



## **Hydra has mammal-like mutation rates facilitating fast adaptation despite its nonaging phenotype**

Arne Sahm, Konstantin Riege, Marco Groth, et al.

*Genome Res.* published online December 4, 2024

Access the most recent version at doi:[10.1101/gr.279025.124](https://doi.org/10.1101/gr.279025.124)

---

**P<P** Published online December 4, 2024 in advance of the print journal.

**Open Access** Freely available online through the *Genome Research* Open Access option.

**Creative Commons License** This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

**Email Alerting Service** Receive free email alerts when new articles cite this article - sign up in the box at the top right corner of the article or [click here](#).

---

Advance online articles have been peer reviewed and accepted for publication but have not yet appeared in the paper journal (edited, typeset versions may be posted when available prior to final publication). Advance online articles are citable and establish publication priority; they are indexed by PubMed from initial publication. Citations to Advance online articles must include the digital object identifier (DOIs) and date of initial publication.

---

To subscribe to *Genome Research* go to:  
<https://genome.cshlp.org/subscriptions>

## Research

# *Hydra* has mammal-like mutation rates facilitating fast adaptation despite its nonaging phenotype

Arne Sahn,<sup>1,2,3</sup> Konstantin Riege,<sup>2</sup> Marco Groth,<sup>4</sup> Martin Bens,<sup>4</sup> Johann Kraus,<sup>5</sup> Martin Fischer,<sup>2</sup> Hans Kestler,<sup>5</sup> Christoph Englert,<sup>6,7</sup> Ralf Schaible,<sup>2</sup> Matthias Platzer,<sup>2,8</sup> and Steve Hoffmann<sup>2,8</sup><sup>1</sup>Computational Phenomics group, IUF—Leibniz Research Institute for Environmental Medicine, 40225 Düsseldorf, Germany;<sup>2</sup>Computational Biology Group, Leibniz Institute on Aging—Fritz Lipmann Institute (FLI), 07745 Jena, Germany; <sup>3</sup>Computational Phenomics group, Faculty of Biology and Biotechnology, Ruhr University Bochum, 44801 Bochum, Germany; <sup>4</sup>Core Facility Next Generation Sequencing, Leibniz Institute on Aging—Fritz Lipmann Institute, 07745 Jena, Germany; <sup>5</sup>Institute of Medical Systems Biology, Ulm University, 89081 Ulm, Germany; <sup>6</sup>Molecular Genetics Lab, Leibniz Institute on Aging—Fritz Lipmann Institute (FLI), 07745 Jena, Germany; <sup>7</sup>Institute of Biochemistry and Biophysics, Friedrich Schiller University Jena, 07745 Jena, Germany

Growing evidence suggests that somatic mutations may be a major cause of the aging process. However, it remains to be tested whether the predictions of the theory also apply to species with longer life spans than humans. *Hydra* is a genus of freshwater polyps with remarkable regeneration abilities and a potentially unlimited life span under laboratory conditions. By genome sequencing of single cells and whole animals, we found that the mutation rates in *Hydra*'s stem cells are even slightly higher than in humans or mice. A potential explanation for this deviation from the prediction of the theory may lie in the adaptability offered by a higher mutation rate, as we were able to show that the genome of the widely studied *Hydra magnipapillata* strain 105 has undergone a process of strong positive selection since the strain's cultivation 50 years ago. This most likely represents a rapid adaptation to the drastically altered environmental conditions associated with the transition from the wild to laboratory conditions. Processes under positive selection in captive animals include pathways associated with *Hydra*'s simple nervous system, its nucleic acid metabolic process, cell migration, and hydrolase activity.

[Supplemental material is available for this article.]

*Hydra* is a genus of small freshwater polyps with a predicted laboratory life span of 1400 years for *Hydra magnipapillata* (Jones et al. 2014). Thus, it is often referred to as “nonsenescent” (Dańko et al. 2015; Schaible et al. 2015; Klimovich et al. 2018), “nonaging” (Sturm et al. 2017; Holtze et al. 2021), or “immortal” (Müller 1996; Boehm et al. 2012; Domazet-Lošo et al. 2014). This controversial attribution is based on the observation that, in stark contrast to most other animals, all *Hydra* species investigated under laboratory conditions avoid senescence while reproducing asexually (Martínez 1998; Schaible et al. 2014). This extraordinary longevity seems to be linked to the simplicity of *Hydra*'s body plan, the homeostatic somatic maintenance of the adult polyp by replacing all cells approximately every 20 days (Campbell 1967), and its high regenerative capacity. The animals can replace any part of their body when dissected into multiple fragments. Each part usually regrows into a new organism within 2–3 days (Fujisawa 2003; Tomczyk et al. 2015), as demonstrated in the most extreme case from a fragment of only 1% of the body mass (Shimizu et al. 1993). This ability to completely rebuild all tissues without leaving scars or evidencing any restrictions is the subject of intensive ongoing research, especially with regard to how wounding, injury signaling, and pattern formation are molecularly linked in this genus. Consequently, *Hydra* is, in addition to being a powerful model for aging studies (Tomczyk et al. 2015), a premier organism for the study of stem cells and regeneration in

general as well as for regenerative medicine (Michael 2019; Vogg et al. 2019b; Holtze et al. 2021).

Contributing factors to these abilities might be the simplicity of *Hydra*'s body plan and the fact that most of its cells are continually proliferating stem cells (Holstein and David 1990). Both faster cycling interstitial cells (I cells) and slower dividing ectodermal and endodermal epithelial cells (E cells) have been described as main stem cell lineages in *Hydra*. While I cells are multipotent and differentiate to neurons, nematocytes, gland cells, and gametes, E cells are unipotent, forming either the ecto- or the endoderm of the body column and the extremities. These lineages cannot replace one another, but combined they can differentiate into all *Hydra* cell types. They renew all the cells of an individual through continuous replacement within 3 weeks (Bosch 2009; Hemmrich et al. 2012; Domazet-Lošo et al. 2014; Siebert et al. 2019). Closely linked to this constant renewal is *Hydra*'s ability to reproduce asexually by budding. During this process, a bud containing one-fifth of the ~50,000–100,000 cells of an adult individual is formed and detached (Dańko et al. 2015; Tomczyk et al. 2015). *Hydra* reproduces asexually in ~19°C water under nutrient-rich conditions, which are usually fulfilled when kept in captivity. Under less favorable conditions, however, *Hydra* regularly switches to sexual reproduction (Martínez and Bridge 2012; Kaliszewicz and Lipińska 2013).

All *Hydra* individuals that originate vegetatively by budding from a single sexually originated ancestor are called a clonal colony or “genet.” In the case of sexual reproduction, soon after fertilization *Hydra* embryonal cells differentiate into three specific stem

<sup>8</sup>Shared senior-authorship.Corresponding author: [arne.sahn@rub.de](mailto:arne.sahn@rub.de)Article published online before print. Article, supplemental material, and publication date are at <https://www.genome.org/cgi/doi/10.1101/gr.279025.124>. Freely available online through the *Genome Research* Open Access option.© 2024 Sahn et al. This article, published in *Genome Research*, is available under a Creative Commons License (Attribution-NonCommercial 4.0 International), as described at <http://creativecommons.org/licenses/by-nc/4.0/>.

cell lineages, including I cells and E cells, that will behave and evolve as separate cell lineages within the genet. Mutations in these lineages that occur during vegetative reproduction can be regarded with respect to the genet as “somatic”—in analogy to mutations occurring in the somatic cells of a multicellular organism with dedicated reproductive cells. Similarly, mutations that have occurred in I-cell lineages and later transform into gametes that give rise to new genets resemble germline mutations in higher organisms. For convenience and simplicity, we denote mutations dating back to before the last sexual reproduction event as “germline” with respect to the genet. Of course, it is important to note that this simplified nomenclature is relative, as “somatic” mutations in a genet may become “germline” for subsequent ones. Nevertheless, for a given genet *Hydra*, “germline” variations are strictly biallelic and are present in all cells of the genet—unless a locus is affected by a second mutation and the resulting allele is transmitted to a subset of cells.

In *Hydra oligactis*, lowering the water temperature can induce aging. Gene expression comparisons between aging and nonaging *Hydra* have highlighted the role of E cells and processes such as apoptosis and autophagy for this nonsenescent phenotype (Sun et al. 2020; Tomczyk et al. 2020). A number of fundamental studies and resources have been published in recent years that examine *Hydra* from a molecular perspective. These include a low-coverage and therefore quasi-monoallelic *H. magnipapillata* strain 105 genome reference (Chapman et al. 2010) and transcriptome sequences (Wenger and Galliot 2013), inducible methods for gene knockdown (Watanabe et al. 2014; Vogt et al. 2019a), topological analysis of morphogenesis (Maroudas-Sacks et al. 2021), as well as quantitative gene expression data from whole individuals and single cells (Krishna et al. 2013; Petersen et al. 2015; Siebert et al. 2019).

Evolutionary forces such as mutation rate and selection have not been explored. Importantly, *Hydra*'s nonsenescent phenotype is not likely to be the result of natural selection—but so-called exaptation (Gould and Vrba 1982)—toward a longer life span but rather the consequence of features that evolved to increase survival under harsh environmental conditions (Schaible et al. 2014). In the wild, due to a high extrinsic mortality, *Hydra*'s average life span is counted in weeks rather than in years (Schaible et al. 2015). The nonsenescent phenotype is observable only under benign laboratory conditions, i.e., an absence of extrinsic mortality risks, steady feeding, and optimal temperature. As a prerequisite for *Hydra*'s regeneration capabilities, a high cell turnover and stem cell omnipotence have been recorded (Schaible et al. 2017). Nonsenescence of *Hydra* individuals, however, does not exclude the possibility that stem cells and differentiated cells could “wear out” and senesce over time, in the manner of individual bacteria (Stewart et al. 2005; Wang et al. 2010).

Somatic mutations are frequently considered as one or even the major factor of the aging process (López-Otín et al. 2013; Vijg 2021; Wolf 2021; Yousefzadeh et al. 2021). Essentially, it is assumed that an increasing mutational burden gradually undermines the integrity of individual cells, followed by that of tissues, organs, and finally the entire organism. This process leads to progressive decay and aging-associated diseases. Accordingly, it has been shown that across a range of very different long-lived mammalian species, mutation rates strongly negatively correlate with life span (Cagan et al. 2022). Out of the species examined by Cagan et al., humans were the longest-lived and had the lowest somatic mutation rate. Thus, the question arises of whether species that live longer than humans have an even lower mutation rate.

Therefore, this paper tests the hypothesis that *Hydra*, specifically the well-studied *H. magnipapillata* strain 105, with its predict-

ed maximal life span of over 1000 years (Schaible et al. 2014), has a significantly lower mutation rate than humans. This purely female strain has reproduced exclusively asexually ever since it was established in 1973 from a single individual (Sugiyama and Fujisawa 1977; Schaible et al. 2015). Furthermore, we investigated the relationship between the mutation rate and adaptability of this strain.

## Results

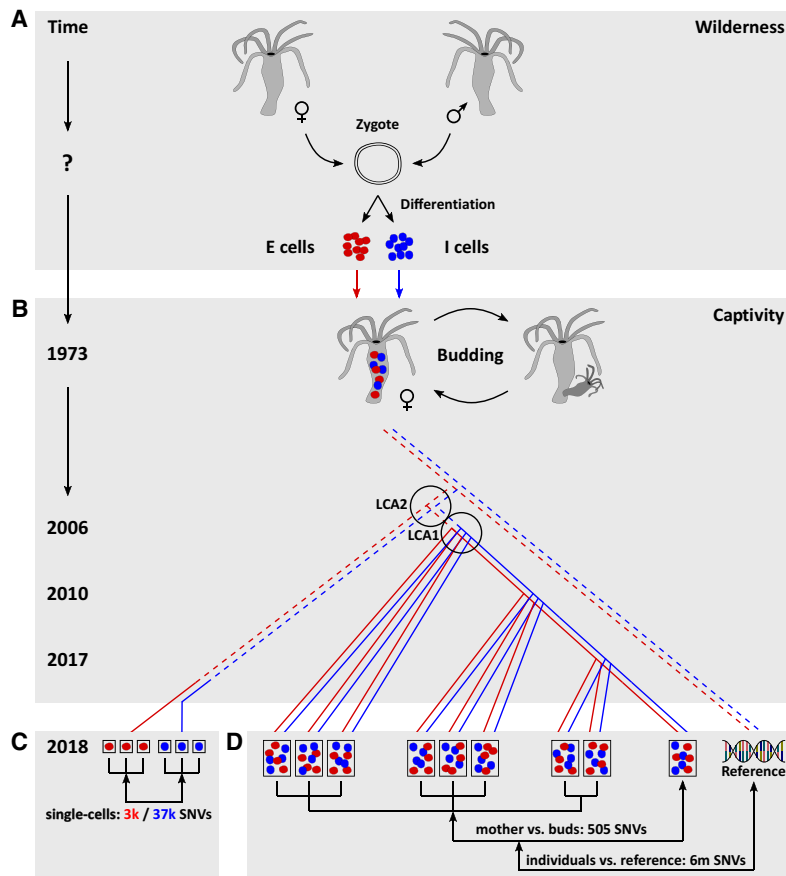
### *Hydra*'s mutation rate is at least as high as that of humans

In mammals, somatic mutations can be identified by the genomic sequencing of single cells, on the one hand, and genomic bulk sequencing of the same individual, on the other hand. Single-cell genomic variations that are present neither in the reference genome nor in the bulk are most likely the result of somatic mutations. This approach is based on the reasonable assumption that the bulk consensus represents the genomic state of the zygote. Consequently, deviations from this state would be caused by somatic mutations (Milholland et al. 2017).

In the common, asexually reproducing *H. magnipapillata* strain 105, however, counting the mutations in this way over the life span of a single individual is not possible: unlike mammals whose life begins with a single cell, the zygote, each of these *Hydra* individuals starts already with at least 10,000 cells forming a genetic mosaic. Nonetheless, the same principle as in mammals could be applied if the mutations were counted over the life span of a genet—in our case the strain 105, whose life started also with a single zygote sometime before the strain's cultivation in 1973 (Fig. 1A).

A challenge, however, is the fact that bulk sequencing in *Hydra* is much less likely to represent the genomic state of the zygote compared to mammals. The reason is that each of the many budding processes that occurred from the cultivation of the strain until today represents a bottleneck, making founder effects likely—which would clearly bias the final mutation count. Therefore, we instead exploited the strict separation of the *Hydra* I- and E-stem cell lineages mentioned above, which occurs very early after the zygote state (Bosch 2009; Hemmrich et al. 2012; Domazet-Lošo et al. 2014; Siebert et al. 2019). Thus, our rationale for counting the somatic mutations in each of the I cells is to search for variants that are present neither in any of the E cells nor in the reference genome of the strain (and vice versa for the E cells). Following this rationale, we identified the somatic single-nucleotide variants (SNVs) in six cells—three I cells and three E cells—taken from a single *Hydra* individual (Fig. 1C). For comparability, we used the same technical approach employed by others to determine mutation rates in mice and humans (Milholland et al. 2017).

We found 37,379 and 3184 unique positions affected by mutations in the stem cell lineages of I cells and E cells, respectively. On average, we found 14,459 SNVs per stem-like I cell and 1083 SNVs per stem-like E cell (Table 1) taking into account only positions that were covered by at least 20 reads. Dividing the number of SNVs by the number of genomic positions reaching this threshold, we observed on average one SNV for every 5770 bases in I cells and one for every 53,492 bases in E cells. To determine the somatic mutation rate on this basis, we needed to estimate the number of cell divisions that had occurred since the zygote state. This was done using two assumptions: firstly, by using published cell division rates—i.e., every 24–30 h and every 3–4 days, for I cells and E cells, respectively (David and Campbell 1972; Campbell and David 1974); secondly, by assuming that the investigated I cells and E cells evolved separately for 44 years—i.e., presuming that the last sexual reproduction took



**Figure 1.** Biological and historical background, experimental setup, and pedigree relationships of analyzed genomic data sets of *H. magnipapillata* strain 105. (A) The last sexual reproduction, representing the strain's genetic origin, occurred in the wild. Shortly after fertilization, a *Hydra* zygote differentiates into noninterchangeable endodermal and ectodermal epithelial (E cells, red) and interstitial (I cells, blue) stem cell lineages. (B) The *H. magnipapillata* strain 105 was derived from a single female polyp that was caught in the wild in 1973 (Chapman et al. 2010; Schaible et al. 2015). The strain has since reproduced exclusively asexually by budding and has been propagated in laboratories worldwide. For generating a reference genome sequence, the strain was recloned from a single polyp in the laboratory of Dr. Hans Bode at UC Irvine in 2004 (Chapman et al. 2010). We obtained the strain from UC Irvine in 2006 and established a pedigree consisting of its mother and eight daughters (buds) generated at three time points: 2006, 2010, and 2017. (C) In 2018, we used single-cell whole-genome sequencing of three E cells and three I cells to identify single-nucleotide variants (SNVs) that had accumulated in each of these cells since their common origin (zygote). (D) In addition, we performed whole-animal sequencing of nine *Hydra* individuals, comprising the pedigree mentioned above. We identified SNVs that arose in each daughter in comparison to the mother. To further identify heterozygous genome positions of the strain, we also compared the data of all nine polyps individually with a reference genome sequence determined in another laboratory from more distantly related individuals (Chapman et al. 2010). The total numbers of uniquely identified SNV positions are displayed: in C for the comparisons of each E cell versus all I cells combined (red) and each I cell versus E cells combined (blue); in D for the comparisons of buds versus mother and individuals versus reference. In our experimental setting, the last common ancestor (LCA) 1 of the nine-individual pedigree and the individual used for single-cell sequencing was budded also in 2006. The LCA of the individuals sequenced by us and those from which the reference genome was derived can be dated back only to a time window of 1974–2004. Pedigree lines: continuous—life history of a *Hydra* individual; dashed—descant relation including multiple *Hydra* individuals derived from each other by budding.

place shortly before strain 105 originated from a single female polyp in 1973 and that reproduction since then had occurred exclusively asexually by budding until our sampling in 2018.

Based on these assumptions, we inferred an average mutation rate per cell division of  $1.21 \times 10^{-8}$  for I cells and  $4.06 \times 10^{-9}$  for E cells (Fig. 2A). In each of the 6 cells studied, the mutation rate in the coding sequence is lower than in the rest of the genome ( $P=0.031$ , paired Wilcoxon test), although the differences are

smaller for E cells. On average, the estimated mutation rate per cell division in the coding sequence is  $1.02 \times 10^{-8}$  for I cells and  $3.46 \times 10^{-9}$  for E cells (Fig. 2A; Supplemental Tables S1, S2).

### Potential error sources are unlikely to affect the order of magnitude of the mutation rate estimates

Our mutation rate calculations are based on an estimate of 14,292 and 4605 mitoses in I cells and E cells, respectively. We investigated the influence of the number of cell divisions—particularly their underestimation—on the magnitude of the mutation rate estimate (Fig. 2B). We found that the number of divisions of E cells would have to be more than 50% higher than assumed to fall below the mutation rate of humans. This would imply that the last sexual reproduction in respect to the polyp that gave rise to strain 105 would have to have occurred more than 20 years before it was captured. In the case of I cells, the number of mitoses would need to be more than four times higher than assumed. In this scenario, the last sexual reproduction would have to have occurred more than 200 years ago, at the beginning of the nineteenth century.

Moreover, we compared the SNVs identified in the six single cells together with their respective 5' and 3' contexts (Supplemental Fig. S1) with known single-cell artifact signatures. Based on this analysis, the total fraction of potential single-cell artifacts in our data set was estimated to range from 1.3% to 3.5% (Miller et al. 2022; Supplemental Table S2). An example of a potential artifact would be spurious mutations caused, e.g., by elevated temperatures during the cell lysis step. Typical for this artifact is a mutation spectrum that is largely dominated by C-to-T transitions due to the possibility of cytosine deamination under high temperatures (Dong et al. 2017). The fraction of these mutations in our data is slightly higher than in humans and mice (Fig. 2C). Furthermore, we analyzed the evenness of single-cell genome amplification by comparing the absolute  $\log_2$  ratios of adjacent bins

(Miller et al. 2022). The median absolute pairwise difference (MAPD) ranged from 0.78 to 1.28 in the single-cell experiments (Table 1).

### Single-nucleotide variants are detectable over different evolutionary timescales

As described above, we used single-cell sequencing to identify somatic variants that would have been present since the last

**Table 1.** SNV calling in single cells (<1973–2018)

	I cells			E cells		
	I cell 1	I cell 2	I cell 3	E cell 1	E cell 2	E cell 3
# Mapped reads (million) <sup>a</sup>	267.10	323.26	357.83	204.68	82.67	123.79
% Mapped reads <sup>a</sup>	69.06	79.93	80.00	63.14	14.89	37.06
Whole genome						
# Covered bases (million) <sup>b</sup>	67.79	74.15	105.54	102.06	22.49	52.69
# Mutations	11,098	12,551	19,729	1855	450	943
Mutations per base	$1.64 \times 10^{-4}$	$1.69 \times 10^{-4}$	$1.87 \times 10^{-4}$	$1.82 \times 10^{-5}$	$2.00 \times 10^{-5}$	$1.79 \times 10^{-5}$
Mutation rate per cell division	$1.15 \times 10^{-8}$	$1.18 \times 10^{-8}$	$1.31 \times 10^{-8}$	$3.95 \times 10^{-9}$	$4.35 \times 10^{-9}$	$3.89 \times 10^{-9}$
Coding sequence						
# Covered bases (million) <sup>b</sup>	4.37	4.689	6.45	6.20	1.41	2.86
# Mutations	633	645	1,007	73	26	50
Mutations per base	$1.45 \times 10^{-4}$	$1.38 \times 10^{-4}$	$1.56 \times 10^{-4}$	$1.18 \times 10^{-5}$	$1.85 \times 10^{-5}$	$1.75 \times 10^{-5}$
Mutation rate per cell division	$1.01 \times 10^{-8}$	$9.63 \times 10^{-9}$	$1.09 \times 10^{-8}$	$2.56 \times 10^{-9}$	$4.01 \times 10^{-9}$	$3.79 \times 10^{-9}$
Validation						
MAPD score	0.783	0.934	0.696	0.687	1.269	1.124
% Single-cell artifact signatures <sup>c</sup>	1.3	1.3	1.8	2.5	3.5	2.9

<sup>a</sup>Uniquely mapping reads after removal of duplicates; lower on average read numbers of E cells versus I cells were caused by overall different outputs of the respective sequencing runs.

<sup>b</sup>(Coverage  $\geq 20$ , both in the single-cell examined and the merged mapping result of the single cells of the opposite stem cell lineage.

<sup>c</sup>Combined SBS scE and SBS scF fractions (Petljak et al. 2019).

occurrence of sexual reproduction before captivity, most likely shortly before 1973. In addition, we examined two kinds of SNVs by whole-animal genome sequencing of a pedigree, consisting of its mother and eight daughters (buds) generated at different time points between 2006 and 2018 (Fig. 1D; Supplemental Figs. S4–S9). These two kinds of variants would have originated either within a shorter timescale (from 2006 to 2018) or over a much longer timescale, before the last sexual reproduction (i.e., well before 1973) (see below and Supplemental Results), respectively.

For the shorter timescale, within the nine-animal pedigree, we compared each daughter (bud) against the mother using strict filtering. This resulted in 809 SNVs across the eight buds covering 505 unique variant positions (Table 2; Supplemental Results; Supplemental Table S3). In this setting, the LCA 1 denotes the mother at her stage of somatic evolution in 2006, i.e., 12 years before sampling (Fig. 1B).

For the longer timescale, we examined each of the sequenced individuals of our nine-individual pedigree for deviations from a monoallelic 850 Mb sized reference genome (Chapman et al. 2010). The identified SNV alleles occurred with high allele frequencies and were often coupled to each other. Therefore, they were most likely already present in the two haploid gamete genomes that combined in the zygote that later gave rise to the individual caught in 1973 from the wild (see Supplemental Results). These germline variants provide information about evolutionary pressures on *Hydra* in the wild. We detected ~4.5 million SNVs per comparison and 6 million unique SNV positions in total across all samples (Supplemental Table S3).

#### Multiple appearance and homozygosity of SNV alleles generated by somatic mutations

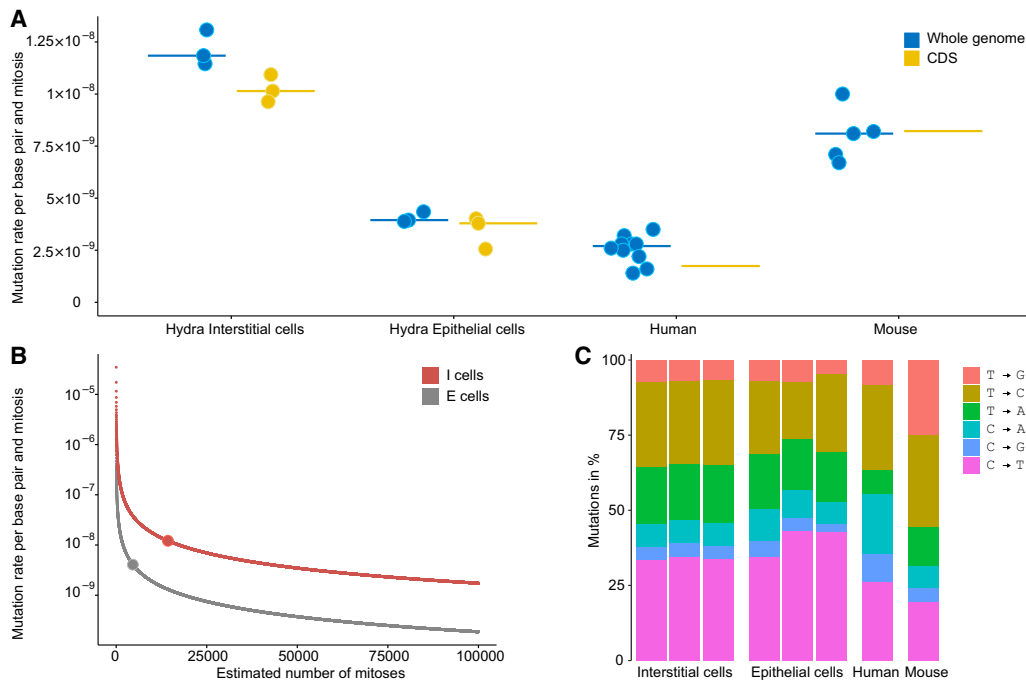
In total, 69.9% of variants identified in single cells were also found in the comparison of the individuals versus the reference. We

found a U-shaped bipartite pattern, with most single-cell SNVs being present either in all nine examined individuals or in none (Fig. 3A). The three I cells examined shared more variants than the E cells (Fig. 3B). Compared to E cells, we found I cell SNVs to be detectable more often at the level of *Hydra* individuals (Fig. 3C) ( $P < 2.2 \times 10^{-16}$ , Wilcoxon test). In addition, somatic SNVs that were identified in two I cells were found in more individuals than those found only in one I cell ( $P < 2.2 \times 10^{-16}$ ). The same relation applies for SNVs found in all three versus two I cells (Fig. 3C) ( $P = 2.9 \times 10^{-3}$ ).

Moreover, 19% of the variants detected in the single-cell experiment were homozygous. Homozygous variants occurred much more frequently in the single-cell experiment than would be expected from their frequency in the genome. Depending on the different classes of shared single-cell SNVs, we observed them 580–22,000 times more frequently than expected (Fig. 3A). We examined the prevalence of SNVs detected at the single-cell level in the comparison of the individuals versus the reference. We found homozygous variants exhibiting higher minor allele frequencies than heterozygous ones (median: 0.43 vs. 0.28,  $P < 2.2 \times 10^{-16}$ , Wilcoxon test). Furthermore, the percentage of homozygosity was considerably higher among single-cell SNVs found in all nine individuals compared to those in none of the individuals (27% vs. 9%,  $P < 2.2 \times 10^{-16}$ , Fisher's exact test).

#### Opposite selection patterns manifest before and after a recent drastic environmental change

We considered how strongly the strain was affected by positive and negative selection across the three evolutionary timescales examined. For this purpose, we examined the variants identified in coding sequences and categorized them into nonsynonymous and synonymous substitutions (Supplemental Tables S2, S3). Based on the respective counts, our analysis then compares the rate of



**Figure 2.** Single-cell whole-genome analysis to determine the mutation rate per cell division in *Hydra*. (A) Mutation rates per cell division in *Hydra* I cells and E cells in genome and coding sequences (CDS) compared with literature data for human and mouse dermal fibroblasts (Milholland et al. 2017). Points represent single sample estimates, lines denote medians (Supplemental Table S2) (for CDS of human and mouse only averages were published). (B) Dependence of mutation rate estimates on the number of mitoses since the zygote. Larger dots represent predicted mutation rates shown in A based on our estimated numbers of mitoses—14,292 and 4605 for I cells and E cells, respectively. (C) Mutation spectra in *Hydra* of the three I cells and three E cells compared to literature data for human and mouse dermal fibroblasts (Milholland et al. 2017).

nonsynonymous variants ( $d_N$ ) as an indicator of selection, against that of synonymous variants ( $d_S$ ) representing the genetic drift. Therefore, a  $d_N/d_S$  ratio  $\gg 1$  indicates positive selection,  $\sim 1$  indicates neutral evolution or genetic drift, and  $\ll 1$  negative selection (Kimura 1980). In line with this, our single-cell comparison covering the intermediate timescale from the zygote onward, i.e., most

likely, mainly the period in captivity (1973–2018) (Fig. 1), and revealing somatic SNVs, showed a genome-wide average  $d_N/d_S$  ratio of 0.63. On the longest evolutionary timescale investigated here and mainly reflecting germline variants accumulated in the wild before sexual reproduction (comparison of whole-animal data versus reference), negative selection was even more pronounced, with

**Table 2.** SNV calling in nine-animal pedigree

Individuals	Youngest budded 2018		Middle budded 2012			Oldest budded 2006			Mother
Age (years)	0.08	0.08	7.20	7.30	8.10	11.10	12.00	12.30	12.60
# Raw reads	193.91	151.63	290.73	304.89	302.57	281.69	301.31	304.27	191.77
# Mapped reads (million) <sup>a</sup>	146.97	114.18	199.57	209.62	206.60	191.17	200.59	217.67	143.28
% Mapped reads <sup>a</sup>	80.08	79.76	72.02	71.98	71.65	70.96	69.52	74.63	79.03
Mother versus buds (2006–2018)									
# Downsampled reads (million)	141.19	141.18	141.18	141.19	141.17	141.18	141.18	141.18	-
# Covered bases <sup>b</sup>	442.90	444.13	440.64	436.60	446.12	444.74	442.49	444.50	-
% Genome covered <sup>b</sup>	0.52	0.52	0.52	0.51	0.52	0.52	0.52	0.52	-
# Combined variants	98	113	86	89	111	107	96	109	-
Individuals versus reference ( $\ll 1973$ until last sexual reproduction)									
# Covered bases (million) <sup>b</sup>	654.78	673.32	475.15	567.36	669.09	684.24	665.69	678.00	564.77
% Genome covered <sup>b</sup>	76.69	78.86	55.65	66.45	78.37	80.14	77.97	79.41	66.15
# Variants (million)	4.60	4.69	3.38	3.98	4.67	5.21	4.64	4.71	4.75
# Variants in CDS (million)	0.18	0.18	0.14	0.16	0.18	0.18	0.18	0.18	0.16

<sup>a</sup>Uniquely mapping reads after removal of duplicates.

<sup>b</sup>(Coverage  $\geq 10$ ).

a  $d_N/d_S$  ratio of 0.49. However, the somatic variants that distinguish the examined buds from their mother and accumulated in the shortest examined timescale (2006–2018) (Fig. 1B; comparison of whole-animal data of buds vs. mother) showed rather a weak sign of positive selection ( $d_N/d_S = 1.06$ ).

### Many biological processes under selection are related to *Hydra's* regenerative capacity

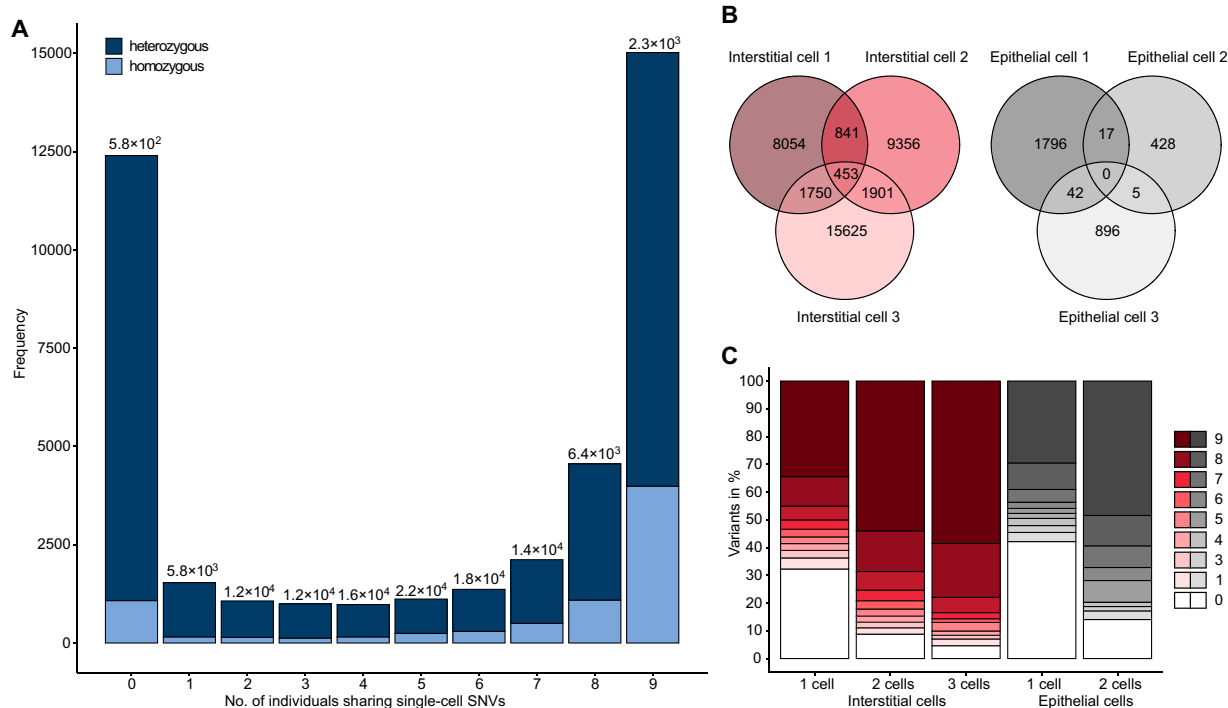
We investigated whether pathways and biological processes were adapted (our criterion for positive selection:  $d_N/d_S > 1$ ) or strongly conserved (our criterion for strong negative selection:  $d_N/d_S < 0.1$ ) during the examined evolutionary timescales. Of course, this analysis was limited to variants associated with coding sequences of ontologically annotated *Hydra* genes. Of the genes affected by the 505 SNVs identified on the shortest examined timescale (2006–2018) only 15 fulfilled these requirements, i.e., too few for any meaningful follow-up. Therefore, further analysis had to be restricted to the intermediate and longest timescales examined (Fig. 1).

Regarding somatic SNVs that accumulated most likely mainly in captivity (after the last sexual reproduction, i.e., probably shortly before 1973, until 2018), 28 biological processes showed signs of positive selection (Fig. 4A; Supplemental Table S4), whereas one process was found to be strongly conserved: “positive regulation of multicellular organismal process” (GO:0051240,  $d_N/d_S = 0.09$ ). For germline SNVs that originated in the wild (before the last sexual reproduction, i.e., mostly well before 1973), exactly the opposite pattern was observed: 274 processes display strong negative selection (Fig. 4B; Supplemental Table S5), while only one process, “heterotrimeric G-protein complex” (GO:0005834,  $d_N/d_S = 1.30$ ), was

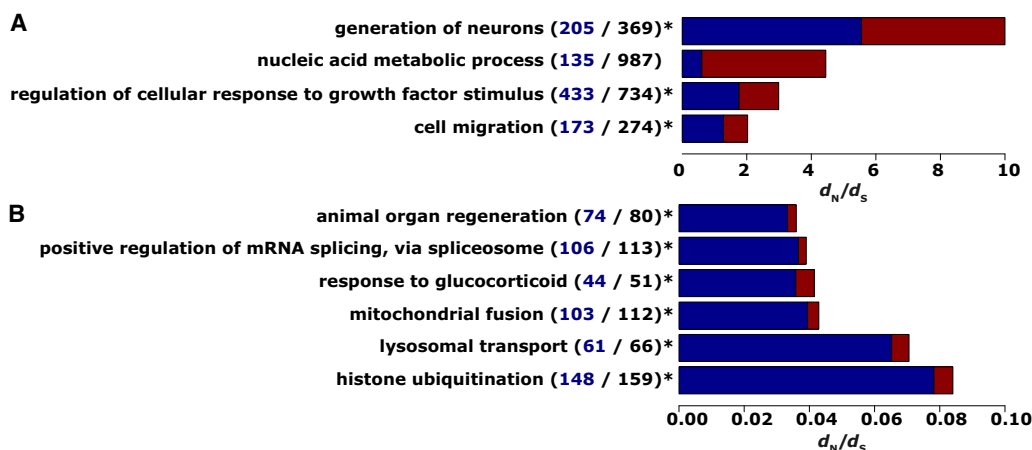
found to be adapted. In addition, two processes exhibited an enrichment for stop-gain mutations for the time in the wild (one-sided Fisher's exact test using a false discovery rate [FDR] threshold of 0.05) (Supplemental Table S8): the “nucleic acid metabolic process” (GO:0090304,  $FDR = 4.5 \times 10^{-7}$ ) and “hydrolase activity” (GO:0016787,  $FDR = 7.0 \times 10^{-5}$ ). Both processes were also identified as positively selected for the time in captivity. We also analyzed start-loss and stop-loss mutations. However, only a small fraction of genes was affected in comparison to nonsynonymous and stop-gain mutations (Supplemental Fig. S2). Thus, no corresponding enrichments for biological processes were found, neither for the time in captivity nor for the time in the wild.

For functional validation of these findings, we determined, for all processes and process categories identified as affected by positive or strong negative selection, how many of the genes involved in them are expressed under laboratory conditions. For this purpose, we sequenced the whole-animal mRNA of three individuals that we had previously examined genomically (Supplemental Table S1). The results showed that slightly more than half of these processes (168/304) (Supplemental Tables S4, S5) and all but one of the process categories formed from them (Fig. 4A,B) showed a significantly higher proportion of expressed genes than would be expected based on the total proportion of expressed genes (11,783/33,820; one-sided Fisher's exact test;  $FDR < 0.05$ ). In particular, the genes of the process categories affected by strong negative selection are expressed to a large extent (Fig. 4B).

For example, “animal organ regeneration” (GO:0031100) was found to be the strongest conserved biological process during the time in the wild: while this process was affected by in total 17 mutations during this time frame, only one of them changed an amino acid, resulting in a  $d_N/d_S$  ratio of only 0.018. At the same time 10



**Figure 3.** Indicators of positive selection among *Hydra* single-cell SNVs. (A) Distribution of hetero- and homozygous single-cell SNVs by their prevalence in the nine individuals. Numbers at the top of the columns represent the observed versus expected ratio of homozygous to heterozygous mutations based on the respective mutation frequency of heterozygous mutations in the genome (see Supplemental Table S7 for details). (B) Venn analysis of single-cell SNVs in I cells and E cells. No overlap was detected between I cells and E cells. (C) Prevalence of single-cell SNVs in 1–9 individuals examined.



**Figure 4.** Summary of processes affected (A) by positive selection during 44 years in captivity after last sexual reproduction and (B) by strong negative selection over a longer timescale before last sexual reproduction in the wild. The analysis of single-cell data/somatic SNVs (A) and the comparison of whole-animal data against the *Hydra* reference genome/germline SNVs (B) resulted in 28 and 274 biological processes affected by positive and strong negative selection, respectively (Supplemental Tables S4, S5). For this overview, processes were summarized by REVIGO (Supek et al. 2011). The total bar length gives the  $d_N/d_S$  value of the respective process; the blue/red part of the bar represents the fraction of expressed/not expressed genes under laboratory conditions. The values in brackets give the number of unique genes that were summarized under the respective term (bold black) and how many of these genes were found to be expressed (blue). An asterisk marks processes with significantly more expressed genes than expected based on the ratio of total expressed genes to total genes examined (11,783/33,820; one-sided Fisher’s exact test; FDR < 0.05). Strong negative selection was defined by  $d_N/d_S < 0.1$ .

out of 11 affected genes were expressed under laboratory conditions—significantly more than expected (FDR =  $9.15 \times 10^{-6}$ ) (Supplemental Table S5). The only two processes that were found to be even more strongly conserved were the “zona pellucida receptor complex” (GO:0002199), which is important for successful fertilization (Hinsch et al. 1997) and “establishment of neuroblast polarity” (GO:0045200). The former was found to be affected by 10 mutations for the time in the wild and the latter by six, none of which changed an amino acid.

The “generation of neurons” (GO:0048699) turned out to be the process most strongly affected by positive selection for the time during captivity with, in total, 11 mutations, all of which changed an amino acid. In addition, with 85 out of 156 genes belonging to this process, significantly more genes were expressed than expected (FDR =  $1.1 \times 10^{-5}$ ).

The second and third most positively selected processes for the time during captivity were “nucleic acid metabolic process” (GO:0090304, 30 out of 32 mutations nonsynonymous,  $d_N/d_S = 4.4$ ) and “growth factor binding” (GO:0019838, 12 out of 13 mutations nonsynonymous,  $d_N/d_S = 3.3$ ), respectively.

## Discussion

Our main finding—that the *H. magnipapillata* strain 105 does not exhibit a significantly lower somatic mutation rate than much more rapidly aging species—is an argument against the generalization of the somatic mutation theory of aging (Morley 1995). Conversely, it could be stated that a mutation rate as high as that of humans could be sufficient to enable life spans of centuries or even more. However, *Hydra* may also be seen as an exception among multicellular organisms, since it somewhat resembles a collection of unicellular organisms (Bosch et al. 2010). Thus, insights into *Hydra*’s biology of aging (Tomczyk et al. 2015) may not be directly transferable to most other animals, including mammals.

To validate the order of magnitude of the mutation rates determined, we examined our results for possible sources of bias: First, our mutation rates are based on an estimate of the number

of cell divisions, which assumes that the last sexual reproduction leading to the strain examined happened not long before cultivation of this strain in 1973 (see also Supplemental Discussion). However, even if this assumption was wrong, our analyses show that the last reproduction would have to have occurred decades or even centuries before the cultivation of the strain to change the order of magnitude of the estimated mutation rate (Fig. 2B). Since *Hydra* is known to usually switch to sexual reproduction when conditions are not within a narrow window of optimality, e.g., water temperatures  $\sim 19^\circ\text{C}$ , it seems highly unlikely that so much time would have elapsed without sexual reproduction in a naturally fluctuating environment (Martínez and Bridge 2012; Kaliszewicz and Lipińska 2013). Second, we analyzed the mutation spectrum for potential single-cell artifacts. Both the estimated fractions of potential artifacts (1.3%–3.5%) and differences to public human and mouse data (Fig. 3C; Supplemental Fig. S1) are relatively small, rendering any significant influence on the order of magnitude of our results unlikely. Third, the sequencing coverage is relatively evenly distributed across the genomes in our experiments. The MAPD metrics, comparing adjacent sequencing bins, range from 0.78 to 1.28, while only values  $> 2$  are seen as potentially problematic (Miller et al. 2022).

Another factor that could potentially lead to overestimated mutation rates is the older genome assembly used (*Hydra* 2.0) (Chapman et al. 2010). However, our approach ensures that an SNV is called only if the single cell that is examined deviates not only from the reference genome but also from the respective opposite stem cell lineage. Furthermore, if low reference genome quality led to false-positive SNV callings, one would expect those SNVs to be shared at least among the single cells of the same lineage. The proportion of common SNVs among the single cells examined is, however, clearly limited (Fig. 3B). Therefore, a significant influence of the genome assembly on the order of magnitude of the mutation rates appears unlikely. In contrast, the majority (69.9%) of SNVs identified by single-cell genome sequencing were also found by comparing the whole-animal genomic data to the reference genome—indicating the robustness of our

variation-calling approaches and their relative independence from possible biases (e.g., by genome amplification, sequencing chemistry, read coverage, or base/variant calling algorithm). Moreover, the similarity of somatic mutation spectra in *Hydra*, human, and mouse (Fig. 2C) argues against deleterious effects of the A/T richness of the polyp's genome. Admittedly, the highly varying mapping rates of 15%–80% (Table 1) are a potential problem. Note, however, that we considered for variant calling only genome positions with a read coverage  $\geq 20$ , both in the single cell examined as well as the combined mapping result of three single cells of the respective opposite stem cell lineage. Given this conservative approach, the number of bases covered by the respective comparisons correlated strongly with the mapping rates (22–106 million bases) (Table 1; see Supplemental Fig. S3 for overlaps in covered bases). Nevertheless, the robustness of our approach is supported by the fact that we found very similar mutation rates within each stem cell lineage, regardless of the mapping rate. For example, for E cell #2, a mapping rate/number of covered bases of only 15%/22 million resulted in a mutation rate per cell division of  $4.35 \times 10^{-9}$ . For E cell #1, a much higher mapping rate/number of covered bases of 63%/102 million resulted in a mutation rate of  $3.95 \times 10^{-9}$  (Fig. 2A; Table 1).

However, the overall lower mapping rates in E cells in comparison to I cells might still have contributed to an underestimation of the number of mutations in the former so that the real difference in the resulting mutation rate between the stem cell lineages could be lower than presented here. This could partly explain the surprising result that I cells with the potential to differentiate to germ cells display a higher mutation rate than E cells that lack this ability. Another explanation would be that I cells are more important for adapting to changing environmental conditions, e.g., due to their ability to differentiate into neurons and gland cells (see below).

Another interesting aspect is the contrasting pattern found with regard to the extent of positive and negative selection in the recent evolution of the examined *Hydra* strain with respect to the time in the wild and in captivity. In the wild, negative selection on the germline SNVs was found in all pathways examined except one; 274 processes were even found to be highly conserved ( $d_N/d_S < 0.1$ ). This was to be expected, since negative selection dominates in coding sequences—even more so if species are highly adapted to a rather constant environment (Yang 2005). For these reasons, it is not surprising that we found extremely few amino acid changes in processes that can be linked directly to highly evolved traits of *Hydra*. For example, the regenerative capacity of *Hydra* is among the most advanced in the animal kingdom (Bosch 2007; Vogt et al. 2019b). Therefore, changes in the corresponding genes are likely to decrease fitness. Similarly, the strong purifying selection of “positive regulation of multicellular organismal processes” (GO:0051240) under laboratory conditions reflects *Hydra*'s delicate balance of reconciling an extremely active stem cell community with the maintenance of the—albeit simple—architecture of the organism (Bosch et al. 2010). Histone ubiquitination, as another example, is a process well known to be strongly conserved throughout evolution from yeast to mammals (Cao and Yan 2012; Wang et al. 2017).

In stark contrast, the evolutionarily extremely short time frame of 44 years in the laboratory is characterized by significant adaptation of a number of biological processes, indicating that for somatic SNVs, positive selection became the prevailing evolutionary force after this severe change of environmental conditions. Fitness adjustments to changes in laboratory conditions are a phe-

nomenon already observed in bacteria, worms, and insects (Large et al. 2016; Hoffmann and Ross 2018; Knöppel et al. 2018). Our finding underlines *Hydra*'s adaptability to environmental changes as providing the basis for its sustainability over an evolutionary period of ~200 million years (Schwentner and Bosch 2015).

Most adapted processes (Fig. 4A) support the hypothesis that a significant proportion of selection in *Hydra* acts on the level of single stem cells rather than only at the level of individuals (Bosch et al. 2010). Following a drastic environmental change, e.g., from the wild to the laboratory, cell clones may increase their prevalence in the cell population of an individual by fitness gains based on positive selection of: (1) “nucleic acid metabolic process” to improve nucleic acids availability for continued cell proliferation (Gross and Rotwein 2016; Zhu and Thompson 2019), (2) “regulation of cellular response to growth factor stimulus” for faster cell cycle progression and higher division rates (Jones and Kazlauskas 2001; Fingar and Blenis 2004; Gross and Rotwein 2016), and (3) faster “cell migration” that is associated with differentiation (Hager and David 1997; Siebert et al. 2019).

The strongest sign for positive selection, summarized by “generation of neurons” (GO:0048699), however, points toward selection acting on the level of individuals, as neuronal integration is essential for the fitness of a multicellular organism. This process includes the production of neuroblasts and their differentiation into neurons. It has long been known that *Hydra* can feed on live animals but not on dead ones, suggesting that they receive from their prey certain signals necessary to initiate the feeding response (Beutler 1924), most likely via their ectodermal nerve net. As food supply and respective feeding response change after the transition from the wild to captivity, it seems plausible that relevant neurophysiological processes also undergo adaptation.

In the comparison of the mother against each of its eight daughters, we found only very small differences. Moreover, the rather similar numbers of 86–113 SNVs identified displayed no correlation with the time of budding (Table 2). This suggests that these variants appeared almost exclusively due to the founder effect, i.e., a shift of allele frequencies as a consequence of random sampling of cells during budding. This would mean that, at least during this late phase of captivity (the final 12 [27%] of the total of 44 years), positive selection did not increase allele frequencies of beneficial alleles to an extent that would be detectable when analyzing an entire animal within the limits of the sequencing resources available to us. In line with the hypothesis that these SNVs were predominantly a result of the founder effect, the rather weak sign of positive selection that was observed ( $d_N/d_S = 1.06$ ) corresponds almost perfectly to neutral selection ( $d_N/d_S \sim 1$ ). Conversely, this means that there is also no evidence for the assumption (discussed by Schaible et al. [2015]) that in order to prevent accumulation of damage in the mother, *Hydra* cells damaged by mutations may be transmitted via budding to the offspring.

In conclusion, the presented first insights into the evolutionary forces that shape the genome of the nonsenescent *H. magnipapillata* strain 105 excluded that its exceptional longevity under benign laboratory conditions is based on a low mutation rate. On the contrary, somatic mutation rates higher than in humans or mice provide the genetic variability underlying *Hydra*'s high adaptability. It remains to be tested whether also other extremely long-lived species deviate from the predictions of the somatic mutation theory of aging. Additional research is needed to decipher the mechanisms behind *Hydra*'s nonsenescence as an adapted, i.e., not brought about by natural selection, exceptional metazoan phenotype.

## Methods

To ensure comparability of results, our methodological approach at the single-cell level almost completely replicated that used to determine human and mouse mutation rates (Milholland et al. 2017; see [Supplemental Table S6](#) for a direct methodical comparison). Briefly, individuals were examined ~24 h after their last feeding. For the isolation of three stem-like E cells and three stem-like I cells (likely a mixture of ecto- and endodermal cells, as we did not differentiate between the two), the first head and foot were dissected. Dissociation of the remaining body column to single-cell suspensions and subsequent classification of cells was then conducted by an established protocol using a maceration solution containing glycerin, glacial acetic acid, and water (David 1973). After isolation of each single cell with a pipette under the microscope at 40× magnification, the remaining glacial acetic acid was evaporated at 60°C for 2 h. The GE Healthcare illustra Single Cell GenomiPhi Kit was used for cell lysis and multiple displacement amplification. Sequencing was performed using Illumina sequencing methodology (Bentley et al. 2008). In detail, 2 µg of amplified DNA was introduced into Illumina's TruSeq DNA PCR-free library preparation. Paired-end sequencing was conducted using Illumina HiSeq 2500 (high-throughput mode, 2 × 101 cycles) and NovaSeq 6000 (Xp workflow, 2 × 251 cycles) devices. Data were obtained using Illumina's bcl2fastq v1.8.4 (HiSeq 2500) and v2.20.0.422 (NovaSeq 6000). Trim Galore! was used for read trimming (<https://github.com/FelixKrueger/TrimGalore>) and BWA-MEM for mapping (Li and Durbin 2009). Duplicate reads were removed using SAMtools rmdup (Li et al. 2009). GATK was used for indel realignment and base quality recalibration (Van der Auwera et al. 2013), and VarScan 2 (Koboldt et al. 2012), MuTect (Cibulskis et al. 2013), and UnifiedGenotyper (Van der Auwera et al. 2013) for variant calling. For each of the 6 cells examined, MuTect and VarScan 2 were invoked such that the cell currently under consideration served as the somatic/tumor input, and the union set of 3 cells of the respective other cell type served as the normal/bulk input. The same caller settings and filter criteria as in Milholland et al. (2017) were used. Only variants predicted by all three callers were considered for further analysis. As in Milholland et al., only positions with a read coverage ≥20 were taken into account for variant calling. Thus, the number of mutations per base was calculated by dividing the number of variant positions by the total number of bases covered by ≥20 reads.

To validate our results further, we implemented a pipeline for single-cell artifact minimization, as described recently in the literature (Miller et al. 2022, Extended Fig. 1G). Briefly, it comprises three steps: (1) calculation of MAPD score as an indicator for potential uneven genome coverage; (2) examination of the known potential single-cell artifact signature SBS scE; and (3) examination of the combined known potential single-cell artifact signatures SBS scE and SBS scF (Petljak et al. 2019). The latter signatures were analyzed together with the COSMIC v3 signatures using the R package sigfit (Gori and Baez-Ortega 2018).

For the estimation of the mutation rate per cell division, the number of mitoses since the time of the zygote was assumed to be the sum of developing mitoses and continuous regeneration mitoses. For developing mitoses, we used  $\log_2(10^5) = 17$ . For continuous regeneration mitoses, given that I cells and E cells are known to divide every 24–30 h and 3–4 days, respectively (David and Campbell 1972; Campbell and David 1974), we used  $44 \times \frac{365}{1.25} = 14,276$  and  $44 \times \frac{365}{3.5} = 4589$ , respectively.

For examining whole *Hydra* individuals, the Genomic DNA Isolation Kit from Norgen Biotek was used for DNA extraction, the NEBNext Ultra II DNA Kit for library preparation (DNA input

10 ng), and an Illumina HiSeq 2500 device for paired-end sequencing (2 × 101 cycles). Data were extracted using Illumina's bcl2fastq v2.20.0.422. Read trimming was conducted with cutadapt (Martin 2011) and Trimmomatic (Bolger et al. 2014), read mapping with BWA-MEM (Li and Durbin 2009), removal of duplicates, indel realignment, and base quality recalibration with GATK (Van der Auwera et al. 2013), and clipped second mate overlaps using *bamUtil* (Jun et al. 2015). Regarding the comparison of the buds versus their mother, for all buds, mapped reads were additionally downsampled to the lowest number of mapped reads among buds before variant calling. Variants were accepted if they were predicted by at least three of the following six callers: BCFtools (Li 2011), FreeBayes (<https://github.com/freebayes/freebayes>), Mutect2 (Van der Auwera et al. 2013), Platypus (Rimmer et al. 2014), VarDict (Lai et al. 2016), and VarScan 2 (Koboldt et al. 2012). For the comparisons of each individual versus reference, “germline” variants were kept that were flagged accordingly. For the comparison of the buds versus their mother, only “somatic” variants were kept that were flagged as somatic by Mutect2, VarScan 2, and VarDict. For the remaining callers, BCFtools FreeBayes, and Platypus, filters were adapted from the Bcbio-next-gen pipeline. Finally, only those somatic variants were kept for downstream analyses that either could be validated in the independent single-cell experiment analysis or had an FDR below 0.05 (based on a one-sided Fisher's exact test). For the multiple test correction, we used the respective number of germline variants ([Supplemental Tables S1, S3](#)). The entire pipeline used for this workflow is available at GitHub (<https://github.com/Hoffmann-Lab/muvac>) (see also [Supplemental Code](#)).

To determine  $d_N/d_S$  ratios, we first categorized the subset of SNVs hitting coding sequences into nonsynonymous and synonymous substitutions using the Hydra 2.0 annotation (Chapman et al. 2010) and a custom script ([Supplemental Code](#)). Using another custom script, we counted the number of potentially nonsynonymous and synonymous sites as described by the literature (Nei and Gojobori 1986). The rate of nonsynonymous substitutions  $d_N$  was then estimated by dividing the total number of nonsynonymous differences observed by the total number of potential nonsynonymous sites and correcting for backmutations using the Jukes–Cantor formula. The same was done for  $d_S$  and synonymous differences/sites, respectively (Nei and Gojobori 1986). The entire procedure described was applied either across all protein-coding genes to determine genome-wide  $d_N/d_S$  ratios or across all genes belonging to a certain GO category to determine GO-category-specific  $d_N/d_S$  ratios ([Supplemental Tables S4, S5](#)). We did not determine  $d_N/d_S$  ratios for single genes, as almost all genes were, if at all, affected by low numbers of variants, rendering  $d_N/d_S$  estimations unreliable.

For the functional annotation of Gene Ontology analyses the Blast2GO (Conesa et al. 2005) resource from the Hydra 2.0 Genome Project Portal was used ([https://research.nhgri.nih.gov/hydra/download/functional\\_annotation/Blast2GO/blast2go\\_results.tsv](https://research.nhgri.nih.gov/hydra/download/functional_annotation/Blast2GO/blast2go_results.tsv)).

To analyze gene expression data of whole *Hydra* individuals, the Norgen Biotek Total RNA Purification Kit was used for RNA extraction. The library preparation was performed with the NEBNext Kits Poly(A) mRNA Magnetic Isolation Module, Ultra II directional RNA Multiplex Oligos for Illumina, using 50–100 ng of total RNA. For sequencing, an Illumina HiSeq 2500 device in paired-end mode was used (51 cycles). Read trimming was conducted with cutadapt (Martin 2011) and Trimmomatic (Bolger et al. 2014), read mapping with segemehl (Hoffmann et al. 2014), and quantification with featureCounts (Liao et al. 2014). We assessed all genes as expressed that were found with a median transcript per million value of at least 5.

To examine whether a GO category X was enriched for expressed genes, we used a one-sided Fisher's exact test with the following input: number of expressed genes belonging to X, number of nonexpressed genes belonging to X, number of expressed genes not belonging to X, and number of nonexpressed genes not belonging to X. To compare the percentage of homozygosity in single-cell SNVs found in all nine individuals versus those in none of the individuals using a one-sided Fisher's exact test (Fig. 3A), the following input data were used: number of homozygous SNVs found in all nine individuals, number of heterozygous SNVs found in all nine individuals, number of homozygous SNVs found in none of the individuals, and number of heterozygous individuals found in none of the individuals. To examine whether a GO category X was enriched for genes with stop-gain/start-loss/stop-loss mutations, we used a one-sided Fisher's exact test with the following input: number of genes belonging to X with stop-gain/start-loss/stop-loss mutations, number of genes belonging to X without stop-gain/start-loss/stop-loss mutations, number of genes not belonging to X with stop-gain/start-loss/stop-loss mutations, and number of genes not belonging to X without stop-gain/start-loss/stop-loss mutations. All read mapping was carried out against the *Hydra vulgaris* 2.0 genome (Chapman et al. 2010).

## Data access

The raw data generated in this study have been submitted to the European Nucleotide Archive (ENA; <https://www.ebi.ac.uk/ena/browser/home>) under accession number PRJEB50344. The variants identified are available as a browser-accessible track hub at [http://genome.ucsc.edu/cgi-bin/hgTracks?genome=Hm105\\_Dovetail\\_Assembly\\_1.0&hubUrl=https://genome.leibniz-fli.de/paras/hydra/hub.txt](http://genome.ucsc.edu/cgi-bin/hgTracks?genome=Hm105_Dovetail_Assembly_1.0&hubUrl=https://genome.leibniz-fli.de/paras/hydra/hub.txt).

## Competing interest statement

The authors declare no competing interests.

## Acknowledgments

We thank Ivonne Görlich, Cornelia Luge, and especially Silke Förste for excellent technical assistance during this project. We thank Penelope Krumm and Debra Weih for proofreading and editing the manuscript. This work was supported by the Joachim Herz Foundation (Add-on Fellowships for Interdisciplinary Life Science, to A.S.).

**Author contributions:** M.P., S.H., R.S., C.E., and A.S. conceived the project. R.S. provided the animals and isolated the cells. M.G., M.F., and M.B. conducted library preparation and sequencing. A.S., K.R., M.P., J.K., and H.K. analyzed the data. M.P. and A.S. wrote the initial manuscript draft. All authors revised the draft and approved the final manuscript.

## References

Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, et al. 2008. Accurate whole human genome sequencing using reversible terminator chemistry. *Nature* **456**: 53–59. doi:10.1038/nature07517

Beutler R. 1924. Experimentelle Untersuchungen über die Verdauung bei hydra. *Z. f. vergl. Physiologie* **1**: 1–56. doi:10.1007/BF00338209

Boehm AM, Khalturin K, Anton-Erxleben F, Hemmrich G, Klostermeier UC, Lopez-Quintero JA, Oberg HH, Puchert M, Rosenstiel P, Wittlieb J, et al. 2012. FoxO is a critical regulator of stem cell maintenance in immortal *Hydra*. *Proc Natl Acad Sci* **109**: 19697–19702. doi:10.1073/pnas.1209714109

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**: 2114–2120. doi:10.1093/bioinformatics/btu170

Bosch TCG. 2009. *Hydra* and the evolution of stem cells. *Bioessays* **31**: 478–486. doi:10.1002/bies.200800183

Bosch TCG. 2007. Why polyps regenerate and we don't: towards a cellular and molecular framework for *Hydra* regeneration. *Dev Biol* **303**: 421–433. doi:10.1016/j.ydbio.2006.12.012

Bosch TCG, Anton-Erxleben F, Hemmrich G, Khalturin K. 2010. The *Hydra* polyp: nothing but an active stem cell community. *Dev Growth Differ* **52**: 15–25. doi:10.1111/j.1440-169X.2009.01143.x

Cagan A, Baez-Ortega A, Brzozowska N, Abascal F, Coorens THH, Sanders MA, Lawson ARJ, Harvey LMR, Bhosle S, Jones D, et al. 2022. Somatic mutation rates scale with lifespan across mammals. *Nature* **604**: 517–524. doi:10.1038/s41586-022-04618-z

Campbell RD. 1967. Tissue dynamics of steady state growth in *Hydra littoralis*. II. Patterns of tissue movement. *J Morphol* **121**: 19–28. doi:10.1002/jmor.1051210103

Campbell RD, David CN. 1974. Cell cycle kinetics and development of *Hydra attenuata*. II. Interstitial cells. *J Cell Sci* **16**: 349–358. doi:10.1242/jcs.16.2.349r

Cao J, Yan Q. 2012. Histone ubiquitination and deubiquitination in transcription, DNA damage response, and cancer. *Front Oncol* **2**: 26. doi:10.3389/fonc.2012.00026

Chapman JA, Kirkness EF, Simakov O, Hampson SE, Mitros T, Weinmaier T, Rattei T, Balasubramanian PG, Borman J, Busam D, et al. 2010. The dynamic genome of *Hydra*. *Nature* **464**: 592–596. doi:10.1038/nature08830

Cibulskis K, Lawrence MS, Carter SL, Sivachenko A, Jaffe D, Sougnez C, Gabriel S, Meyerson M, Lander ES, Getz G. 2013. Sensitive detection of somatic point mutations in impure and heterogeneous cancer samples. *Nat Biotechnol* **31**: 213–219. doi:10.1038/nbt.2514

Conesa A, Götz S, García-Gómez JM, Terol J, Talón M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* **21**: 3674–3676. doi:10.1093/bioinformatics/bti610

Dańko MJ, Kozłowski J, Schaible R. 2015. Unraveling the non-senescence phenomenon in *Hydra*. *J Theor Biol* **382**: 137–149. doi:10.1016/j.jtbi.2015.06.043

David CN. 1973. A quantitative method for maceration of hydra tissue. *Wilhelm Roux Arch Entwickl Mech Org* **171**: 259–268. doi:10.1007/BF00577724

David CN, Campbell RD. 1972. Cell cycle kinetics and development of *Hydra attenuata*. I. Epithelial cells. *J Cell Sci* **11**: 557–568. doi:10.1242/jcs.11.2.557

Domazet-Lošo T, Klimovich A, Anokhin B, Anton-Erxleben F, Hamm MJ, Lange C, Bosch TCG. 2014. Naturally occurring tumours in the basal metazoan *Hydra*. *Nat Commun* **5**: 4222. doi:10.1038/ncomms5222

Dong X, Zhang L, Milholland B, Lee M, Maslov AY, Wang T, Vijg J. 2017. Accurate identification of single-nucleotide variants in whole-genome-amplified single cells. *Nat Methods* **14**: 491–493. doi:10.1038/nmeth.4227

Fingar DC, Blenis J. 2004. Target of rapamycin (TOR): an integrator of nutrient and growth factor signals and coordinator of cell growth and cell cycle progression. *Oncogene* **23**: 3151–3171. doi:10.1038/sj.onc.1207542

Fujisawa T. 2003. *Hydra* regeneration and epitheliotopes. *Dev Dyn* **226**: 182–189. doi:10.1002/dvdy.10221

Gori K, Baez-Ortega A. 2018. sigfit: flexible Bayesian inference of mutational signatures. bioRxiv doi:10.1101/372896

Gould S, Vrba ES. 1982. Exaptation—a missing term in the science of form. *Paleobiology* **8**: 4–15. doi:10.1017/S0094837300004310

Gross SM, Rotwein P. 2016. Unraveling growth factor signaling and cell cycle progression in individual fibroblasts. *J Biol Chem* **291**: 14628–14638. doi:10.1074/jbc.M116.734194

Hager G, David CN. 1997. Pattern of differentiated nerve cells in hydra is determined by precursor migration. *Development* **124**: 569–576. doi:10.1242/dev.124.2.569

Hemmrich G, Khalturin K, Boehm A-M, Puchert M, Anton-Erxleben F, Wittlieb J, Klostermeier UC, Rosenstiel P, Oberg H-H, Domazet-Lošo T, et al. 2012. Molecular signatures of the three stem cell lineages in *Hydra* and the emergence of stem cell function at the base of multicellularity. *Mol Biol Evol* **29**: 3267–3280. doi:10.1093/molbev/mss134

Hinsch E, Hägele W, Schill W-B, Hinsch K-D. 1997. The zona pellucida “receptors”. In *The fate of the male germ cell* (ed. Ivell R, Holstein A-F), pp. 313–328. Springer, Boston, MA.

Hoffmann AA, Ross PA. 2018. Rates and patterns of laboratory adaptation in (mostly) insects. *J Econ Entomol* **111**: 501–509. doi:10.1093/jeet/toy024

Hoffmann S, Otto C, Doose G, Tanzer A, Langenberger D, Christ S, Kunz M, Holdt LM, Teupser D, Hackermüller J, et al. 2014. A multi-split mapping algorithm for circular RNA, splicing, trans-splicing and fusion detection. *Genome Biol* **15**: R34. doi:10.1186/gb-2014-15-2-r34

- Holstein TW, David CN. 1990. Cell cycle length, cell size, and proliferation rate in hydra stem cells. *Dev Biol* **142**: 392–400. doi:10.1016/0012-1606(90)90360-U
- Holtze S, Gorshkova E, Braude S, Cellerino A, Dammann P, Hildebrandt TB, Hoeflich A, Hoffmann S, Koch P, Terzibasí Tozzini E, et al. 2021. Alternative animal models of aging research. *Front Mol Biosci* **8**: 660959. doi:10.3389/fmolb.2021.660959
- Jones SM, Kazlauskas A. 2001. Growth factor-dependent signaling and cell cycle progression. *FEBS Lett* **490**: 110–116. doi:10.1016/S0014-5793(01)02113-5
- Jones OR, Scheuerlein A, Salguero-Gómez R, Camarda CG, Schaible R, Casper BB, Dahlgren JP, Ehlén J, García MB, Menges ES, et al. 2014. Diversity of ageing across the tree of life. *Nature* **505**: 169–173. doi:10.1038/nature12789
- Jun G, Wing MK, Abecasis GR, Kang HM. 2015. An efficient and scalable analysis framework for variant extraction and refinement from population scale DNA sequence data. *Genome Res* **25**: 918–925. doi:10.1101/gr.176552.114
- Kaliszewicz A, Lipińska A. 2013. Environmental condition related reproductive strategies and sex ratio in hydras. *Acta Zool* **94**: 177–183. doi:10.1111/j.1463-6395.2011.00536.x
- Kimura M. 1980. A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences. *J Mol Evol* **16**: 111–120. doi:10.1007/BF01731581
- Klimovich A, Rehm A, Wittlieb J, Herbst E-M, Benavente R, Bosch TCG. 2018. Non-senescent *Hydra* tolerates severe disturbances in the nuclear lamina. *Aging* **10**: 951–972. doi:10.18632/aging.101440
- Knöppel A, Knopp M, Albrecht LM, Lundin E, Lustig U, Näsvall J, Andersson DI. 2018. Genetic adaptation to growth under laboratory conditions in *Escherichia coli* and *Salmonella enterica*. *Front Microbiol* **9**: 756. doi:10.3389/fmicb.2018.00756
- Koboldt DC, Zhang Q, Larson DE, Shen D, McLellan MD, Lin L, Miller CA, Mardis ER, Ding L, Wilson RK. 2012. VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res* **22**: 568–576. doi:10.1101/gr.129684.111
- Krishna S, Nair A, Cheedipudi S, Poduval D, Dhawan J, Palakodeti D, Ghanekar Y. 2013. Deep sequencing reveals unique small RNA repertoire that is regulated during head regeneration in *Hydra magnipapillata*. *Nucleic Acids Res* **41**: 599–616. doi:10.1093/nar/gks1020
- Lai Z, Markovets A, Ahdesmaki M, Chapman B, Hofmann O, McEwen R, Johnson J, Dougherty B, Barrett JC, Dry JR. 2016. VarDict: a novel and versatile variant caller for next-generation sequencing in cancer research. *Nucleic Acids Res* **44**: e108. doi:10.1093/nar/gkw227
- Large EE, Xu W, Zhao Y, Brady SC, Long L, Butcher RA, Andersen EC, McGrath PT. 2016. Selection on a subunit of the NURF chromatin remodeler modifies life history traits in a domesticated strain of *Caenorhabditis elegans*. *PLoS Genet* **12**: e1006219. doi:10.1371/journal.pgen.1006219
- Li H. 2011. A statistical framework for SNP calling, mutation discovery, association mapping and population genetical parameter estimation from sequencing data. *Bioinformatics* **27**: 2987–2993. doi:10.1093/bioinformatics/btr509
- Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**: 1754–1760. doi:10.1093/bioinformatics/btp324
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup. 2009. The Sequence Alignment/Map format and SAMtools. *Bioinformatics* **25**: 2078–2079. doi:10.1093/bioinformatics/btp352
- Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* **30**: 923–930. doi:10.1093/bioinformatics/btt656
- López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2013. The hallmarks of aging. *Cell* **153**: 1194–1217. doi:10.1016/j.cell.2013.05.039
- Maroudas-Sacks Y, Garion L, Shani-Zerbib L, Livshits A, Braun E, Keren K. 2021. Topological defects in the nematic order of actin fibres as organization centres of Hydra morphogenesis. *Nat Phys* **17**: 251–259. doi:10.1038/s41567-020-01083-1
- Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet.J* **17**: 10–12. doi:10.14806/ej.17.1.200
- Martínez DE. 1998. Mortality patterns suggest lack of senescence in hydra. *Exp Gerontol* **33**: 217–225. doi:10.1016/S0531-5565(97)00113-7
- Martínez DE, Bridge D. 2012. *Hydra*, the everlasting embryo, confronts aging. *Int J Dev Biol* **56**: 479–487. doi:10.1387/ijdb.113461dm
- Michael PS. 2019. Hydra as a unique model for the study of regenerative mechanisms in metazoans. *MOJ Anat Physiol* **6**: 151–153. doi:10.15406/mojap.2019.06.00264
- Milholland B, Dong X, Zhang L, Hao X, Suh Y, Vijg J. 2017. Differences between germline and somatic mutation rates in humans and mice. *Nat Commun* **8**: 15183. doi:10.1038/ncomms15183
- Miller MB, Huang AY, Kim J, Zhou Z, Kirkham SL, Maury EA, Ziegenfuss JS, Reed HC, Neil JE, Rento L, et al. 2022. Somatic genomic changes in single Alzheimer's disease neurons. *Nature* **604**: 714–722. doi:10.1038/s41586-022-04640-1
- Morley AA. 1995. The somatic mutation theory of ageing. *Mutat Res* **338**: 19–23. doi:10.1016/0921-8734(95)00007-5
- Müller WA. 1996. Pattern formation in the immortal *Hydra*. *Trends Genet* **12**: 91–96. doi:10.1016/0168-9525(96)81419-3
- Nei M, Gojobori T. 1986. Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Mol Biol Evol* **3**: 418–426.
- Petersen HO, Höger SK, Looso M, Lengfeld T, Kuhn A, Warnken U, Nishimiya-Fujisawa C, Schnölzer M, Krüger M, Özbek S, et al. 2015. A comprehensive transcriptomic and proteomic analysis of Hydra head regeneration. *Mol Biol Evol* **32**: 1928–1947. doi:10.1093/molbev/msv079
- Petljak M, Alexandrov LB, Brammeld JS, Price S, Wedge DC, Grossmann S, Dawson KJ, Ju YS, Iorio F, Tubio JMC, et al. 2019. Characterizing mutational signatures in human cancer cell lines reveals episodic APOBEC mutagenesis. *Cell* **176**: 1282–1294.e20. doi:10.1016/j.cell.2019.02.012
- Rimmer A, Phan H, Mathieson I, Iqbal Z, Twigg SRF, Wilkie AOM, McVean G, Lunter G, Consortium WGS. 2014. Integrating mapping, assembly and haplotype-based approaches for calling variants in clinical sequencing applications. *Nat Genet* **46**: 912–918. doi:10.1038/ng.3036
- Schaible R, Sussman M, Kramer BH. 2014. Aging and potential for self-renewal: *Hydra* living in the age of aging - a mini-review. *Gerontology* **60**: 548–556. doi:10.1159/000360397
- Schaible R, Scheuerlein A, Daňko MJ, Gampe J, Martínez DE, Vaupel JW. 2015. Instant mortality and fertility over age in *Hydra*. *Proc Natl Acad Sci* **112**: 15701–15706. doi:10.1073/pnas.1521002112
- Schaible R, Ringelhan F, Kramer BH, Scheuerlein A. 2017. *Hydra*: evolutionary and biological mechanisms for non-senescence. In *The evolution of senescence in the tree of life* (ed. Shefferson RP, et al.), Part II: Senescence in animals, Chap. 12, pp. 238–253. Cambridge University Press, Cambridge. doi:10.1017/9781139939867.012
- Schwentner M, Bosch TC. 2015. Revisiting the age, evolutionary history and species level diversity of the genus *Hydra* (Cnidaria: Hydrozoa). *Mol Phylogenet Evol* **91**: 41–55. doi:10.1016/j.ympev.2015.05.013
- Shimizu H, Sawada Y, Sugiyama T. 1993. Minimum tissue size required for hydra regeneration. *Dev Biol* **155**: 287–296. doi:10.1006/dbio.1993.1028
- Siebert S, Farrell JA, Cazet JF, Abeykoon Y, Primack AS, Schnitzler CE, Juliano CE. 2019. Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science* **365**: eaav9314. doi:10.1126/science.aav9314
- Stewart EJ, Madden R, Paul G, Taddei F. 2005. Aging and death in an organism that reproduces by morphologically symmetric division. *PLoS Biol* **3**: e45. doi:10.1371/journal.pbio.0030045
- Sturm A, Perczel A, Ivics Z, Vellai T. 2017. The Piwi-piRNA pathway: road to immortality. *Aging Cell* **16**: 906–911. doi:10.1111/acel.12630
- Sugiyama T, Fujisawa T. 1977. Genetic analysis of developmental mechanisms in hydra I. Sexual reproduction of *Hydra magnipapillata* and isolation of mutants. *Dev Growth Differ* **19**: 187–200. doi:10.1111/j.1440-169X.1977.00187.x
- Sun S, White RR, Fischer KE, Zhang Z, Austad SN, Vijg J. 2020. Inducible aging in *Hydra oligactis* implicates sexual reproduction, loss of stem cells, and genome maintenance as major pathways. *Geroscience* **42**: 1119–1132. doi:10.1007/s11357-020-00214-z
- Supek F, Bošnjak M, Škunca N, Šmuc T. 2011. REVIGO summarizes and visualizes long lists of gene ontology terms. *PLoS One* **6**: e21800. doi:10.1371/journal.pone.0021800
- Tomczyk S, Fischer K, Austad S, Galliot B. 2015. *Hydra*, a powerful model for aging studies. *Invertebr Reprod Dev* **59**: 11–16. doi:10.1080/07924259.2014.927805
- Tomczyk S, Suknovic N, Schenkelaars Q, Wenger Y, Ekundayo K, Buzgariu W, Bauer C, Fischer K, Austad S, Galliot B. 2020. Deficient autophagy in epithelial stem cells drives aging in the freshwater cnidarian *Hydra*. *Development* **147**: dev177840. doi:10.1242/dev.177840
- Van der Auwera GA, Carneiro MO, Hartl C, Poplin R, Del Angel G, Levy-Moonshine A, Jordan T, Shakir K, Roazen D, Thibault J, et al. 2013. From FastQ data to high confidence variant calls: the Genome Analysis Toolkit best practices pipeline. *Curr Protoc Bioinformatics* **43**: 11.10.11–11.10.33. doi:10.1002/0471250953.bi1110s43
- Vijg J. 2021. From DNA damage to mutations: all roads lead to aging. *Ageing Res Rev* **68**: 101316. doi:10.1016/j.arr.2021.101316
- Vogg MC, Beccari L, Iglesias Olle L, Rampon C, Vriz S, Perruchoud C, Wenger Y, Galliot B. 2019a. An evolutionarily-conserved Wnt3/β-catenin/Sp5 feedback loop restricts head organizer activity in *Hydra*. *Nat Commun* **10**: 312. doi:10.1038/s41467-018-08242-2
- Vogg MC, Galliot B, Tsiaris CD. 2019b. Model systems for regeneration: *Hydra*. *Development* **146**: dev177212. doi:10.1242/dev.177212

- Wang P, Robert L, Pelletier J, Dang WL, Taddei F, Wright A, Jun S. 2010. Robust growth of *Escherichia coli*. *Curr Biol* **20**: 1099–1103. doi:10.1016/j.cub.2010.04.045
- Wang L, Cao C, Wang F, Zhao J, Li W. 2017. H2b ubiquitination: conserved molecular mechanism, diverse physiologic functions of the E3 ligase during meiosis. *Nucleus* **8**: 461–468. doi:10.1080/19491034.2017.1330237
- Watanabe H, Schmidt HA, Kuhn A, Höger SK, Kocagöz Y, Laumann-Lipp N, Özbek S, Holstein TW. 2014. Nodal signalling determines biradial asymmetry in *Hydra*. *Nature* **515**: 112–115. doi:10.1038/nature13666
- Wenger Y, Galliot B. 2013. RNAseq versus genome-predicted transcriptomes: a large population of novel transcripts identified in an Illumina-454 *Hydra* transcriptome. *BMC Genomics* **14**: 204. doi:10.1186/1471-2164-14-204
- Wolf AM. 2021. The tumor suppression theory of aging. *Mech Ageing Dev* **200**: 111583. doi:10.1016/j.mad.2021.111583
- Yang Z. 2005. The power of phylogenetic comparison in revealing protein function. *Proc Natl Acad Sci* **102**: 3179–3180. doi:10.1073/pnas.0500371102
- Yousefzadeh M, Henpita C, Vyas R, Soto-Palma C, Robbins P, Niedernhofer L. 2021. DNA damage—how and why we age? *eLife* **10**: e62852. doi:10.7554/eLife.62852
- Zhu J, Thompson CB. 2019. Metabolic regulation of cell growth and proliferation. *Nat Rev Mol Cell Biol* **20**: 436–450. doi:10.1038/s41580-019-0123-5

Received January 23, 2024; accepted in revised form October 18, 2024.