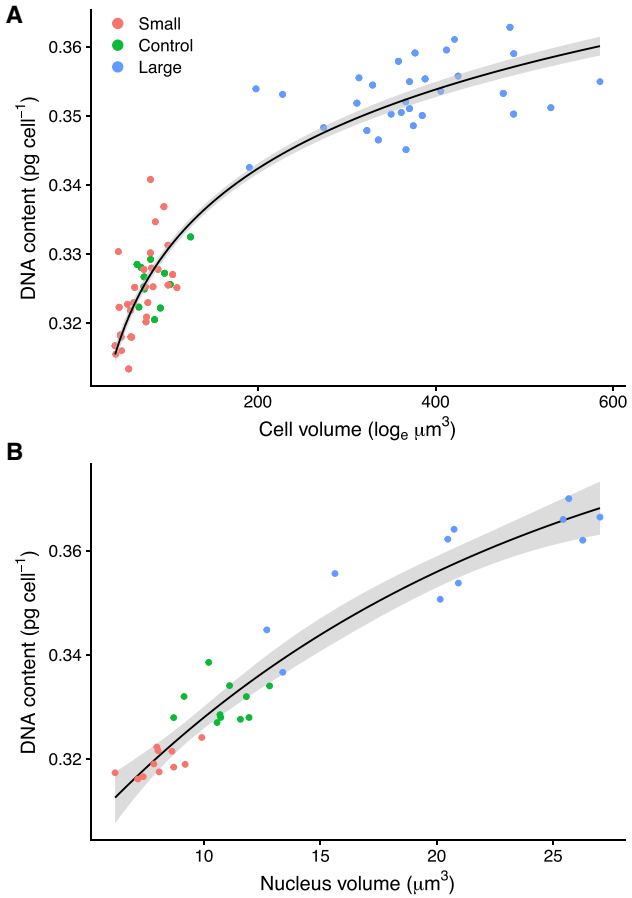

Figure 1. Positive Relationships among Cell DNA Content, Mean Cell Volume, and Mean Nucleus Volume.

Lines (±95% confidence interval [CI]) indicate best-fitting ordinary least squares (OLS) models following AIC (A: *DNA content* = 0.25 $e^{0.016\ Cell\ Volume}$, R^2 = 0.9; B: *DNA content* = 0.39 − 0.11 $e^{-0.055\ Nucleus\ Volume}$, R^2 = 0.92). Each dot represents the mean value measured from an independent lineage, with the color indicating the artificial size-selection treatment. Mean cell DNA content was measured using flow cytometry, whereas cell size and nucleus size were measured with fluorescent microscopy. See Figure S1 for cell size trajectories across generations.

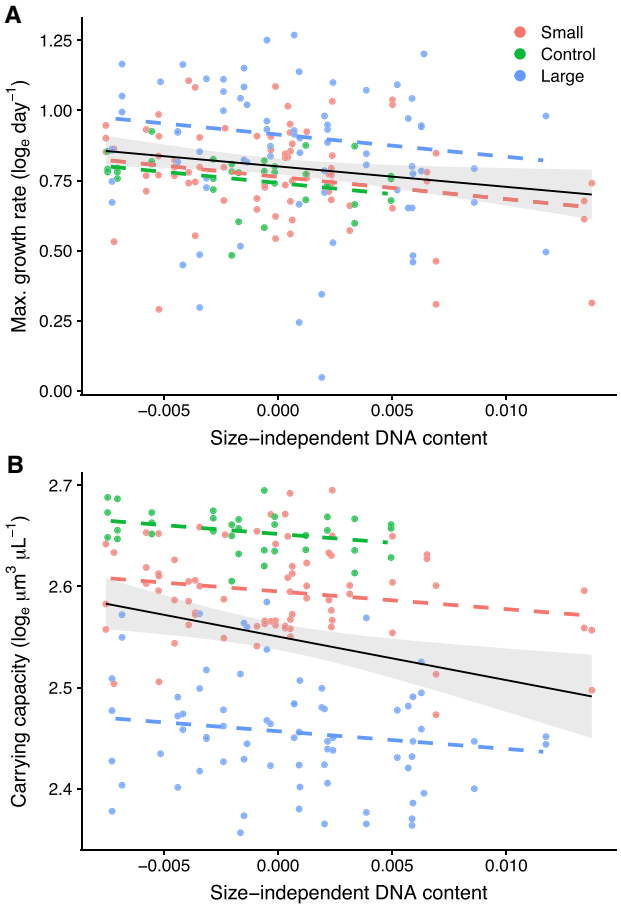

Figure 2. Growth Rate and Carrying Capacity Decrease With Increasing Size-Independent DNA Content

Size-independent DNA content had statistically significant negative effects on both population-level traits (type II Wald chi-square test; maximum intrinsic growth rate: df = 1, chi-square = 6.31, p = 0.012; carrying capacity: df = 1, chi-square = 3.94, p = 0.047) and consistent across size-selection treatments (i.e., no sign. interaction among dashed lines in both panels). Values on x axis are the mean size-independent DNA content of a cell, calculated as the residuals from the model in Figure 1A. Colors indicate the size-selection treatment. Dashed lines indicate the best-fitting linear mixed-effect models at each size-selection treatment, whereas continuous lines indicate the overall effects across all data (± 95% CI). Each dot represents the mean value measured from an independently selected lineage—after correcting for fixed (i.e., plate and initial biovolume density) and random (lineage identity) covariates. See ANOVA in Table S1, Figure S1 for cell size trajectories across generations, and Figure S2 for independent confirmatory experiments.

a species across a much larger range of cell sizes than would normally be possible.

Evolution of Cell Size, Nucleus Size, and DNA Content

Trials for this experiment took place between 350 and 500 generations of artificial selection, when cell sizes showed little directional change over time—albeit with occasional fluctuations (see Figure S1 for evolutionary trajectories in cell size across 3 years). Mean cell volumes were on average 97 μm^3 for small-selected lineages, 177 μm^3 for control, and 915 μm^3 for large-selected lineages (see arrows in Figure S1). Hence, large-selected cells were 9.4 times larger than small-selected cells and 5.2 times larger than control cells. Cell size was highly positively correlated with nucleus volume (R^2 = 0.9; Figure 1A). Nucleus volumes ranged from 5 to 25 μm^3 and were highly, positively correlated with cell DNA content (R^2 = 0.92; Figure 1B), which ranged from 0.30 to 0.37 pg among lineages (Figure 1). Such strong positive correlations among cell size, nucleus size, and total DNA content are often found in photosynthetic cells [20], whereas in other microbes (e.g., yeast), total DNA content is unrelated to nucleus size [21, 22].

Fitness and Relative Genome Size

The size-independent DNA content (pg) quantifies the genome size relative to the cell volume of a lineage (i.e., model residuals in Figure 1A). We found that both the maximum growth rate and the total biovolume reached at carrying capacity of a lineage decreased with increasing size-independent DNA content

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

CellPress

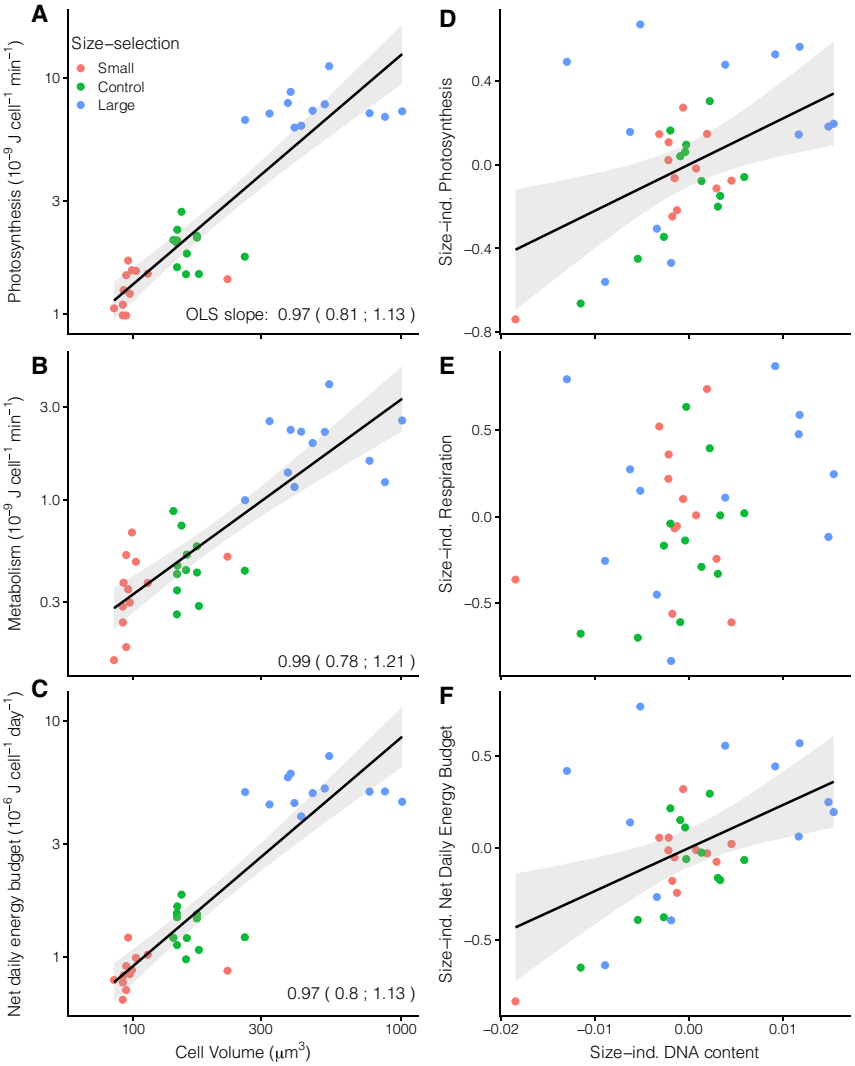

Figure 3. Effects of Size-Independent (Relative) DNA Content on Total and Size-Independent Energy Rates of a Cell

(A–C) The left column shows the relationships between mean cell volume (μm^3) and per-cell (A) photosynthesis ($F_{1,34}$ = 145.6; p < 0.001; R^2 = 0.81), (B) respiration ($F_{1,34}$ = 86.28; p < 0.001; R^2 = 0.72), and (C) net energy ($F_{1,34}$ = 139.1; p < 0.001; R^2 = 0.8).

(D–F) The right column shows the effect of size-independent DNA content on size-independent (D) photosynthesis ($F_{1,34}$ = 9.41; p = 0.004), (E) respiration ($F_{1,34}$ = 2.68; p = 0.11), and (F) net daily energy budget ($F_{1,34}$ = 10.35; p = 0.003).

The size-independent energy rate of a culture was calculated as the residuals from the models in the left column. The size-independent DNA of a culture was calculated as the residuals from the model in Figure 1A. Each dot represents the mean value calculated from an independent lineage, with the color representing the artificial size-selection treatment. Lines indicate the best-fitting OLS linear model, with size-scaling slopes reported for the left column (±95% CI). See Figure S1 for cell size trajectories across generations.

(Figure 2). This indicates that cells with relatively smaller genomes grew faster and accumulated more biomass than cells with relatively larger genomes, regardless of cell size. Also, the best-fitting models following AIC did not include any 2-way interactions between size-independent DNA content and the artificial size-selection treatment, indicating that the effects of DNA content on both fitness proxies were consistent among control, small-selected, and large-selected treatments (i.e., similar slopes among dashed lines in Figure 2; see Table S1 for ANOVA table). We repeated our experiment with a higher initial per-cell resource regime and again found a negative relationship between fitness and relative genome size (see Figure S2).

This finding offers some of the first direct, empirical support for the fitness benefits of reducing genome size relative to cell size, as predicted by the selfish DNA hypothesis [8, 9]. The fitness benefits of relatively smaller genomes are consistent with empirical findings of prokaryotic cells showing increased growth after removing non-essential DNA [23]. Although these results indicate that "genome streamlining" is often beneficial, the mechanisms remain controversial. It could be that newly formed repeated sequences are initially maladaptive but might later become functionally integrated within the genome [23]. If so, the genome size of a species may result from a balance between the costs of genome expansion and the advantages of a greater potential cooperation among sequences [23]. Because nucleotides are particularly rich in nitrogen (N) and phosphorous (P) [24], the disadvantages of relatively larger genomes should be accentuated when these essential nutrients are scarce. This hypothesis predicts that selection against larger relative genomes should be strong under resource limitation. Our finding that relatively smaller genomes can reach higher biovolumes at carrying capacity (i.e., the size of the population when growth stops due to resource limitation) is consistent with this prediction.

Energy Rates and Relative Genome Size

Larger cells recorded higher per-cell energy rates (Figures 3A–3C), which is consistent with previous studies on this model system [25–27]. The size-independent energy rate indicates energy use relative to cell size (i.e., model residuals in Figures 3A–3C). Theory predicts that the increase in fitness with decreasing size-independent DNA content should arise from decreasing metabolic costs associated with replicating superfluous DNA [8, 9]. Instead, we found the opposite: both the size-independent photosynthesis and the net daily energy budget increased with the size-independent DNA content of a lineage, whereas size-independent respiration showed no relationship with relative genome size (Figures 3D–3F). These relationships were consistent among control, small-selected, and large-selected lineages (i.e., no significant interaction in Figures 3D–3F).

CURBIO 16591

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

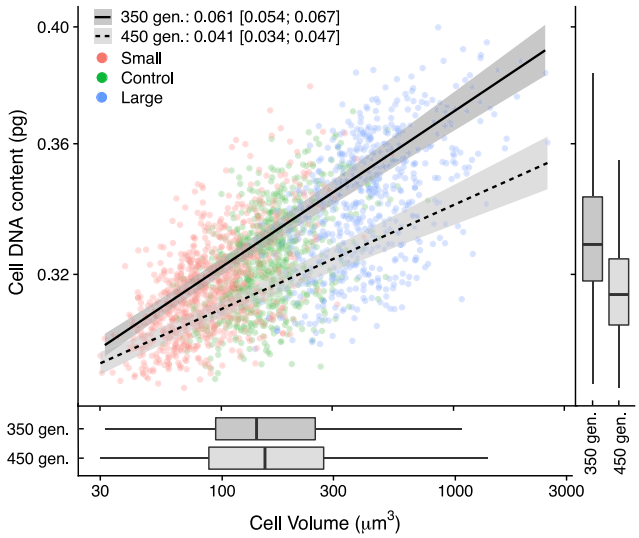

Figure 4. Relationships between Cell Volume and Cell DNA Content in Lineages of *Dunaliella tertiolecta* after 350 (Solid Line) and 450 (Dashed Line) Generations of Artificial Selection for Size

Lines (±95% CI) indicate the best-fitting linear mixed-effect model on $\log_{10}$-$\log_{10}$ axes, with slope coefficients reported in the legend. The model revealed a significant interaction between cell size and generation time (type II Wald chi-square test, $\chi_1 = 29.3$, $p < 0.001$; slope coefficients and CI reported in legend). Each point is a cell after correcting for random covariates (i.e., lineage identity nested within generation), and the color represents the size-selection treatment. The DNA content (pg) of each cell was inferred from measurements of nucleus volume (μm^3) in Figure S4 by using the calibration curve in Figure 1B ($R^2 = 0.92$)—hence, these trends are unaffected by organelle DNA. Boxplots show the distributions of cell size (bottom) and cell DNA content (right) after 350 (n = 1,332) and 450 (n = 1,058) generations. See Methods S1 for step-by-step guide and sample images, Figure S1 for cell size trajectories across generations, and Figure S3 for the relationship between cell volume and cell DNA content divided by generation.

These results show that lineages with relative larger genomes had more net energy available. Energy use (respiration) was unaffected by relative genome size, but photosynthesis rates increased with relative genome size, leading to net energy availability increasing with relative genome size. So lineages with relatively larger genomes grow more slowly when resources are abundant and have lower biovolumes at carrying capacity (Figure 2) but have more energy available than lineages with relatively smaller genomes (Figures 3D and 3F). We are at a loss to explain these apparently contradictory results. It seems there are complex interactions between genome size, energy use, and performance that current theory does not capture. Other studies should explore the covariance between these factors to establish their generality so that we can modify theory regarding genome size evolution accordingly.

Evolutionary Trajectories of Absolute Genome Size

The cell volume of size-selected cells showed little directional change after 250 generations of artificial selection (Figure S1). After cell size had stabilized under each selection treatment, cells reduced their relative DNA content from 350 to 450 generations (see boxplots on right of Figure 4), but the magnitude of this reduction depended on cell size (i.e., significant interaction cell size × generation time in Figure 4; see Figure S3 for model fits within each artificial selection treatment). Specifically, the DNA content of a large cell (2,500 μm^3) decreased by 0.012 standard deviations per generation from 350 to 450 generations ($t_{70} = -6.22$; $p < 0.001$), whereas the DNA content of a small cell (30 μm^3) showed no difference over time ($t_{70} = -1.85$; $p = 0.068$; Figure 4). Importantly, this reduction in DNA content occurred without cells altering their cell volumes ($t_{2180} = -0.3$; $p = 0.76$; see overlapping boxplots at the bottom of Figure 4). The evolutionary trajectories of cell DNA content were inferred from changes in nucleus volume (shown in Figure S4) using the calibration curve in Figure 1B—hence, our inferences were unaffected by differences in organelle DNA among cells (see Methods S1).

These results show that the capacity to reduce the relative genome size of a cell increased with cell size and absolute genome size. Despite comparable directional selection (i.e., parallel lines in Figure 2), larger cells reduced their DNA content by up to 11% in 100 generations (+0.012 standard deviations per generation), whereas smaller cells (with absolutely smaller genomes) showed no change over time. There appears to be a minimum absolute size to the genome of this species, which constrains the scope for smaller cells to reduce their genome. Possibly, the physical dimensions of the nucleus cannot be reduced below a certain size, which is consistent with small-selected cells having relatively larger nuclei [28] (see size-scaling exponents <1 in Figure S4). Alternatively, evolutionary rates may be proportional to absolute genome size, with larger genomes expected to evolve faster than smaller genomes [29]. If so, reducing the genome below a certain size (through a correlated reduction in cell size) slowed down evolutionary rates, meaning that small-selected cells require many more generations to achieve an equivalent evolution than large-selected cells. Among species, an increase in evolutionary rates with increasing genome sizes and cell size might be offset by generation time also decreasing with cell size [30] (but see [31]). Yet this is not the case for our lineages, where instead, larger cells tend to grow faster than smaller cells (see blue line above red line in Figure 2A; see [32] for a discussion on fitness differences across cell sizes for this model system). Regardless of the underlying mechanisms, that the adaptive capacity of a cell increases with its absolute genome size suggests a potential advantage of evolving larger cells with larger genomes, but this hypothesis remains untested.

The evolution of smaller genomes continued well after cells had achieved a stable size distribution. Specifically, all cells retained their size after the first 250 generations, whereas genome size continued to evolve until 450 generations (cf. boxplots on the right with those at the bottom of Figure 4). We think that the longer evolutionary response of genome size might be a consequence of the strong artificial selection pressure applied on cell size. Genotypic covariance with cell size may have initially "dragged" genome size to unfavorable phenotypes, due to pleiotropy or correlational selection. Only later, after cells had achieved a stable size (~250 generations), the genome could adapt to smaller, more favorable sizes. Consistent with this explanation, the size-independent DNA content was more variable in small- and large-selected lineages (ranging from −0.5 to 1) than in control lineages (from −0.5 to 0.4; cf. x axis range among treatments in Figures 2A and 2B). Traits can differ in their

CURBIO 16591

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

ability to respond to correlational selection on body size [33–37]. Yet delayed evolution is usually inferred from phylogenetic comparative methods, which are based on pattern recognition and have a limited ability to infer cause-effect relationships [37]. Instead, here, we showed that manipulating cell size using artificial selection can lead to a temporally decoupled response of genome size lasting for several hundred generations. An important next step will be to determine how general these evolutionary lags in genome size are under natural conditions and among other species.

Why do we observe persistent variation in relative genome size in the face of directional selection on all of our measured fitness proxies? Assuming no evolutionary constraints, consistent selection for relative smaller genomes should reduce among-individual variation [38, 39], but relatively larger genome sizes were present after 350 generations and variation among lineages persisted. Initially, we hypothesized that each observed combination of cell size and genome size may be under selection in a specific environment. For example, relative larger genomes may enhance competitive abilities at high light and high nutrients, whereas relatively smaller genomes might have higher fitness as resources decrease. This explanation was consistent with our finding that larger genomes show greater size-independent energy fluxes and poorer performance in high-competition, low-resource environments (as measured by biovolume reached at carrying capacity). But this explanation is inconsistent with lineages with relatively larger genomes also showing slower growth under resource-unlimited conditions (see Figure 2A). We confirmed this finding in a second, independent experiment, in which we grew lineages under even higher initial per-capita resources (see Figure S2). Thus, lineages with relatively smaller genomes are not only more ''thrifty'' when it comes to resources, they are also able to grow faster when resources are abundant—such results contradict classic ecological theory [40–42].

Conclusions

By using evolved lineages of the same eukaryotic species, we provide direct evidence for fitness benefits associated with reducing relative genome size—as predicted by theory—and we also see indications of countervailing forces imposing a minimum genome size in a species, which are yet to be incorporated into theory. Overall, our results suggest that the positive relationship between cell size and genome size in nature may be the product of contrasting evolutionary pressures, not only to minimize redundant DNA but also to maintain a minimum level of essential function. Energy fluxes are clearly affected by genome size, but not in the ways that theory anticipates, as lineages with smaller genomes appear able to avoid the classic trade-off between population growth rate and carrying capacity [42].

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
  - Lead Contact
  - Materials Availability
  - Data and Code Availability
- EXPERIMENTAL MODEL AND SUBJECT DETAILS
- METHOD DETAILS
  - Artificial selection for cell size
  - Experimental trials
  - Cell size
  - DNA content
  - Nucleus size
  - Energy use
  - Population growth
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental Information can be found online at https://doi.org/10.1016/j.cub.2020.06.033.

A video abstract is available at https://doi.org/10.1016/j.cub.2020.06.033#mmc4.

ACKNOWLEDGMENTS

We thank Melanie Loveass, Tormey Reimer, Lucy Chapman, Belinda Comerford, and Michaela Parascandalo for help with laboratory procedures. We also thank Dr. Michael McDonald, Dr. Chris Greening, the handling editor, and two anonymous reviewers for insightful comments and suggestions. We would like to express our gratitude to Lesley Wiadrowski, Stewart Crowley, and John Arvanitakis for logistical support at Monash University. Finally, we are particularly grateful to the Australian Research Council for financial support.

AUTHOR CONTRIBUTIONS

All authors contributed to conceptualization of the study. M.E.M. and G.G. developed the methodology and collected the data. M.E.M. carried out formal analysis, developed all figures and tables, and wrote the initial draft of the manuscript. All authors provided substantial comments and edits and gave final approval for publication.

DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: March 8, 2020
Revised: April 28, 2020
Accepted: June 9, 2020
Published: July 16, 2020

REFERENCES

1. Vinogradov, A.E., and Anatskaya, O.V. (2006). Genome size and metabolic intensity in tetrapods: a tale of two lines. Proc. Biol. Sci. *273*, 27–32.
2. Gregory, T.R. (2002). A bird's-eye view of the C-value enigma: genome size, cell size, and metabolic rate in the class aves. Evolution *56*, 121–130.
3. Vinogradov, A.E. (1997). Nucleotypic effect in homeotherms: body-mass independent resting metabolic rate of passerine birds is related to genome size. Evolution *51*, 220–225.
4. Pagel, M., and Johnstone, R.A. (1992). Variation across species in the size of the nuclear genome supports the junk-DNA explanation for the C-value paradox. Proc. Biol. Sci. *249*, 119–124.
5. Knight, C.A., Molinari, N.A., and Petrov, D.A. (2005). The large genome constraint hypothesis: evolution, ecology and phenotype. Ann. Bot. *95*, 177–190.

6. Gregory, T.R. (2005). Genome size evolution in animals. In The Evolution of the Genome, T.R. Gregory, ed. (Elsevier), pp. 3–87.
7. Baetcke, K.P., Sparrow, A.H., Nauman, C.H., and Schwemmer, S.S. (1967). The relationship of DNA content to nuclear and chromosome volumes and to radiosensitivity (LD50). Proc. Natl. Acad. Sci. USA *58*, 533–540.
8. Doolittle, W.F., and Sapienza, C. (1980). Selfish genes, the phenotype paradigm and genome evolution. Nature *284*, 601–603.
9. Orgel, L.E., and Crick, F.H. (1980). Selfish DNA: the ultimate parasite. Nature *284*, 604–607.
10. Cavalier-Smith, T. (2005). Economy, speed and size matter: evolutionary forces driving nuclear genome miniaturization and expansion. Ann. Bot. *95*, 147–175.
11. Gregory, T.R. (2001). The bigger the C-value, the larger the cell: genome size and red blood cell size in vertebrates. Blood Cells Mol. Dis. *27*, 830–843.
12. Gregory, T.R. (2001). Coincidence, coevolution, or causation? DNA content, cell size, and the C-value enigma. Biol. Rev. Camb. Philos. Soc. *76*, 65–101.
13. Beaulieu, J.M., Moles, A.T., Leitch, I.J., Bennett, M.D., Dickie, J.B., and Knight, C.A. (2007). Correlated evolution of genome size and seed mass. New Phytol. *173*, 422–437.
14. Rayburn, A.L., Dudley, J.W., and Biradar, D.P. (1994). Selection for early flowering results in simultaneous selection for reduced nuclear DNA content in maize. Plant Breed. *112*, 318–322.
15. Womack, M.C., Metz, M.J., and Hoke, K.L. (2019). Larger genomes linked to slower development and loss of late-developing traits. Am. Nat. *194*, 854–864.
16. Tenaillon, M.I., Manicacci, D., Nicolas, S.D., Tardieu, F., and Welcker, C. (2016). Testing the link between genome size and growth rate in maize. PeerJ *4*, e2408.
17. Vinogradov, A.E. (2003). Selfish DNA is maladaptive: evidence from the plant Red List. Trends Genet. *19*, 609–614.
18. Vinogradov, A.E. (2004). Genome size and extinction risk in vertebrates. Proc. Biol. Sci. *271*, 1701–1705.
19. Greilhuber, J. (1998). Intraspecific variation in genome size: a critical reassessment. Ann. Bot. *82*, 27–35.
20. Jovtchev, G., Schubert, V., Meister, A., Barow, M., and Schubert, I. (2006). Nuclear DNA content and nuclear and cell volume are positively correlated in angiosperms. Cytogenet. Genome Res. *114*, 77–82.
21. Neumann, F.R., and Nurse, P. (2007). Nuclear size control in fission yeast. J. Cell Biol. *179*, 593–600.
22. Jorgensen, P., Edgington, N.P., Schneider, B.L., Rupes, I., Tyers, M., and Futcher, B. (2007). The size of the nucleus increases as yeast cells grow. Mol. Biol. Cell *18*, 3523–3532.
23. Kurokawa, M., Seno, S., Matsuda, H., and Ying, B.W. (2016). Correlation between genome reduction and bacterial growth. DNA Res. *23*, 517–525.
24. Pellicer, J., Hidalgo, O., Dodsworth, S., and Leitch, I.J. (2018). Genome size diversity and its impact on the evolution of land plants. Genes (Basel) *9*, 88.
25. Malerba, M.E., and Marshall, D.J. (2019). Size-abundance rules? Evolution changes scaling relationships between size, metabolism and demography. Ecol. Lett. *22*, 1407–1416.
26. Malerba, M.E., Palacios, M.M., Palacios Delgado, Y.M., Beardall, J., and Marshall, D.J. (2018). Cell size, photosynthesis and the package effect: an artificial selection approach. New Phytol. *219*, 449–461.
27. Malerba, M.E., White, C.R., and Marshall, D.J. (2018). Eco-energetic consequences of evolutionary shifts in body size. Ecol. Lett. *21*, 54–62.
28. Malerba, M.E., and Marshall, D.J. (2020). Testing the drivers of the temperature-size covariance using artificial selection. Evolution *74*, 169–178.
29. Oliver, M.J., Petrov, D., Ackerly, D., Falkowski, P., and Schofield, O.M. (2007). The mode and tempo of genome size evolution in eukaryotes. Genome Res. *17*, 594–601.
30. Kempes, C.P., Dutkiewicz, S., and Follows, M.J. (2012). Growth, metabolic partitioning, and the size of microorganisms. Proc. Natl. Acad. Sci. USA *109*, 495–500.
31. Marañón, E. (2015). Cell size as a key determinant of phytoplankton metabolism and community structure. Annu. Rev. Mar. Sci. *7*, 241–264.
32. Malerba, M.E., Palacios, M.M., and Marshall, D.J. (2018). Do larger individuals cope with resource fluctuations better? An artificial selection approach. Proc. Biol. Sci. *285*, 1–9.
33. Nunn, C.L., and Cooper, N. (2015). Investigating evolutionary lag using the species-pairs evolutionary lag test (SPELT). Evolution *69*, 245–253.
34. Giannini, N.P., and Goloboff, P.A. (2010). Delayed-response phylogenetic correlation: an optimization-based method to test covariation of continuous characters. Evolution *64*, 1885–1898.
35. Thorén, S., Lindenfors, P., and Kappeler, P.M. (2006). Phylogenetic analyses of dimorphism in primates: evidence for stronger selection on canine size than on body size. Am. J. Phys. Anthropol. *130*, 50–59.
36. Deaner, R.O., and Nunn, C.L. (1999). How quickly do brains catch up with bodies? A comparative method for detecting evolutionary lag. Proc. Biol. Sci. *266*, 687–694.
37. Blomberg, S.P., and Garland, T., Jr. (2002). Tempo and mode in evolution: phylogenetic inertia, adaptation and comparative methods. J. Evol. Biol. *15*, 899–910.
38. Futuyma, D.J., and Moreno, G. (1988). The evolution of ecological specialization. Annu. Rev. Ecol. Syst. *19*, 207–233.
39. Walsh, B., and Blows, M.W. (2009). Abundant genetic variation + strong selection = multivariate genetic constraints: a geometric view of adaptation. Annu. Rev. Ecol. Evol. Syst. *40*, 41–59.
40. Pianka, E.R. (1970). On r- and K-selection. Am. Nat. *104*, 592–597.
41. Reznick, D., Bryant, M.J., and Bashey, F. (2002). r- and K-selection revisited: the role of population regulation in life-history evolution. Ecology *83*, 1509–1520.
42. Wilson, E.O., and MacArthur, R.H. (1967). The Theory of Island Biogeography (Princeton University).
43. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source platform for biological-image analysis. Nat. Methods *9*, 676–682.
44. R Development Core Team (2019). R: A language and environment for statistical computing (R Foundation for Statistical Computing). http://www.R-project.org/.
45. Pinheiro, J., Bates, D., DebRoy, S., and Sarkar, D.; R Core Team (2016). nlme: linear and nonlinear mixed effects models. R package version 3.1-128. http://CRAN.R-project.org/package=nlme.
46. Fox, J., and Weisberg, S. (2019). An R Companion to Applied Regression, Third Edition (Sage). https://socialsciences.mcmaster.ca/jfox/Books/Companion/.
47. Wickham, H., Averick, M., Bryan, J., Chang, W., McGowan, L.D., François, R., Grolemund, G., Hayes, A., Henry, L., Hester, J., et al. (2019). Welcome to the Tidyverse. J. Open Source Softw. *4*, 1686.
48. Wickham, H. (2011). The split-apply-combine strategy for data analysis. J. Stat. Softw. *40*, 1–29.
49. Wickham, H. (2009). ggplot2: Elegant Graphics for Data Analysis (Springer-Verlag).
50. Wilke, C.O. (2016). cowplot: streamlined plot theme and plot annotations for 'ggplot2'. R package version 0.7.0. https://CRAN.R-project.org/package=cowplot.
51. Guillard, R.R.L. (1975). Culture of phytoplankton for feeding marine invertebrates. In Culture of Marine Invertebrate Animals, W.L. Smith, and M.H. Chanley, eds. (Plenum), pp. 26–60.

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

52. Sun, J., and Liu, D. (2003). Geometric models for calculating cell biovolume and surface area for phytoplankton. J. Plankton Res. *25*, 1331–1346.

53. Boucher, N., Vaulot, D., and Partensky, F. (1991). Flow cytometric determination of phytoplankton DNA in cultures and oceanic populations. Mar. Ecol. Prog. Ser. *71*, 75–84.

54. Button, D.K., and Robertson, B.R. (2001). Determination of DNA content of aquatic bacteria by flow cytometry. Appl. Environ. Microbiol. *67*, 1636–1645.

55. Holm-Hansen, O. (1969). Algae: amounts of DNA and organic carbon in single cells. Science *163*, 87–88.

56. Malerba, M.E., White, C.R., and Marshall, D.J. (2017). Phytoplankton size-scaling of net-energy flux across light and biomass gradients. Ecology *98*, 3106–3115.

57. Williams, P.J.B., and Laurens, L.M.L. (2010). Microalgae as Biodiesel & Biomass Feedstocks: Review & Analysis of the Biochemistry (Energetics & Economics).

58. Burnham, K.P., and Anderson, D.R. (2004). Multimodel inference: understanding AIC and BIC in model selection. Sociol. Methods Res. *33*, 261–304.

59. Gingerich, P.D. (1993). Quantification and comparison of evolutionary rates. Am. J. Sci. *293*, 453–478.

60. Haldane, J.B.S. (1949). Suggestions as to quantitative measurement of rates of evolution. Evolution *3*, 51–56.

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological Samples		
Dunaliella tertiolecta (Butcher)	Australian National Algae Culture Collection (ANACC)	CS-14
Chemicals, Peptides, and Recombinant Proteins		
Lugol solution	Sigma-Aldrich	L6146-1L
DAPI (4′,6-diamidino-2-phenylindole)	Sigma-Aldrich	D9542
Glutaraldehyde	Sigma-Aldrich	G5882
CountBright counting beads	Thermofisher	C36950
Deposited Data		
Codes, data, and microscopy photos	Mendeley data	https://doi.org/10.17632/yz536nxyst.1
Experimental Models: Cell Lines		
Size-evolved lineages of the green microalga *Dunaliella tertiolecta*	Centre for Geometric Biology, Monash University	N/A
Software and Algorithms		
Fiji version 2.0.0	[43]	https://imagej.net/Fiji/Downloads
R version 3.6.1	[44]	CRAN mirror
nlme (R package)	[45]	CRAN mirror
car (R package)	[46]	CRAN mirror
tidyverse (R package)	[47]	CRAN mirror
plyr (R package)	[48]	CRAN mirror
ggplot2 (R package)	[49]	CRAN mirror
cowplot (R package)	[50]	CRAN mirror
Other		
PreSens Sensor Dish Readers	AS-1 Scientific	N/A
BD LSRFortessa X-20 cell analyzer	BD Biosciences	N/A
Leica DMi8 fluorescent inverted microscope	Leica	N/A
SPECTROstar ® Nano plate reader	BMG labtech	N/A

RESOURCE AVAILABILITY

Lead Contact

Further information for resources should be directed and will be fulfilled by the Lead Contact, Martino E. Malerba (martino.malerba@gmail.com).

Materials Availability

This study did not generate new unique reagents and used standard materials and methods. Unless stated otherwise, all chemicals were analytical grade and sourced from Sigma-Aldrich (Sigma Chemical Co., St Louis, MO, USA).

Data and Code Availability

All data and codes generated during this study have been deposited to Mendeley Data: https://doi.org/10.17632/yz536nxyst.1 (https://doi.org/10.17632/yz536nxyst.1).

EXPERIMENTAL MODEL AND SUBJECT DETAILS

We sourced the cosmopolitan, fast-growing green microalgal species *Dunaliella tertiolecta* (Butcher) from the Australian National Algae Culture Collection (ANACC; strain code CS-14). We cultured the mother cultures in a temperature-controlled room at 21 ± 1°C using autoclaved f/2 medium (no silica) from 0.45 μm-filtered seawater [51]. Light intensity was at 150 μM photos m^{-2} s^{-1} with a photoperiod of 14-10 day-night cycle, using low-heat 50 W LED flood lights (Power-lite™, Nedlands Group, Bedfordale, Australia).

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

CellPress

METHOD DETAILS

Artificial selection for cell size

For more details on the artificial selection protocols, refer to Malerba et al. [27]. The method relies on larger cells forming a pellet at the bottom of test tubes at lower centrifugal forces compared to smaller cells, which instead will remain in solution (i.e., differential centrifugation). On 25th April 2016, we inoculated 72 lineages from the same ancestral population of *D. tertiolecta* into aseptic 75 cm^2 plastic cell culture flasks (Corning, Canted Neck, Nonpyrogenic). Since then, we selected all lineages twice a week, each Monday and Thursday: 30 lineages were large-selected, 30 small-selected and 12 were the control. Control cultures experienced identical conditions (including centrifugation) without being size-selected. At the end of selection, we reinoculated all cultures into fresh f/2 medium. Lineages were not axenic, but we kept bacterial loads to low levels by resuspending pelleted cells in autoclaved medium twice a week and by handling samples using sterile materials under a laminar-flow cabinet (Gelman Sciences Australia, CF23S, NATA certified).

Experimental trials

We sampled cells after 350, 450 or 500 generations of artificial selection for size (see arrows in Figure 1). To remove any environmental effects and non-genetic phenotypic differences from artificial selection, we exposed cells to three generations (a week) of common garden conditions with no centrifugation (neutral selection) before starting trials. At generations 350 and 450, we analyzed cell size and nucleus size. At generation 500, we measured cell size, population growth, and DNA content.

Cell size

Following neutral selection, we measured the mean cell volume of all lineages sampled for this experiment, using light microscopy at 400X after staining cells with Lugol's iodine at 2%. We calculated the cell volume from around 200 cells per culture in Fiji 2.0.0 [43] assuming prolate spheroid shape, as recommended for this species by Sun and Liu [52].

DNA content

We used flow cytometry to estimate the mean cell DNA content in each lineage before starting growth trials. The intensity of the blue fluorescence (excitation: 355 nm, detection: 525 ± 50 nm) from DAPI (4′,6-diamidino-2-phenylindole) is proportional to the total DNA content in a phytoplankton cell [53, 54]. Samples were first fixed with 2% glutaraldehyde, then stained with DAPI (0.1 μg mL^{-1}), and after 30 minutes of dark-incubation were placed in a well plate and analyzed with the BD LSRFortessa X-20 cell analyzer (BD Biosciences, San Jose, CA, US). We used CountBright counting beads (Thermofisher, Waltham, MA, United States) as internal standards and we standardized all mean fluorescent values for the mean bead fluorescence. To present results in absolute units and facilitate comparisons among studies, optical values of bead-standardized fluorescence were centered to 0.336 pg DNA cell^{-1} – reported for this species in Holm-Hansen [55] – and scaled to an equivalent coefficient of variation (e.g., 1 ± 0.1 SE in fluorescent unit was converted to 0.336 ± 0.0336 pg DNA cell^{-1}).

The size-independent DNA content of a lineage was then calculated as the residuals on the linear relationship between the mean genome size of a lineage and its mean cell volume [$DNA\ content = 0.25\ e^{0.016 \times Cell\ Volume}$; $R^2 = 0.9$; Figure 1A]. Also, we ensured that alternative models to calculate residuals (i.e., major-axis regressions) yielded equivalent results ($r = 0.98$). A positive size-independent DNA content indicates a culture with a higher-than-average DNA content for its mean cell volume, whereas a negative value indicates a lower-than-average DNA content. In this way we could evaluate the effects of DNA content after removing all variance associated with cell size – a conservative approach compared to partitioning variance among correlated variables, such as DNA content and cell size.

Nucleus size

We measured nucleus volume for individual cells using microscopy. For each artificial selection treatment, we sampled 12 lineages after 350 and 450 generations (total of 72 lineages). Samples were first fixed in 2% glutaraldehyde and then rinsed and resuspended into growth medium. Fixed samples were diluted to approximately 3×10^6 cells mL^{-1} and stained with DAPI to attain a final dye concentration of 0.1 μg mL^{-1}. This dye can penetrate fixed cells and bind with DNA to form a fluorescence with an absorption maximum at 358 nm (ultraviolet) and an emission maximum at 461 nm (blue). We incubated all stained cells in the dark for 30 minutes and imaged them on a slide with a fluorescent inverted microscope (Leica DMi8). For each lineage, we took 20 photos with both brightfield view for cell size and with DAPI channel (excitation = 325-375nm; emission = 435-485nm) for nucleus size. We analyzed the photos in Fiji 2.0.0 to estimate the volume of cells and nuclei, assuming spheroid shapes for both. Finally, we converted nucleus volume (μm^3) into cell DNA content (pg) using a calibration curve across 36 lineages ($DNA\ content = 0.39 - 0.11\ e^{-0.055\ Nucleus\ Volume}$, $R^2 = 0.92$; Figure 1B). Hence, values of cell DNA content inferred from nucleus volume were unaffected by differences in organelle DNA among cells (see Methods S1).

Energy use

Rates of oxygen exchange were measured for all lineages at 21°C using 5×24-channel PreSens Sensor Dish Readers (SDR; AS-1 Scientific, Wellington, New Zealand). Methods were adapted from Malerba et al. [56]. Briefly, samples were placed in 5 mL sealed vials and randomly allocated to the top row of pre-calibrated SDR that were placed horizontally under the light source. Blanks

Please cite this article in press as: Malerba et al., Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*, Current Biology (2020), https://doi.org/10.1016/j.cub.2020.06.033

containing medium with no cells were placed in the row underneath (pilot study showed there was no row effect on blank measurements). Rates of oxygen exchange were recorded in the dark and at saturating light regimes (~200 μmol^{-1} quanta m^{-2} s^{-1}), with each lineage replicated four times. Mean rates of light and dark metabolism for each sample were calculated from averaging linear rates of O_2 change over time, standardized by population density, and converted into units of Joules $cell^{-1}$ min^{-1}, using the conversion factor of 512 × 10^{-3} J (μmol O_2)$^{-1}$ from Williams and Laurens [57]. Finally, the daily net energy budget of a cell (J $cell^{-1}$ day^{-1}) was calculated by adding the photosynthetic rate during the day and removing energy used at night (assuming a 14–10 h day–night photocycle).

Population growth

We collected estimates for max. rate of biomass production and total cell biovolume at carrying capacity for all 72 size-selected lineages after 500 generations of artificial selection with the methods described in Malerba et al. [32] and Malerba and Marshall [25]. Briefly, we loaded each lineage into three independent 96-well plates (Corning® polystyrene, flat bottom, with lid, sterile, non-treated, Sigma-Aldrich) after resuspending cells into standard fresh f/2 medium and standardizing initial populations to the same blank-corrected optical density (at 750 nm) – which we showed to be a reasonable proxy for the total biovolume in a culture [27, 32]. We loaded each sample into a 250 μL well after randomizing the position within the plate. We used light saturated conditions (> 200 μM m^{-2} sec^{-1}) at 14-10 h day-night cycle to grow all plates and we monitored the blank-corrected optical density (750 nm) for five days (at the same time into the photoperiod) using a plate reader SPECTROstar ® Nano (BMG labtech, Offenburg, Germany). This indirect way to measure biomass production allowed for far more frequent (non-destructive) monitoring compared to direct methods (e.g., flow cytometer). A pilot study showed that evaporation in the wells was low (~1% per day) and was therefore ignored in the analysis.

QUANTIFICATION AND STATISTICAL ANALYSIS

We used three simple linear regression models to analyze the effects of size-independent DNA content and size-selection treatments on three size-independent energy rates: photosynthesis, respiration and net-daily energy budget. Refer to figure legend for more statistical details of each analysis.

For the growth curves, we fitted three non-linear logistic-type models to describe the change in total biovolume production over time in each well: a logistic sinusoidal curve with lower asymptote forced to 0, a more general logistic curve with a non-zero asymptote, and a "Gompertz with mortality" sinusoidal model with a population decline after reaching carrying capacity (for a graphical explanation and more details on the models, see Figures S2 and S3 in [32]). From the best-fitting growth model following AIC [58], we extracted the maximum predicted value of total biovolume (K; unit of μm^3 μL^{-1}), which represents the total biomass reached at the end of the time-series. From the first derivative of the best-fitting growth model, we extracted the maximum intrinsic growth rate (r; unit of day^{-1}), which represents the maximum observed growth rate of the population.

Two separate linear mixed models were used to analyze total biomass (K) and maximum intrinsic growth (r) as a function of size-independent DNA content, the artificial selection treatment (i.e., control, small- or large-selected lineages), and their interaction. We also included initial population density and plate ID as fixed covariates, and lineage ID as a random intercept. All continuous variables were centered and standardized. We incorporated different classes of variance structures (functions *varFun* in the *nlme* package in R) and we used AIC model selection to determine the best-fitting model. After ensuring all assumptions were met, we used Wald chi-square test to calculate probability values for the explanatory variables in each best-fitting linear mixed-models (function *Anova* in the *car* package in R).

Finally, evolutionary rates of cell volume and DNA content were quantified using Haldane units, as $d/SD \times t$ where d is the difference of $\log_e$-transformed cell volumes between two means, SD is the pooled standard deviation of all $\log_e$ -transformed cell volumes and t is the generation interval. One Haldane unit is defined as a change by a factor of one standard deviation per generation [59, 60].

All analyses in this study were carried out in R version 3.6.1 [44] using packages nlme [45], car [46], tidyverse [47], and plyr [48] for model fitting and ggplot2 [49] and cowplot [50] for plotting.

Current Biology, Volume *30*

Supplemental Information

Genome Size Affects Fitness in the Eukaryotic Alga *Dunaliella tertiolecta*

Martino E. Malerba, Giulia Ghedini, and Dustin J. Marshall

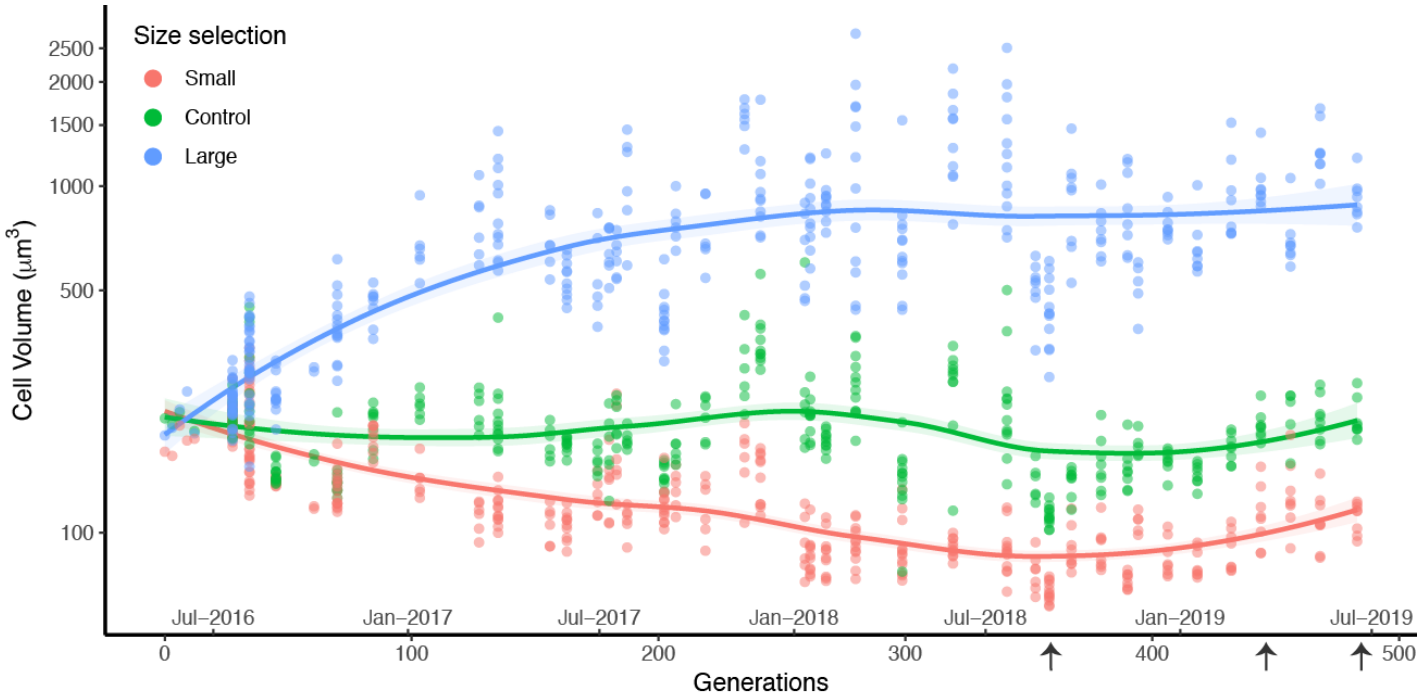

Figure S1: Three years (ca. 500 generations) of artificial selection to evolve the cell size of the green microalga *Dunaliella tertiolecta*. Related to Figures 1-4. Each point represents the mean cell volume of an independent lineage. Coloured lines indicate the fit of a LOESS smoother for each size-selection treatment (either small-selection, large-selection or control). Arrows on the x-axis indicate the trials for this experiment. We measured cell size and nucleus size after 350 and 450 generations, energy rates after 350 generations, and population growth and DNA content after 500 generations. Throughout all assays, the cell volume of large-selected cells was on average between 6.5 and 11.7 times larger than small-selected cells.

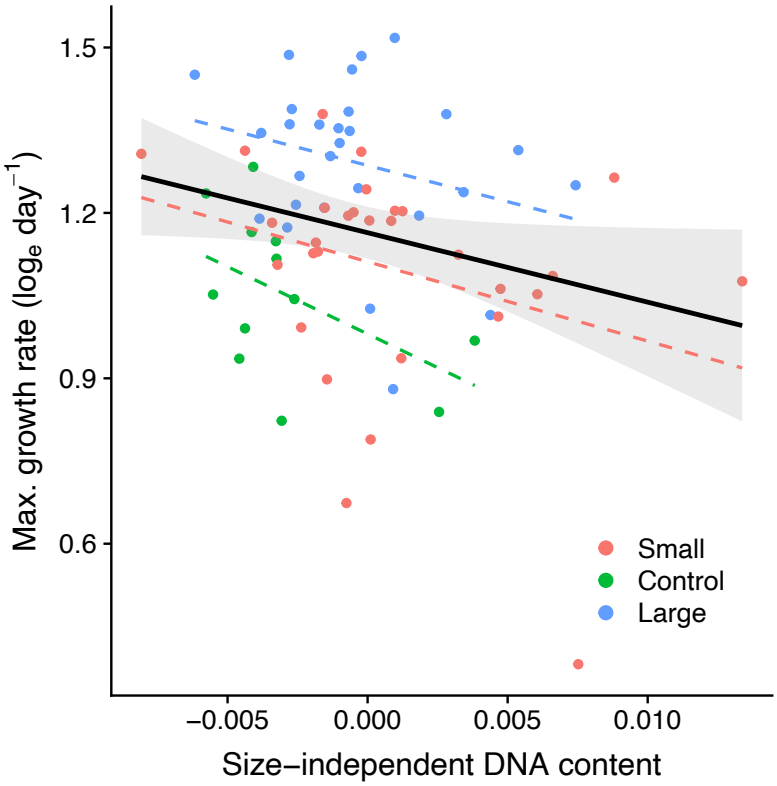

Figure S2: To confirm the negative relationship between max. growth rate and relative genome size, we ran a second experiment with higher initial per-capita resource availability where we monitored the first 24 hours of growth in all lineages. Related to Figure 2A. The qualitative differences remained identical to Figure 2A: all negative slopes were significant ($F_{1,64} = 5.66$, $p = 0.02$) and there was no sign. interaction among dashed lines ($F_{2,64} = 0.17$, $p = 0.84$). Values on x-axis are the mean size-independent DNA content of a lineage. Colours indicate the size-selection treatment. Dashed lines indicate the best-fitting linear mixed-effect models at each size-selection treatment, whereas continuous lines indicate the overall effects across all data (±95% C.I.). Each dot represents the mean value measured from an independently selected lineage.

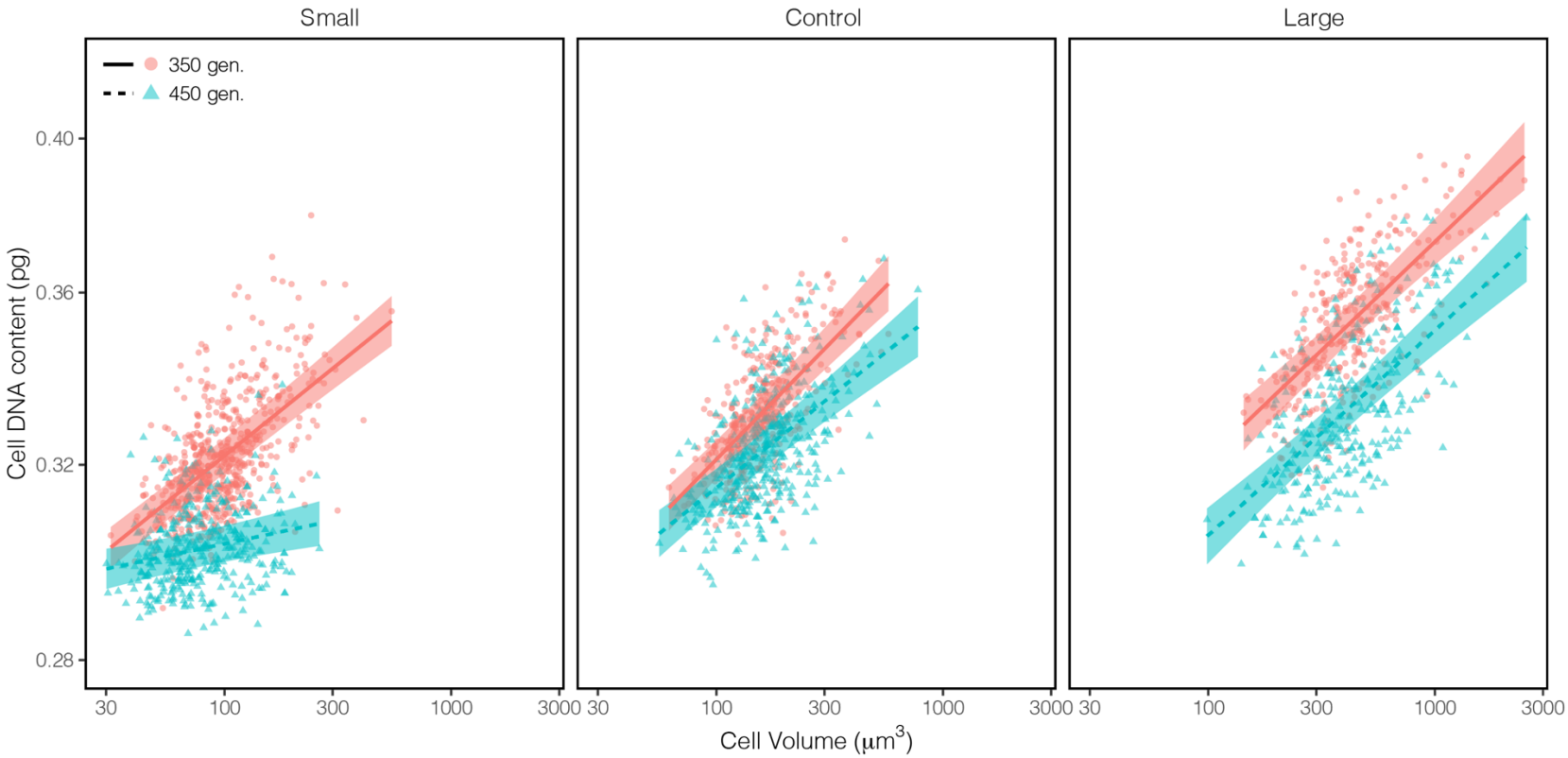

Figure S3: Relationship between cell volume and cell DNA content divided by generation (line type and colour) and artificial size-selection treatment (columns). Relate to Figure 4. Each point is a cell after correcting for random covariates (i.e. lineage identity nested within generation). Lines [±95% C.I.] indicate the best-fitting linear mixed-effect model. These data are the same as those presented in Figure 4.

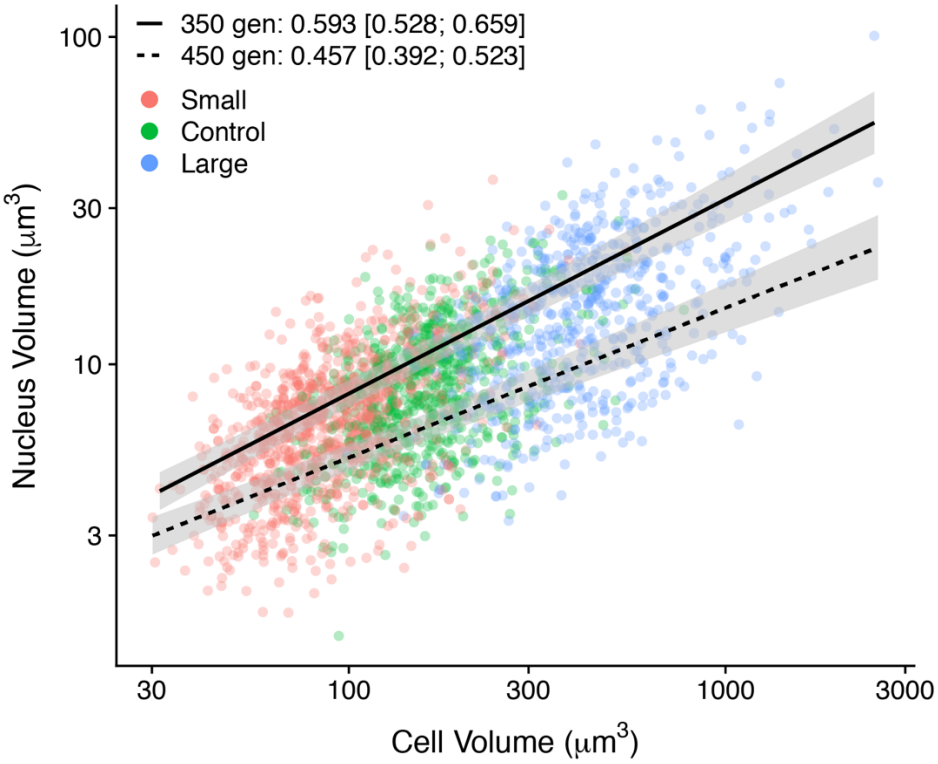

Figure S4: Relationship between nucleus volume (µm^3) and cell volume (µm^3) in lineages of *Dunaliella tertiolecta* after 350 (solid line) and 450 (dashed line) generations of artificial selection for size. Related to Figure 4. Lines [±95% C.I.] indicate the best-fitting linear mixed-effect model, with size-scaling coefficients reported in the legend. Because the size-scaling exponent is <1 (i.e. hypo-allometric), it implies that larger cells have lower nucleus volume to cell volume ratios compared to smaller cells. Each point is a cell after correcting for random covariates (i.e. lineage identity nested within generations), with the colour representing the artificial size-selection treatment. See Figure 4 for the same data converted from nucleus volume to cell DNA content using the model in Figure 1B.

Response: $\log_e(K)$

Explanatory:	*Chisq*	*Df*	*Pr(>Chisq)*
Size-Ind. DNA	3.9423	1	0.047
Treatment	362.2128	2	< 0.001
Plate	49.1832	2	< 0.001
Init. Biovol.	67.1069	1	< 0.001

Response: $\log_e(r)$

Size-Ind. DNA	6.3008	1	0.012
Treatment	23.056	2	< 0.001
Plate	16.5667	2	< 0.001
Init. Biovol.	140.8615	1	< 0.001

Table S1: ANOVA table with type II Wald chi-square test for the best-fitting models (following AIC) for population-level traits of maximum specific growth rate (*r*) and maximum total biovolume (*K*). Related to Figure 2. Fixed effects are the mean size-independent DNA content (i.e. the residuals in Figure 1A; *Size-Ind. DNA*), the artificial selection treatment (*Treatment*), the plate ID (*Plate*), and the initial biovolume in the culture (*Init. Biovol.*). Lineage identity was included as a random intercept in both models. The interaction between Size-Ind. DNA and Treatment was included in the initial model, but was later selected against by AIC model selection.