Transgenerational chromosome repair in the asexual bdelloid rotifer *Adineta vaga*

Antoine Houtain[*1,2], Alessandro Derzelle[†1,2], Marc Llirós[†3,1,4], Boris Hespeels[1], Émilien Nicolas[2,4], Paul Simion[5,1], Julie Virgo[1], Anne-Catherine Heuskin[6,7], Thomas Lenormand[8], Bernard Hallet[*‡4] and Karine Van Doninck[*‡2]

[1]LEGE, URBE, Université de Namur, Namur, 5000, Belgium
[2]Molecular Biology and Evolution (MBE), Université libre de Bruxelles (ULB), Brussels, 1050, Belgium
[3]Present address: Bioinformatics and Bioimaging research group, Biosciences Department, Faculty of Sciences, Technology and Engineerings, Universitat de Vic – Universitat Central de Catalunya, Vic, Catalunya, 08500, Spain
[4]LIBST, Croix du Sud 4/5, Université Catholique de Louvain (UCLouvain), Louvain-la-Neuve, 1348, Belgium
[5]Present address: EcoBio - Ecosystems, Biodiversity, Evolution, Université de Rennes 1, 35042 Rennes, France
[6]Namur Research Institute for Life Sciences (NARILIS)
[7]Laboratory of Analysis By Nuclear Reactions (LARN), University of Namur, Namur, Belgium
[8]CEFE, Univ Montpellier, CNRS, Univ Paul Valéry Montpellier 3, EPHE, IRD, Montpellier, France

January 25, 2024

[*]Corresponding authors:
antoine.houtain@gmail.com
bernard.hallet@uclouvain.be
karine.van.doninck@ulb.be

[†]These authors contributed equally to this work
[‡]These authors contributed equally to this work

Abstract

Homologous recombination plays a fundamental role in the evolution of organisms. It serves as a DNA repair mechanism which, in sexual organisms, contributes to genetic diversity through the shuffling of alleles during meiosis. Here we investigate the two functions of homologous recombination in the bdelloid rotifer *Adineta vaga*, an ancient asexual species also known for its tolerance to extreme genotoxic stresses. Genomic analyses reveal that *A. vaga* retained meiotic recombination mechanisms, both for DNA repair and occurrence of spontaneous crossovers during oogenesis. Our study introduces a novel transgenerational DNA repair mechanism termed break-induced homologous extension repair (BIHER). BIHER operates on single DNA ends, enabling the repair of fragmented chromosomes. Our findings suggest that the BIHER mechanism, combined with a holocentric structure of chromosomes and a modified meiosis, constitutes a key adaptation for life in extreme environments. Identifying such a mechanism in bdelloid rotifers sheds a new light on the strategies that evolved to maintain genome structure in asexually reproducing species.

Keywords— Parthenogenesis, experimental evolution, chromosome fragmentation, DNA repair, transgenerational chromosome repair, genome evolution, loss of heterozygosity, holocentric chromosomes

Main

Homologous recombination (HR) is a universal genetic mechanism that operates as a high-fidelity, template-dependent repair mechanism of DNA double-strand breaks (DSBs). In addition, during sexual reproduction, HR generates variability by shuffling alleles through crossovers (CO)[1]. This genetic mixing limits selective interference among loci, improving the response to selection in sexual species[2–4].

In this study, we investigated whether meiotic recombination serves both functions of DNA repair and genetic mixing within the unique context of an asexual, extremo-tolerant non-model organism, the bdelloid rotifer *Adineta vaga*[5–8]. Our recent work unveiled the genome structure and the oogenesis process of *A. vaga*, showing six pairs of homologous chromosomes that pair during a non-reductional, abortive meiosis I, producing an unreduced egg[9,10]. Throughout oogenesis, germline DNA repair activity was detected in response to irradiation[10], suggesting the potential occurrence of meiotic HR for chromosome repair. Furthermore, natural populations of bdelloid rotifers showed several indirect genetic signatures of recombination supporting a role in genome evolution[9,11–13].

Here, we detected genomic changes, including direct signatures of HR accumulating over generations, under hydrated conditions, and following cycles of desiccation in *A. vaga* in the course of a "micro-evolution" (ME) experiment for over 75 generations (Supp. Fig. 1). The repair function of HR was examined in a "genome repair" (GR) experiment following acute DNA damage induced by proton irradiation (Supp. Fig. 2). In both experiments, selection was reduced by collecting a single descendant at each generation.

Meiotic recombination in the germline

The ancestral laboratory strain of *A. vaga* has a diploid genome structure containing 6 pairs of homologous chromosomes with an average heterozygosity of 1.7%[9]. In the ME experiment, a total of 22 lines were sequenced at the endpoint (i.e.: after $\sim$ 80 or $\sim$ 130 generations for desiccated and hydrated lines, respectively). Eleven lines remained hydrated throughout, while another 11 lines experienced 4 cycles of 10-days of desiccation (Supp. Fig. 1). Among these 22 lines, 14 were sequenced at the midpoint of the experiment (after $\sim$40 and $\sim$70 generations, respectively). For both hydrated and desiccated lines, the average heterozygosity (H) decreased with the number of generations, with an average loss of 4.23% at endpoint (Fig. 1**a**). However, there is a significant difference in LOH accumulation rate between the endpoints of both

conditions (Fig. 1**b**; Mann-Whitney U rank test, p = 0.03; CLES=0.22; RBC=0.55), showing that desiccated samples accumulated LOH faster than hydrated samples.

LOH clusters were detected in all ME lines (Fig. 1**c**). Clusters with fewer than 5 contiguous sites were ignored, as they could arise from independent LOH events rather than from a genetic event such as gene conversion or a reciprocal exchange (Supp. Fig. 3). Most clusters (85.2%) are smaller than 100 contiguous sites, but cover only 14.2% of all LOH sites in ME. Clusters larger than 100 consecutive sites are less frequent (14.8%) but encompass most LOH 85.8%. Among the largest LOH clusters (>100 sites), 19.4% are potentially deletions (Δ coverage below -20), covering 11.4% of all LOH sites in ME (Supp. Fig. 4).

In all samples, allele frequency (AF) and coverage were more variable towards the end of the chromosomes (representative example in Supp. Fig. 5) correlating with higher repetitive content[9]. To address signal noise in LOH detection across chromosome-scale regions (e.g. heterozygosity signal due to mismappings in repetitive content and variability in the reference genome), a LOH score was calculated along the chromosomes by computing the AF difference between the ancestor and the offspring samples using a window size of 50 to 100Kb corresponding to 1/250 of the chromosome length. A LOH score close to 1.0 results from an effective LOH event, while a LOH score around 0.0 indicates no change (Supp. Methods, Supp. Fig. 3**b**). All ME lines showed LOH tracts ranging from *ca.* 100 bp to several Mb. The longer tracts (>50 Kb) often corresponded to multiple LOH clusters (Fig. 1**c**) interrupted by repetitive content. Some of the LOH tracts extended to the chromosome ends, while others are in the middle. Examples of such LOH tracts are presented in Fig. 2**a**, where a LOH tract of *ca.* 3 Mb (1) is found at the extremity of chromosome 6 in sample D5C3, and another of *ca.* 115 Kb (2) is in the middle. Smaller LOH tracts (<50 Kb) were also detected (examples in Supp. Fig. 6), as exemplified by tract (3) in Fig. 2**b**. An illustrative depiction of the potential occurrence of various types of LOH tracts in eggs resulting from a non-reductional meiosis I is provided in 2**c**. Long LOH tracts, ranging from 50 Kb to over 1 Mb, were identified in 7 out of 11 hydrated lines and in 3 out of 11 desiccated lines. Long LOH tracts extending to the chromosome tips resulted from a single crossover (CO) with segregation of only one recombinant chromatid (case 1 in Fig. 2**c**). Other long LOH tracts, in the middle of chromosomes, likely arose from double COs (case 2 in Fig. 2**c**). Shorter tracts (from 100 bp to 50 Kb) were found in all ME lines, and those of up to 2 Kb probably resulted from gene conversions (GC). GCs are produced by a CO followed by the segregation of both recombinant chromatids, or after HR resolution without a CO (e.g.: case 3 in Fig. 2**c**).

All long LOH tracts (>50 Kb) identified are present at both the midpoint and endpoint of the ME experiment, with two exceptions in samples D2B3 and D2C1 where a LOH is found at the midpoint but absent at the endpoint (Supp. Fig. 7). In those cases, HR likely generated the LOH tracts in the culture used for the midpoint sequencing after its separation from the rest of the line.

LOH tracts provide evidence that DSBs, programmed or not, occur frequently and induce HR in hydrated *A. vaga* lines subjected to low selection pressure. The slightly higher LOH rate detected under desiccated conditions could be attributed to a higher frequency of DSBs in these lines[14].

Taken together, these results strongly support the recent finding that *A. vaga* oocytes undergo non-reductional meiosis I during which homologous chromosomes pair to produce crossovers[10].

Inheritance of deletions and duplications

Besides signatures of reciprocal exchanges between homologous chromosomes, the genomes of ME lines exhibit regions where changes in the LOH score is associated with variations in coverage compared to the ancestor (Supp. Fig. 3**b**; Supp. Table 1). LOH tracts with a decrease in coverage correspond to deletions (Fig. 3**a-b**), while tracts where coverage increases correspond to duplications (Fig. 3**c-d**). There are 11 large (>50 Kb) duplications and 16 large (>50 Kb) deletions in the chromosomes of 13 (8 hydrated and 5 desiccated) of the 22 sequenced ME lines. The distribution of these large events in ME is not random, as it appears to accumulate towards the end of the chromosomes and as samples that present recombination events have more events than would be expected if these events occurred independently (Chi2 = 112.8; p = 5.02e-19) (Fig. 4; Supp. Fig. 8).

Deletions can be homozygous when both parental alleles are absent (example in Fig. 3**a**) or hemizygous if one allele is absent (example in Fig. 3**b**; Expected patterns in Supp. Fig. 3). Among the large deletions (>50 Kb) observed in the ME lines (17), the majority are at chromosomes ends (12) and are homozygous (8), with a few terminal deletions being hemizygous. The remaining deletions (5) are found in the middle of chromosomes and are hemizygous (Fig. 4) (summary based on Supp. Fig. 7).

A coverage increase indicates the presence of additional copies of chromosome portions (e.g.: Fig. 3**c**). Because some homologous chromosomes of the ancestral genome are of unequal length, in one case in ME lines, a coverage increase indicates a restoration of diploidy caused

by the duplication of the longer homolog of chromosome 1 (Fig. 3**d**). The other duplications (13 in total) represent additional copies of a genomic region (Fig. 4; Supp. Fig. 7).

Restorations (1/13) and homozygous deletions (8/17) likely result from a single CO between homologous chromosomes of different sizes. Whether the result is a deletion or a restoration is determined by the subsequent segregation of chromatids (Fig. 3**e**). The mechanism behind the other deletions and duplications remains unclear, one possibility being that they are caused by unequal crossovers[15] (Fig. 3**f**). However, some deletions have patterns of AF and coverage that are not expected after an unequal CO. For example, one large deletion has a LOH tract that extends past the coverage drop tract, as if partially repaired (Supp. Fig. 9**a-b**). Other deletions display a progressive drop of coverage over dozens of Kb (Supp. Fig. 9**c-d**) where a breakpoint would be expected (e.g.: Supp. Fig. 9**e-f**). In one particular case, a deletion had a progressive coverage decrease on one side and a breakpoint on the other side (Supp. Fig. 9**g**).

Large duplications spanning several Mb were identified in four ME lines: H2B4, H4A4, H5A2 and D2C1 (Fig. 4), consistently at chromosome ends. For instance, in H4A4, chromosome 1 appears to be fully duplicated, while another duplication of approximately 1.8 Mb is present at the end of chromosome 6. These duplications of H4A4 were detected at both the 61 and 118 generations timepoints (Supp. Figs. 7; 10**a-b**), indicating their early appearance in the ME experiment and transmission across generations. At 118 generations (endpoint), chromosome 1 exhibits further duplications of one portion, with higher Δ coverage at the endpoint than the midpoint (Supp. Fig. 10**a**). Similarly, the duplication on chromosome 6 has expanded by roughly 500 Kb at endpoint compared to the midpoint indicating that duplications accumulated over generations (Supp. Fig. 10**b**).

In most cases, duplications and deletions are probably due to unequal crossovers between repetitive sequences. However, the mapping of short reads prevents determining whether the duplications are actually in tandem or independent fragments. In one particular case, shifts in coverage and AF (Supp. Fig. 10**c**) correspond to a ratio of 5:4 copies, resulting in a coverage level of 1.25 and an AF of 0.6 (tetraploid), rather than the expected values for diploid (0.5) or triploid (0.66). A single duplication of chromosome 1 (triploidy) cannot explain these observations. Furthermore, the observed AF shifts do not span the entire chromosome. This result suggests that line H4A4 underwent a whole genome duplication (WGD) followed by a duplication of the shorter homolog of chromosome 1.

An additional copy of a chromosome suggests that aneuploidy can be transmitted through generations in *A. vaga*, which is consistent with a parthenogenetic mode of reproduction by a

non-reductional meiosis I.

Radiation-induced large deletions

In the ME experiment, short-term (10 days) desiccation events slightly increased the rate of LOH in *A. vaga*, suggesting that desiccation-induced DNA damage could enhance HR activity by triggering DSB repair in the germline. To explore whether HR mechanisms play a role in maintaining genome integrity under more severe DNA damaging conditions, cultures were subjected to longer periods of desiccation (30 days) or irradiated with increasing doses of proton IR (Supp. Fig. 2).

The distributions of LOH clusters in the non-irradiated (0 Gy) and non-desiccated (hydrated) samples consist of short clusters with fewer than 100 contiguous sites, distributed around a Δ coverage of 0 (Fig. 1**d**). As discussed earlier, these tracts are likely the result of GC events accompanying HR in the germline. Their presence in untreated samples confirms that recombination between homologous chromosomes is constitutive and occurs spontaneously, independently of external stressors. The higher prevalence of GC signatures in the ME lines compared to the GR samples simply reflects their accumulation over a greater number of generations (Fig. 1**c-d**). Compared to the hydrated sample (3,169 LOH sites), progeny from the 30-day desiccated cultures exhibited a slightly more extended distribution (4,955 LOH sites), with two LOH clusters exceeding 100 contiguous sites (Fig. 1**d**). This suggests that the level of DNA damage and DSBs induced during 1 month of desiccation (estimated at $\sim$15–20 DSBs per genome[14]) marginally increased the level of HR already present under hydrated conditions. However, irradiation with 100, 250 and 500 Gy of proton (corresponding to $\sim$88, 220, and 440 DSBs per genome, respectively[14]) resulted in a dose-dependent amplification of both the number and size of LOH clusters relative to the non-irradiated sample (Fig. 1**d** and Supp. Fig. 4). In particular, the 500 Gy IR exposure produced long LOH stretches spanning up to 400 Kb (Fig. 4; Supp. Fig. 11).

A significant proportion of both the short and long LOH tracts induced by IR is distributed around a low coverage (Δ coverage below -20; Fig. 1**d**). Specifically, in the 100, 250 and 500 Gy IR conditions, 42.1%, 37.5% and 31.24%, respectively, of the total number of LOH clusters were potential deletions, while in the non-irradiated control and the ME lines (Supp. Fig. 4) deletions represented only 15.85% and 20.7% of LOH tracts, respectively. This observed pattern of increased deletion signatures likely reflects the choice of DNA repair pathways. Indeed, single DSBs could be repaired by non-homologous end joining (NHEJ) or single-strand

annealing (SSA), resulting in DNA losses at the repair site[16]. When two distant DSBs occur on the same chromosome, the ligation of the distal breaks could result in a larger deletion spanning several Kb. Some LOH tracts with Δ coverage higher than -20, also found in the 500 Gy samples (Fig. 4), may similarly arise from two distant DSBs. In such case, HR could promote the exchange of the intervening chromosomal segment, with no change in coverage, akin to a double CO as observed in the ME lines. However, these LOH tracts also display AF variability, which is not expected after a double CO (Supp. Fig. 11).

In the 500 Gy samples, most large clusters (>100 LOH sites) correspond to hemizygous deletions, where large chromosomal regions (up to 400 Kb) are lost. Unlike the ME lines where deletions tended to be at chromosome ends, IR-induced deletions are scattered along the chromosomes, reflecting the extent of DNA fragmentation and the stochasticity of IR-induced damage (Fig. 4; Supp. Fig. 11). The increased DNA fragmentation induced by proton radiation in one generation could be linked to a decrease in growth rate of irradiated cultures, with the largest effect observed at 500 Gy IR (cf. Supp. Results; Supp. Fig 12). Similarly, the decrease in fertility is likely caused by this increased genomic fragmentation following the higher IR doses.

Hemizygous deletions, resulting from the loss of a chromosome fragment between two DSBs, are expected to present a characteristic breakpoint with an abrupt change in coverage from 1.0 to 0.5, and an AF increase from 0.5 to 1.0 due to the loss of a parental allele (Supp. Fig. 3). However, the deletions found in the progeny of 500 Gy irradiated individuals present a different pattern, with gradual variations of AF and coverage at the boundaries of the corresponding LOH tract (Fig. 5**a**; Supp. Fig. 13). In all cases, AF progressively increases, stabilizing at 1.0, while coverage gradually decreases to various extents towards the center of the deletion. The maximum change in coverage was from $\sim$0.5 to $\sim$0.85, depending on deletion size, suggesting a mechanism of progressive DNA repair acting from the borders.

Since these observations correspond to the state of the offspring populations at the time of Illumina sequencing (i.e., after about one month of growth), the 500 Gy irradiated P500 C18 E3 clone (which contains 8 large intrachromosomal deletions) was cultured for an additional six months, and its genome was re-sequenced using long-read Nanopore technology. All deletions have the same AF profile as before the extended growth (with a shift to a plateau of 0.9 instead of 1.0 caused by the higher rate of sequencing errors in Nanopore reads). In contrast, the coverage level has increased from the LOH tract borders until reaching a value of $\sim$1.0 for the shortest deletions, showing a complete recovery of some of the deleted fragments

(Fig. 5**a,c**; Supp. Fig. 13). This result is consistent with a dynamic process of deletion lengthening and shortening, producing heterogeneous offspring in the culture (Fig. 5**b**). Partial degradation of the chromosomes from the initial breakpoints of deletions introduced variability in LOH tract size, contributing to the progressive AF shifts observed. Besides degradation, the progressive recovery of coverage suggests that a transgenerational DNA repair mechanism gradually replaced the missing genomic fragment using the homologous chromosome as a template (Fig. 5**b**).

Remarkably, the gain in coverage after an extended 6-month of culture is consistent across all intrachromosomal deletions (approximately 20%, Fig. 5**b**). The fitting of a mathematical model of diffusion to the coverage data yielded estimates for the rates of chromosome degradation and repair (Θ_1 and Θ_2, respectively) for all deletions separately. The estimated parameters were very similar across deletions, supporting the hypothesis that a common repair mechanism is acting on the different deletions (Supp. Fig. 14).

Transmission and repair of broken chromosomes over generations

The above scenario implies the inheritance of broken chromosomes from one generation to the next to allow DNA repair to continue. Consistent with this hypothesis, the long Nanopore reads from the P500 C18 E3 sample formed a continuous alignment across the entire deleted segments. None of them matched with reduced deletion intermediates in which the distal parts of the fragments were joined, indicating that the borders of the deletion were not resealed during the repair process (Fig. 5**d**). Supporting this assumption, a simulation of 50 deletion intermediates for the P500 C18 E3 deletion (of Figure 5**a**) revealed that reads containing re-ligated DNA sequences should have been detected in the alignment if they were indeed present in the actual dataset (Fig. 5**c**; Supp. Fig. 15). This result shows that IR-induced DNA breakage generated fragmented chromosomes, and that the resulting fragments remained unsealed during successive steps of transgenerational repair.

To confirm the inheritance of broken chromosomes, single-cell eggs from descendants of the P500 C18 E3 clone were arrested in metaphase of the first division, and their karyotype was compared to that of the parental strain (Fig. 6). Eggs collected from the control line displayed an average of 12 DAPI-stained bodies, consistent with previous karyotyping experiments and genome assemblies showing that *A. vaga* has six pairs of homologous chromosomes[9]. In contrast, the average number of chromosomal entities identified by independent experimenters in the P500 C18 E3 progeny varied between ∼12 and ∼20 (Fig. 6). This result provides

conclusive evidence that broken chromosomes can be maintained and transmitted over multiple generations in *A. vaga* until they become fully reconstituted by recombinational repair. The variable karyotype established for different eggs aligns with the genomic heterogeneity observed in the P500 C18 E3 line, where large chromosomal deletions were partially or totally resolved at the time of sequencing (Fig. 5**b**; Supp. Fig. 13).

The observed stable inheritance of fragmented chromosomes during both meiotic and mitotic cell divisions in *A. vaga* is difficult to reconcile with classical mechanisms of monocentric chromosome segregation in eukaryotes. Typically, acentric chromosome fragments, lacking the centromere, are rapidly lost as they do not attach to the mitotic spindle and cannot separate properly in daughter cells. Instead, we propose that *A. vaga* chromosomes have a holocentric structure with kinetochores dispersed along their axis, enabling isolated fragments to segregate independently[17]. This proposition aligns with previous Hi-C (High throughput Chromosome Conformation Capture) sequencing data, showing the absence of centromeric signatures in the contact maps of *A. vaga* genome[9,18].

Discussion

Previous studies established the diploid nature of *A. vaga*, highlighting meiotic chromosome pairing, recombination signatures, and a delayed DNA repair activity in the germline[9,10]. Our current study highlights the role of meiotic HR in generating LOH at both small scales (gene conversions) and large scales (crossovers). This process contributes to an increased homozygosity, irrespective of DNA damage induction. However, after over 75 generations, LOH accumulated slowly in *A. vaga* (on average 3.24e-04 per bp per generation in hydrated lines) and remained constrained (4.23% on average and at most 13.2%), despite experimental conditions minimizing selection.

Our results also support a novel DNA repair mechanism called break-induced homologous extension repair (BIHER) (Fig. 5**e**). This mechanism is inspired from break-induced replication (BIR), which primarily acts to restart stalled replication forks during vegetative growth and telomere maintenance in certain forms of cancer[19]. We propose that BIHER specifically operates between paired homologous chromosomes during the modified meiosis I of *A. vaga*, coinciding with DNA synthesis signatures that were detected in maturing oocytes following DNA damage induction[10]. BIHER initiates at one-end DSBs (half DSBs) resulting from genome fragmentation and chromosome end degradation. The process then proceeds by DNA strand

invasion into the template, followed by DNA synthesis, elongating the broken chromosome. Because of limited processivity, elongation stops before reaching the opposite border of large deletions, and the complementary lagging strand is synthesized asynchronously. This process occurs at both deletion ends producing extended chromosome fragments which are inherited through mitosis of somatic cell lines if the chromosomes are holocentric, until they reach meiosis I of the next generation where the BIHER mechanism can restart. Over generations, BIHER-mediated elongation is faster than chromosome degradation, leading to a reduction in deletion size and subsequent coverage recovery, as observed for large deletions in the GR experiment on *A. vaga*. Eventually, the extended chromosome fragments overlap and become able to fuse back by recombination.

BIHER may also explain patterns of AF variation and coverage observed in some deletions found in ME lines. These patterns could not be explained by unequal CO but could represent various stages of repair by BIHER (Supp. Fig. 9**a-g**).

Similarly, in the GR experiment, successful BIHER explains the few cases of LOH tracts displaying AF variability with a Δ coverage close to 0 (Supp. Fig. 16). These signatures correspond to the pattern observed in small, fully repaired deletions from Nanopore sequencing of sample P500 C18 E3 (Supp. Fig. 13**a,f**). This implies that these LOH tracts actually originated from deletions and were effectively repaired by BIHER prior to the Illumina sequencing.

Intra-individual meiotic HR, as observed here in *A. vaga*, may play a dual role by mitigating the accumulation of deleterious alleles[20] but also introducing diversity by switching alleles and creating mutation hotspots[21]. HR can also cause copy-number and structural variations, such as deletions and duplications. Indeed, *A. vaga* displays a high genomic plasticity, tolerating these variants but also fragmented chromosomes, aneuploidy and potentially whole genome duplications. The combination of a non-reductional meiosis I, genomic plasticity, transgenerational repair by BIHER, and efficient DNA repair proteins[8,22] contribute to a more in-depth understanding of the remarkable genomic resilience of *A. vaga*, and likely, bdelloids in general. Our results also support the hypothesis that holocentric chromosomes evolved as an adaptive response to survive increased DNA damage, particularly in extreme environments[23]. More generally, our research highlights how chromosomal features and reproductive modes may contribute to genome plasticity, fostering diversification and adaptability. We propose that similar mechanisms may extend beyond bdelloid rotifers to species with comparable genomic and reproductive characteristics. By preventing catastrophic chromosome rearrangements, BIHER may serve as a pivotal process enabling asexual organisms, such as bdelloids, to adapt in

challenging environments while maintaining genome structure. Future research should explore the consequences of HR and BIHER on population dynamics, considering larger population sizes and stronger selection pressures in the wild.

References

1. Zickler, D. & Kleckner, N. Recombination, pairing, and synapsis of homologs during meiosis. *Cold Spring Harb. Perspect. Biol.* **7,** a016626 (2015).
2. Otto, S. P. The evolutionary enigma of sex. *Am. Nat.* **174,** S1–S14 (2009).
3. Glemin, S. & Galtier, N. in *Evolutionary Genomics* 311–335 (Springer, 2012).
4. Alberts, B. *Molecular biology of the cell* (WW Norton & Company, 2017).
5. Gladyshev, E. & Meselson, M. Extreme resistance of bdelloid rotifers to ionizing radiation. *Proc. Natl. Acad. Sci.* **105,** 5139–5144 (2008).
6. Krisko, A., Leroy, M., Radman, M. & Meselson, M. Extreme anti-oxidant protection against ionizing radiation in bdelloid rotifers. *Proc. Natl. Acad. Sci.* **109,** 2354–2357 (2012).
7. Hespeels, B. *et al.* Iron Ladies–How Desiccated Asexual Rotifer Adineta vaga Deal With X-Rays and Heavy Ions? *Front. Microbiol.* **11,** 1792 (2020).
8. Moris, V. C. *et al.* Ionizing radiation responses appear incidental to desiccation responses in the bdelloid rotifer Adineta vaga. *BMC Biology* **22,** 11 (2024).
9. Simion, P. *et al.* Chromosome-level genome assembly reveals homologous chromosomes and recombination in asexual rotifer Adineta vaga. *Sci. Adv.* **7,** eabg4216 (2021).
10. Terwagne, M. *et al.* DNA repair during nonreductional meiosis in the asexual rotifer Adineta vaga. *Sci. Adv.* **8,** eadc8829 (2022).
11. Debortoli, N. *et al.* Genetic exchange among bdelloid rotifers is more likely due to horizontal gene transfer than to meiotic sex. *Curr. Biol.* **26,** 723–732 (2016).
12. Vakhrusheva, O. A. *et al.* Genomic signatures of recombination in a natural population of the bdelloid rotifer Adineta vaga. *Nat. Commun.* **11,** 1–17 (2020).
13. Laine, V. N., Sackton, T. B. & Meselson, M. Genomic signature of sexual reproduction in the bdelloid rotifer Macrotrachella quadricornifera. *Genetics* **220** (2021).

14. Hespeels, B. *et al.* Gateway to genetic exchange? DNA double-strand breaks in the bdelloid rotifer Adineta vaga submitted to desiccation. *J. Evol. Biol.* **27,** 1334–1345 (2014).

15. Smith, G. P. Evolution of Repeated DNA Sequences by Unequal Crossover: DNA whose sequence is not maintained by selection will develop periodicities as a result of random crossover. *Science* **191,** 528–535 (1976).

16. Ceccaldi, R., Rondinelli, B. & D'Andrea, A. D. Repair pathway choices and consequences at the double-strand break. *Trends Cell Biol.* **26,** 52–64 (2016).

17. Senaratne, A. P., Cortes-Silva, N. & Drinnenberg, I. A. *Evolution of holocentric chromosomes: drivers, diversity, and deterrents* in *Semin. Cell. Dev. Biol.* **127** (2022), 90–99.

18. Muller, H., Gil, J. & Drinnenberg, I. A. The impact of centromeres on spatial genome architecture. *Trends Genet.* **35,** 565–578 (2019).

19. Liu, L. & Malkova, A. Break-induced replication: unraveling each step. *Trends Genet.* **38,** 752–765 (2022).

20. Engelstädter, J. Asexual but not clonal: evolutionary processes in automictic populations. *Genetics* **206,** 993–1009 (2017).

21. Hinch, R., Donnelly, P. & Hinch, A. G. Meiotic DNA breaks drive multifaceted mutagenesis in the human germ line. *Science* **382,** eadh2531 (2023).

22. Nicolas, E. *et al.* Horizontal acquisition of a DNA ligase improves DNA damage tolerance in eukaryotes. *Nat. Commun.* **14,** 7638 (2023).

23. Mandrioli, M. & Manicardi, G. C. Holocentric chromosomes. *PLoS Genet.* **16,** e1008918 (2020).

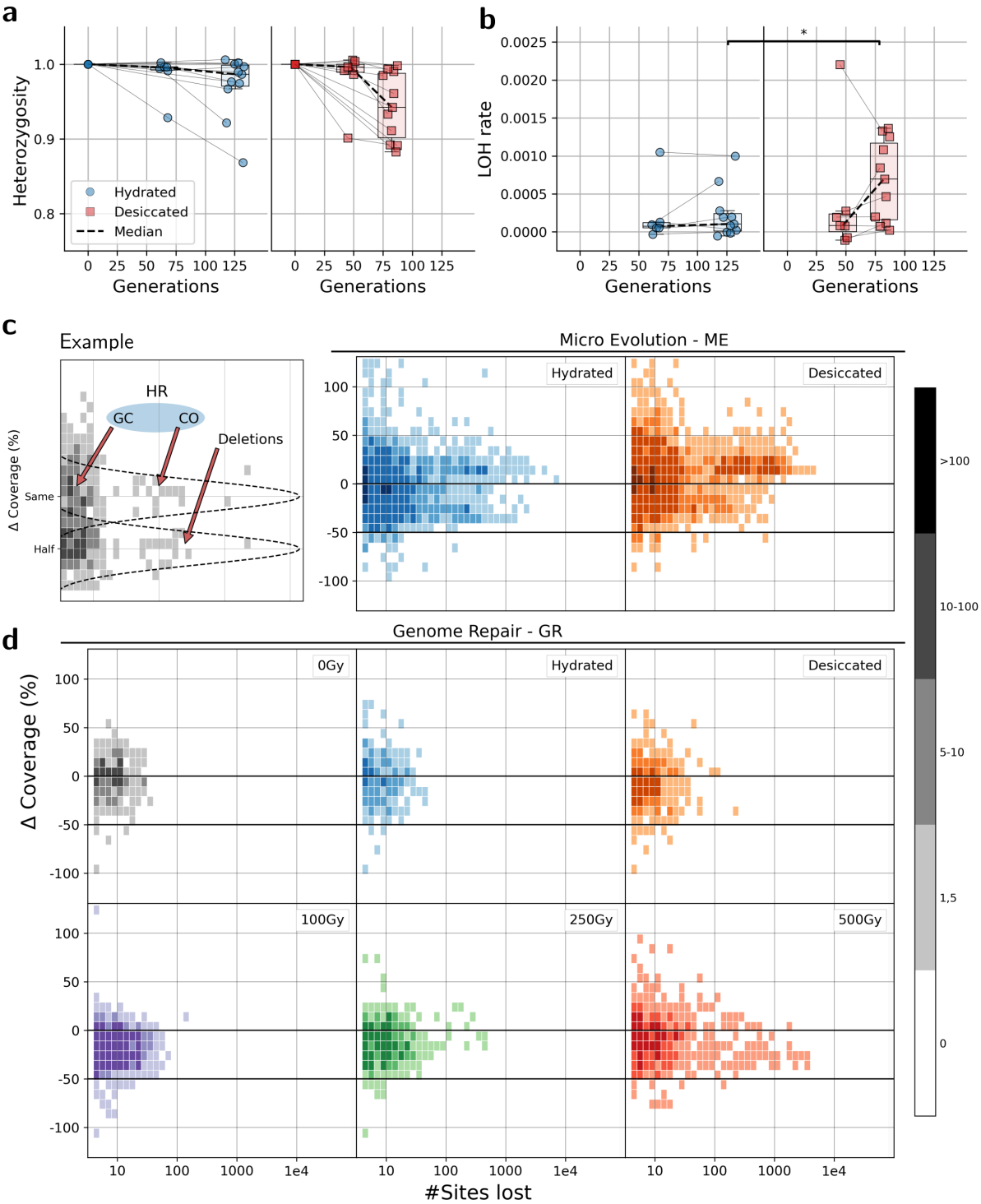

Figure 1. LOH accumulation in ME and GR. a: Heterozygosity over generations in ME samples. **b**: LOH rate over generations in ME samples (Mann-Whitney U rank test; p=0.03; CLES=0.22; RBC=0.55). **c-d**: Histograms of LOH clusters size represented by the number of heterozygous sites losses (logged $x-$axis) frequency (color, $z-$axis) and $\Delta\%$ coverage ($y-$axis) per experimental condition in ME (**c**) and in GR (**d**). An example shows where LOH due to deletions and RE are expected on the histogram.

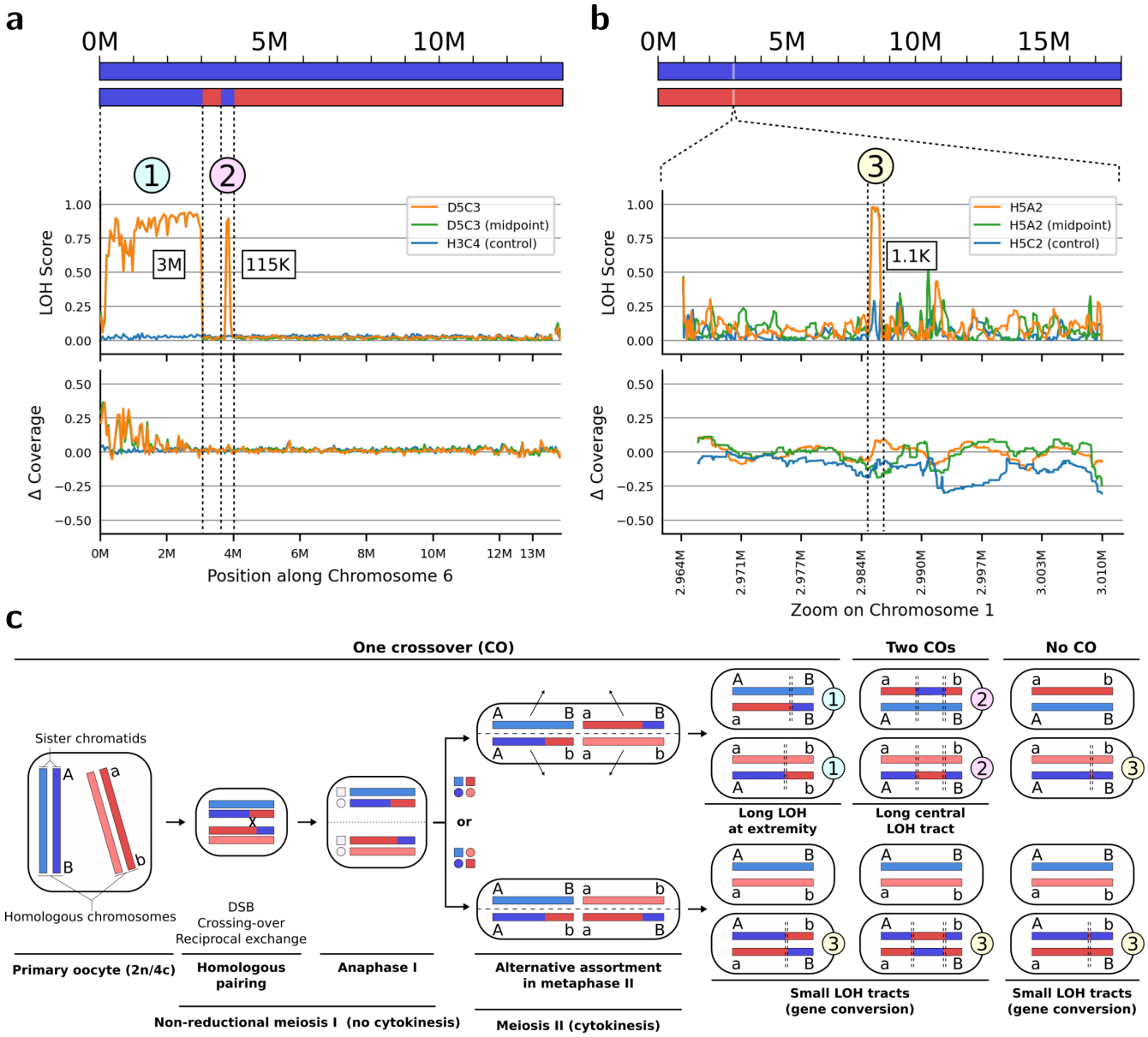

Figure 2. Examples of LOH tracts due to HR in ME. a: LOH score and Δ coverage along chromosome 6 of ME sample D5C3 (midpoint and endpoint) showing examples of LOH tracts resulting from a reciprocal exchange after one CO (1) or two COs (2). ME sample, H3C4 is shown as a negative control, without LOH at the considered positions. **b**: LOH score and Δ coverage along a small portion of chromosome 1 of ME sample H5A2 (midpoint and endpoint) showing an example of a LOH tract probably produced by GC (3). ME sample, H5C2 is shown as a negative control, without LOH at the considered positions. **c**: Description of oogenesis by non-reductional meiosis I and expected LOH tracts following homologous recombination with reciprocal exchange or gene conversion. Different cases are shown with one, two or no crossovers. Some cases of COs followed by segregation produce a gene conversion and a reciprocal exchange (e.g. the case 3 with only one crossover) but only the LOH tract due to GC can be detected and not the exchange.

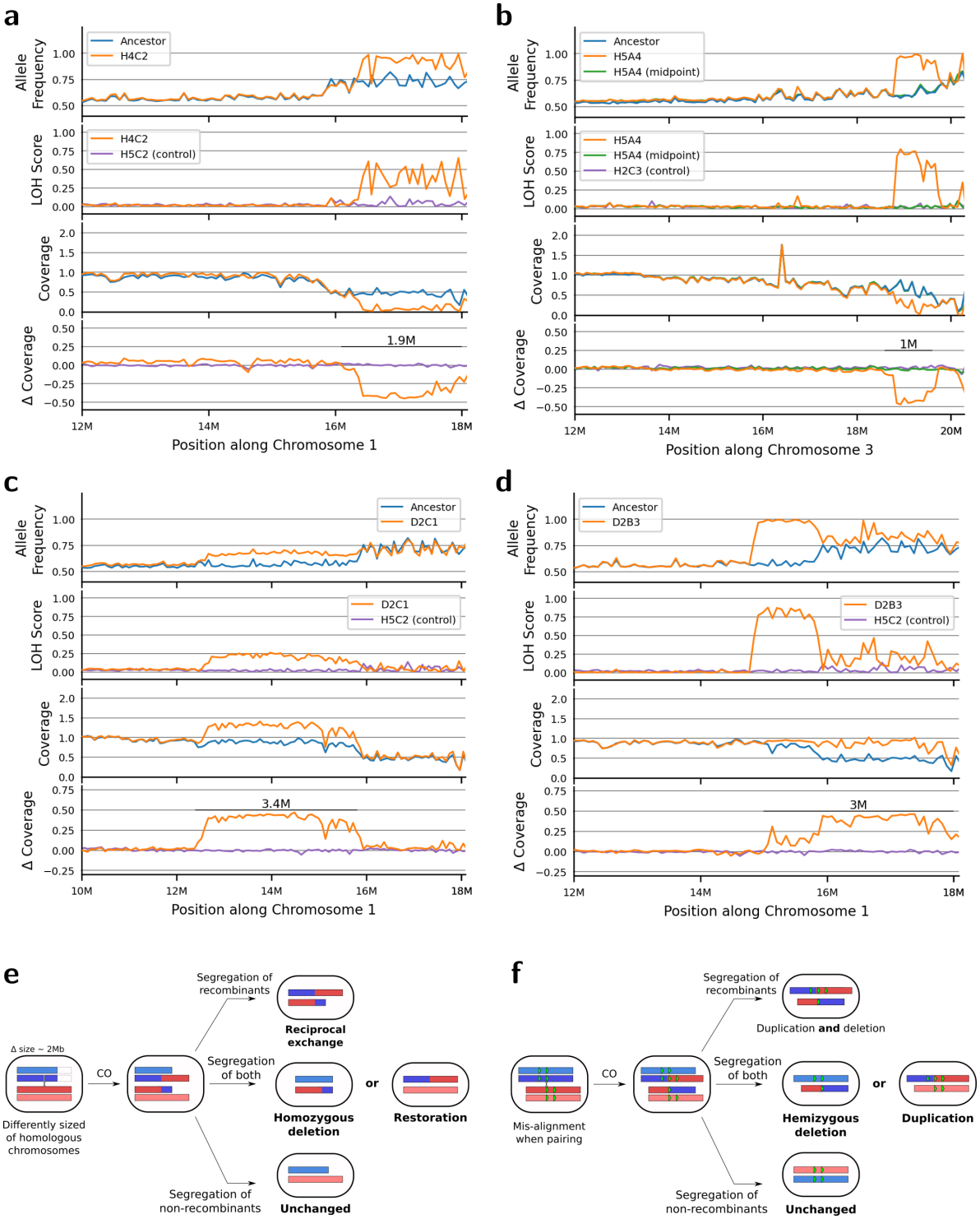

Figure 3. Large deletions and duplications after generations of ME. a-d: Allele frequency, LOH score, Coverage and Δ coverage in non-overlapping windows at the end of chromosomes. **a**: Homozygous deletion on chromosome 1 of sample H4C2 (window = 72.6 Kb). **b**: Hemizygous deletion on chromosome 3 of sample H5A4 (window = 72.6 Kb). There are only 4 hemizygous deletions at the ends of chromosomes. They are difficult to distinguish from homozygous deletions based on the presented plots because of noise in the alignment. All large deletions were checked for coverage retention, hemizygous deletions retained coverage in the region, whereas homozygous deletions have lost all coverage. **c**: Duplication of a region on chromosome 1 of sample D2C1 (window = 81.4 Kb). **d**: Restoration of an ancestral deletion on chromosome 1 in sample D2B3 (window = 72.6 Kb). **e**: Recombination between homologous chromosomes of different sizes (e.g.: chromosome 1 of *A. vaga*) can produce homozygous deletions or restorations of coverage at chromosome extremities depending on how recombinant chromatids segregate. **f**: Recombination between mis-aligned chromosomes (i.e.: at repetitive regions) can produce segmental duplications or deletions.

Figure 4. Summary of positions of larger events (>100 Kb for visibility) in ME and GR per sample per chromosome (based on Supp. Figs. 7 and 11).

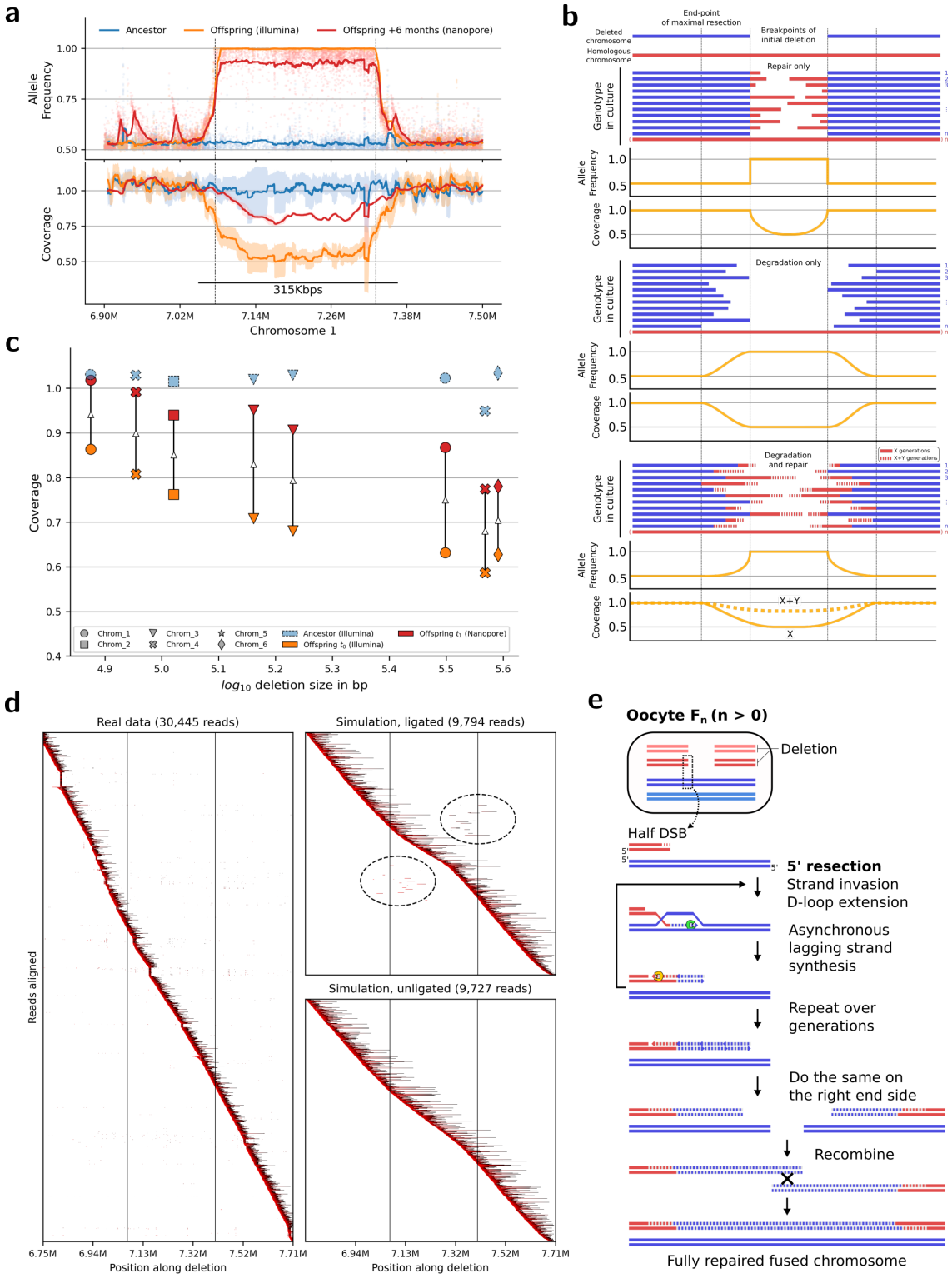

Figure 5. Progressive DNA repair of fragmented chromosomes. **a**: LOH score and Δ coverage along a deleted region of chromosome 1 in the GR 500 Gy IR sample P500 C18 E3. **b**: Comparison of expected patterns of AF and coverage in a deleted region with and without extensive DNA ends resection. **c**: Evolution of coverage (y-axis) sorted by deletion size (logged x-axis) in sample P500 C18 E3. Two timepoints are show: Illumina t_0 (blue) and Nanopore t_1 (orange). **d**: Long reads mapping on the reference region of the deletion shown in **a**. The alignment color corresponds to the sequencing read portion that aligned (red: head, black: tail), the x-axis is the position along the reference sequence. The vertical black lines indicate the LOH tract boundaries of the deletion. Reads alignment from the sequencing data (left panel) and both simulated data (right panels, cf. Methods) are shown. **e**: Proposed model of progressive repair of single DNA ends (half-DSB) by BIHER.

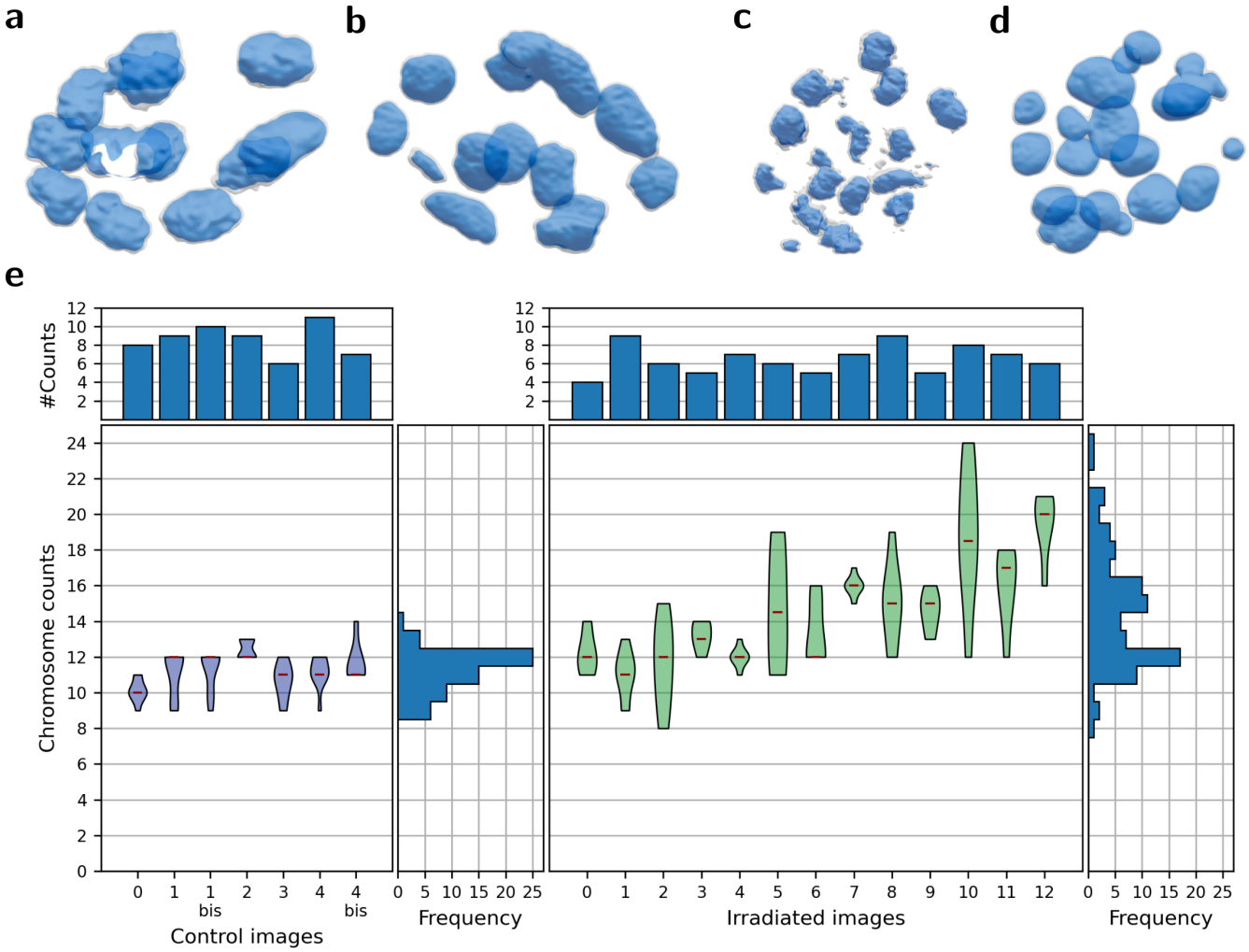

Figure 6. Higher number of chromosome fragments in GR sample P500 C18 E3. a-b 3D representations obtained from a z-stack microscopy of a one-celled egg laid by a non-irradiated control mother (in panel e, these images correspond to counts of control images 1 and 1 bis (panel a) and images 4 and 4 bis (panel b)). **c-d**: 3D representation obtained from a z-stack microscopy of a one-celled egg laid by a mother from the P500 C18 E3 culture (500 Gy irradiated) (in panel e, these images correspond to "irradiated" image 9 (c) and 11 (d)). **e**: Results of independent chromosome counting based on the 3D karyotypes. The top panels show the number of reported counts per image. The bottom panels show the distribution and medians of counted DAPI-stained bodies per image (violin plots) and the frequency of reported counts (histograms) per experimental condition. Counters were blind to the experimental condition. In control images, two identical images (indicated by *bis*) were presented separately to some counters in order to check for consistency. Two out of 6 counters presented with both images reported different counts but failed in only one test. In both consistency tests, the median value of both counts is identical.

Methods

Ancestral strain and cultures

One *A. vaga* individual from the AD008 (COI sequence accession number is KM043184) laboratory strain was isolated and cultured until a population of several thousands individuals was established. This population served as the ancestral line for both the ME and GR experiments. All cultures were maintained in SPA® water and were fed with sterile (0.45 μm pore-size filtering and further autoclaved for 30 min at 121°C) extract of lettuce juice. Subsequent lineages derived from both experiments were cryopreserved at −80°C and stored in 1.5 mL Eppendorf tubes in SPA water®.

Micro-evolution (ME) experimental setup

The experimental design for ME is outlined in Supp. Fig. 1. Initially, 120 *A. vaga* individuals were randomly selected and isolated from the ancestral culture. These individuals were divided into two batches across five 12-well plates, with one individual per well. Five plates were assigned to the H (hydrated) condition, remaining continuously hydrated, while the other five plates were allocated to the D (desiccated) condition. In the D condition, the 12-well plates were submitted to four cycles of 10 days of desiccation followed by re-hydration. Before each desiccation cycle, the cultures were left to grow until they reached *ca.* 1000 individuals to increase the survival rate after desiccation[1]. Every fifth and sixth day, the plates were inspected for the presence of laid eggs. Then, a hatchling was randomly selected for isolation, serving as the starting point for the next generation. In cases where the isolated hatchling failed to reproduce, the wells of the previous generations were kept as backups from which another individual was isolated (up to five attempts) to ensure lineage propagation before declaring extinction. The number of failed attempts was recorded and taken into account in the final calculation of the number of generations. Since bacterial, fungi, and other debris contaminants accumulated during the experiment, we regularly refreshed the culture medium to prevent a contaminant overgrowth. Lineages were collected and cryostored at the midpoint of the experiment when they reached the isolation number 50 for desiccated and 70 for hydrated lines. Because of failed reproduction in several cases, the actual generation number is different from the isolation number. At this stage, lineages reached on average 48 actual generations for desiccated lines and 64 for hydrated lines. Lineages were collected and cryostored at the endpoint of the experiment when they reached the isolation number 88 for the desiccated condition and 137 for the hydrated condition. At this stage, lines reached on average 83 (desiccated) and 128 (hydrated) actual generations. Lines that became extinct during the experiment were not counted in these averages. Desiccation was induced at isolation numbers 20, 42, 62, and 82. Supp. Table 2 summarizes the number of actual generations reached by each lineage at every isolation step.

Genome repair (GR) experimental setup

A summary of the experimental design of GR is shown in Supp. Fig. 2. In GR, the ancestral culture was split into several culture lines:

- 30H: In hydrated conditions during 30 days

- 30D: Desiccated conditions during 30 days
- 0G: (control) Desiccated for 1 day, then placed into the irradiation chamber without irradiation.
- 100G: Desiccated for 1 day then irradiated with 100 Gy protons
- 250G: Desiccated 1 day and then irradiated with 250 Gy protons
- 500G: Desiccated 1 day and then irradiated with 500 Gy protons

After each experimental treatment, *A. vaga* cultures were re-hydrated and individuals were immediately isolated in multi-well plates. The third, fourth, and fifth eggs laid by the isolated mothers were collected and isolated in individual Petri dishes. The first two eggs were discarded in case they already started embryonic development and were not at a unicellular oocyte stage when undergoing the treatment. The isolated eggs were then left to grow to a dense culture (*ca.* 15,000 adults per plate). These cultures were split to obtain triplicates. Then, they were washed with SPA water and harvested. The collected rotifer pellets were cryostored. The code used to name GR samples was the following: C## refers to the clone number (*i.e.:* the number of the isolated mother), E## refers to the egg number (*i.e.:* the number specifying the order of the laid egg from the isolated mother).

Fertility experiment The fertility and viability of cultures were evaluated for the GR experiment. 120 females were isolated in 12-well cell culture plates (Greiner bio-one®) and kept at 19°C in 2.0 mL SPA water during 15 days and regularly fed with 5μL of sterile lettuce juice. The number of adult females and laid eggs per well was recorded for each condition after 7 and 15 days of experiment. Wells containing fewer than 2 adults or showing signs of fungal or bacterial growth were discarded from the analysis. For practical reasons, when the count of laid eggs exceeded 10 per well, it was recorded as 10+. Statistical analysis, including two-sided independent t-tests corrected for unequal variance and Glass' δ statistics (Supp. Fig. 12), was conducted using `SciPy`[2].

Desiccation protocol and irradiation protocol Desiccation and irradiation of cultures were performed as described in Hespeels *et al.* (2014). Individuals desiccated for 1 day were submitted to 1.7 MeV proton radiation (delivering 25 keV μm^{-1})[1].

Culture collection of GR lines Cultures of each experimental condition were collected in 1.5 mL Eppendorf tubes with SPA water in quadruplicates. Each collection tube contained approximately 15,000 individuals. Tubes were centrifuged and stored at −80°C. The frozen rotifers pellets from these tubes were used for genomic DNA extraction.

Culture and collection of ME lines for sequencing

ME lines that were selected for sequencing were fetched from the freezer and cryopreserved rotifers were thawed at $4°C$. Few individuals were collected at midpoint and endpoint of the ME experiment. Therefore, thawed rotifers were cultured until the culture reached *ca.* 15,000 adults. The cultures were then incubated overnight in 15 mL SPA water and 150 μL of PenStrep 10,000 U. Cultures were conserved in 75 cm^2 cell flasks or in 100 mm Petri-dishes in both case with 15 mL SPA water. For collection, the Petri dishes or cell flasks of one sample were vortexed then transferred to 15 mL Falcon centrifugation tubes and centrifuged at 3,220 g. Supernatants were discarded and pellets from tubes containing the same sample were pooled.

Then the Falcon tubes were leveled at 15 mL with filtered SPA water (0.45 μm pore-size) and centrifuged at 3,220 g. Supernatants were discarded, and pellets were washed with 15 mL filtered SPA water and centrifuged at 3,220 g. Then the supernatants were discarded and pellets were transferred to 1.5 mL Eppendorf tubes, leveled at 500 μL, and centrifuged at 16,048 g. These tubes were cryostored until used for genomic DNA extraction.

Genomic DNA extraction

Thawed samples stored at $-80°$C were brought to 4°C, the culture pellets were used for DNA extraction with the QIAGEN Gentra Puregene Tissue Kit. The 1.5 mL Eppendorf tubes containing each one rotifer pellet were centrifuged at 16,048 g for 5 minutes. The supernatant was discarded, and the pellet was kept on ice for 5 minutes. 300 μL lysis solution was added over the pellet and left to incubate for 15 minutes. Then 1.5 μL proteinase K solution was added, and the pellet was resuspended. This solution was incubated overnight at 56°C. The following day, 1.5 μL RNAse A was added, and the solution was slowly mixed and incubated for 30 minutes at 37°C in a shaking dry bath followed by 2 minutes on ice. Then, 100 μL of the protein precipitation solution were added, and the tube was vortexed for 20 seconds. The sample was centrifuged at 4°C 16,048 g for 4 minutes and the supernatant was transferred into a DNA low bind Eppendorf tube containing 300 μL of isopropanol. The solution was gently mixed by inversion and centrifuged at 4°C 16,048 g for 2 minutes. The supernatant was discarded, and the pellet was washed with 300 μL of cold ethanol 70% (stored at $-20°$C). The solution was mixed by inversion and centrifuged at 4°C 16,048 g for 2 minutes. The supernatant was discarded, and the pellet was left to dry at room temperature for *ca.* 10 minutes. Finally, they were resuspended in 60 μL of the DNA hydration solution and stored at 4°C until use.

Illumina sequencing

PCR-free 250-bp paired-end sequencing was performed at the Genomics Core (UZLeuven) on the Illumina HiSeq2000/2500 and NovaSeq platforms, depending on the sequencing run, aiming for a sequencing depth of at least 100X per sample.

Nanopore sequencing

One GR 500 Gy (P500 C18 E3) sample was cultured for an additional 6 months and subsequently sequenced using an Oxford Nanopore PromethION at KeyGene Netherlands. KeyGene received genomic DNA, and it was treated with a Circulomics SRE kit to remove smaller fragments. 2.5 μg of Ampure purified genomic DNA were used as starting input for library construction. PromethION 1D Genomic DNA by ligation SQK-LSK109 library prep kit was used to construct a 1D library from the sample. Approximately 800 ng were loaded on a FLO-PRO002 (R9.4.1 pore) version flow cell. The sample was sequenced on two flow cells on a PromethION P24 in two 72 hr runs, since the first flow cell turned out to be of a relatively low pore number. Basecalling was performed in real-time PromethION compute module using package 20.06.18 using the high-accuracy basecalling model. Approximately 123 Gb of ONT data was produced. It was sequenced in two runs. The sequencing reads that passed QC ($Q \geq 7$) had runs N50s of 26,140 and 27,035 respectively.

Bioinformatic analyses

QC

Genomic reads were analyzed using FastQC[3] and multiQC[4]. Detailed information regarding the `.fastq` files produced is available in Supp. Tables 2. Certain samples exhibited notable adapter sequence contamination at the read extremities. Similarly, some sequencing reads had poor quality metrics at extremities. Both adapters and poor quality bases were trimmed prior to alignment using `fastp`. Some libraries contained some signatures of contaminant sequences (detected by the presence of a second %GC peak). Contaminants were primarily bacteria (mostly Enterococcus and *E. coli*) and fungi (mostly *Saccharomycodes ludwigii*). Despite the presence of these contaminants, the reads were retained in the libraries as they would not align to the reference genome.

Variant calling of Illumina data

The different steps of the variant calling pipeline were performed with a python wrapper pipeline (detailed in Supp. Methods) that runs variant calling with `GATK4`[5] (available at github.com/AntoineHo/runGATK).

Genotype change detection

The variant file, containing information about all reference sites, was annotated to include comparative information between ancestor and offspring genotypes using a custom python script. The added information were, 1) how many samples share a common variant (added as a format field SH), 2) the type of variant (added as a format field TP) such as uncalled, biallelic SNP, biallelic indel, multiallelic SNP, etc. 3) a tag indicating how the site changed compared to the ancestor sample (added as a format field GET) such as C for conserved genotype, L for loss of heterozygosity, N for new heterozygous site, etc. A site was considered multi-allelic if more than two alleles are represented by >10% of the reads aligning at the considered position.

Variant calling of nanopore data

Nanopore reads were aligned on the reference genome using `ngmlr`[6]. The resulting alignments were sorted using `samtools`[7]. The coverage was obtained using the `sambamba depth base`[8] command and variants were called using `bcftools mpileup`[7] and `bcftools call` (cf. Supp. methods for the parameter set used). The final `.vcf` file was analyzed with custom Python scripts.

`bcftools mpileup` was run with the following parameters:

```
--config ont
--annotate "FORMAT/DP,FORMAT/AD"
--ignore-RG
```

`bcftools call` was run with the following parameters:

```
--c -P 0.1 -v
```

Allele frequency, coverage, LOH tracts and LOH cluster definition

Allele frequency (AF) was defined as the ratio between the number of reads supporting the most represented allele at a given genomic position and the total number of reads aligned at the same position. The genomic positions considered for AF calculation are all the sites where heterozygosity is detected in the ancestor sample. AF should be close to 0.5 at heterozygous positions, while it should reach 1.0 when only one allele is represented at a position. The depth of sequencing at a position was obtained using the `sambamba depth base` command[8]. The depth data were used per sample to determine the average and median depths of sequencing (Supp. Fig. 17). AF and coverage information of heterozygous sites were parsed from the `.vcf` using `pysam`. In subsequent analyses, coverage was defined as the total number of reads aligned at a given genomic position divided by the average homozygous coverage of the considered sample in central chromosome regions.

The average heterozygosity H was calculated using the number of heterozygous sites passing these filters: $0.35 \leq \text{AF} \leq 0.65$ and $25\% \leq \text{coverage} \leq 1000\%$ (% coverage is relative to the sample average). The filtered number is then divided by the size of the genome. Additionally, in the ME experiment only, the LOH rate was computed as $\frac{1-H}{g}$ where g represents the number of generations in the specific ME line under consideration. Average heterozygosity and LOH rates were compared between ME hydrated and desiccated lines at endpoint and midpoint. Endpoint and midpoint values within one condition were compared. Across all pairwise comparisons (e.g.: hydrated vs. desiccated, midpoint vs. endpoint), Student's t-test without corrections were performed if both samples passed a normality test (Shapiro-Wilk test) and had equal variance (Levene's test). Student's t-test with correction for unequal variance were performed if samples had unequal variances. If samples failed the normality test, a Mann-Whitney U (MWU) rank test was performed instead. Effect sizes are reported as Cohen's d for Student's t-tests and as Rank Biserial Correlation for MWU rank tests.

Loss of heterozygosity were detected as clusters of consecutive heterozygous sites in the ancestor that became homozygous in the offspring. The LOH **clusters** were described using three metrics: their genomic position, their size (counted in number of ancestrally heterozygous sites), and coverage. The Δ coverage is calculated as follows: $\Delta C = C_{Off} - C_{anc}$ where C is the average coverage of a LOH cluster (example provided in Supp. Fig. 3**a**).

On a chromosome scale, LOH cluster analysis fails due to their division, mostly by repetitive content that leads to erroneous heterozygous calls (paralogous mappings). To visualize long LOH **tracts** at the chromosome scale, AF and coverage were averaged within large non-overlapping windows (1/250 of chromosome size) along the chromosomes. Window sizes per chromosome were 72587 bp (Chromosome 1), 65099 (2), 81419 (3), 60899 (4), 67722 (5) and 55572 (6). LOH, duplications and deletions tracts were identified by observing the changes in AF, coverage, LOH score (defined as $2 \times |AF_{Off} - AF_{anc}|$) and Δ coverage. Supp. Table 1 and Supp. Fig. 3**b** summarize the expected patterns of AF, LOH score, coverage and Δ coverage in various scenarios involving LOH, deletions and duplications.

LOH tracts and clusters detection were performed using custom Python scripts with a few filters being applied. The LOH clusters were stored in a dataframe that was used to perform all subsequent analyses and representations. Depending on the analysis or representation of the clusters, various filters were applied. Most importantly, the number of samples sharing a LOH cluster must be lower than 2 (unique to a sample) to avoid false positives. We filtered clusters with a size lower than 5 sites based on the results of independent LOH clustering (cf. Supp. Methods). We filtered out low (<25%) and high (>1000%) coverage clusters using the average

coverage of the sample. In cluster representation along the chromosomes, we merged clusters that were separated by less than 1 Kb after filtering to avoid cluster splits due to paralogous mapping on local repetitive content.

Simulation of independent recombination events

.

To model the distribution of events across ME samples at endpoint; 53 independent events (successes) were distributed across 22 samples in 10,000 simulations using Numpy. The expected number of samples carrying n recombination events was computed by averaging the result of all simulations. The average number for each n was used as the expected frequencies for the Null hypothesis that events are independent, and compared to observed values of n with a Chi2 test.

Diffusion model of deletion repair

The deletions are repaired gradually by a regular molecular mechanism. To investigate this hypothesis, we fitted the coverage data along the chromosomes at two timepoints (first Illumina sequencing and second nanopore sequencing) to estimate this rate of repair for each deletion. To do so, we modeled coverage f at a position p and time t as a function: $f(p,t)$. We assumed that the repair process could be represented by a diffusion-reaction process (Equation 1).

$$\frac{\delta f(p,t)}{\delta t} = \theta_1 \frac{\delta^2 f(p,t)}{\delta p^2} + \theta_2 f(p,t)(1 - f(p,t)) \tag{1}$$

$$\begin{array}{llllll} \text{a)} & f(p,0) = 1 & \text{for} & p<p_{min} & \text{or} & p>p_{max} \\ \text{b)} & f(p,0) = 1/2 & \text{for} & p>p_{min} & \text{and} & p<p_{max} \\ \text{c)} & f(0,t) = f(1,t) = 1 & & & & \end{array} \tag{2}$$

θ_1 represent the diffusion coefficient associated with the repair and θ_2 the rate of repair, and where p_{min} and p_{max} represent the initial boundary positions of the deletions along the chromosome. The parameters θ_1 and θ_2 were adjusted by least square. The coverage data were smoothed using a moving average with a window width of 50 positions. The coverage data was slightly shifted up or down, so that their average value was near one at both extremities, as assumed in Eq. 2b. This shift was computed as one minus the average frequency observed in the windows defined from $[0, p_{min}/2]$ to $[1/2 + p_{max}/2, 1]$.

Deletion and nanopore reads simulations

`samtools faidx` was used to extract the reference sequence corresponding to a large deletion observed in sample P500 C18 E3 on chromosome 1. `minimap2` was used to align nanopore long reads with splice-aware settings (`-ax splice -G400k`). `nanosim`[9] was used to analyze of the error profile of the real nanopore dataset and then simulate two new sets of long reads. These two sets were based on new reference sequences generated by simulating 50 pairs of breakpoints around the actual deletion breakpoints. The 50 pairs of breakpoints have variable positions to account for variability in deletion size, as is observed in the actual data. Thus, both sets

of reference sequences contain the same deletions but in one set the deletions were ligated *in silico* (set 1) and in the other set, the deletions were left unligated (set 2). In addition, both sets also contained 50 un-deleted reference sequences to account for the presence of a non-deleted allele. The two sets of long reads obtained were aligned with `minimap2` using the same splice-aware settings. A custom Python script was used to plot the alignments obtained with the real dataset and both simulated sets (unligated vs. ligated deletions). A summary of the methodology is represented in Supp. Fig. 15.

Karyotypes

A clonal culture was initiated using a rotifers back-up of the GR sample P500 C18 E3. The culture was maintained in a 75 cm^2 cell flask until it reached *ca.* 40,000 individuals. Similarly, a clean lab strain culture was maintained and used as a control. Then, individuals were concentrated in a smaller volume (10 cm^2 petri dish) and cultured for two days at 25°C in food-rich water with demecolcine 10 μg/mL. After two days, eggs were collected in an Eppendorf 1.5 mL tube and centrifuged at 7,379 g. The supernatant was discarded and 500 μL of SPA water (put at pH 10 with NaOH) was added. Then 500 μL of a diluted bleach solution was added (14 mL SPA water with 1 mL bleach). The tubes were vortexed then the tubes were centrifuged at 7,379 g for 2 minutes, rinsed with 1 mL of SPA water and the supernatant was removed. These last three steps were repeated three times. The egg pellet was resuspended in 1 mL of a hypotonic solution (0.05 M sucrose in distilled water). The mix was transferred to a 15 mL Falcon tube. The tube is then placed on a vortex and shaken at high speed, while 4 mL of the hypotonic solution is slowly added drop by drop. The eggs are left in the incubator overnight on an agitator. The tubes were then centrifuged for 10 minutes at 3,220 g and the supernatant was removed. Then 5 mL of a fixating solution (3 parts of methanol and 1 part of acetic acid) was slowly added drop by drop while the tubes were vortexed at high speed. The tubes were left on ice for 1 hour and then centrifuged at 3,220 g for 10 minutes. The supernatant was discarded and 5 mL of the fixating solution was added drop by drop (more rapidly) while the tubes were vortexed at high speed. The tubes were left on ice for 20 minutes and centrifuged at 3,220 g for 10 minutes. The supernatant was discarded. *ca.* 15 μL of the fixated eggs were dropped from *ca.* 15 cm on a microscope slide. The drops were dried and, for some slides, gently squashed by pushing on the coverslip. The drops were colored using DAPI (1 μg/mL in PBS) and washed twice with PBS. Finally, the slides were mounted with fluoromount-G (left to solidify for at least one hour at 37°C) and stored at 4°C until observation. Z-stacks images of one-celled eggs were then acquired using a Zeiss LSM Airyscan confocal microscope. Images were taken only when chromosomes were clearly visible and when at least 10 chromosomes were initially counted. Z-stacks were then threshold and converted to 3D images with a custom python script. These 3D images were displayed on a public website, on which users can input a chromosome counts. The experimental condition was masked for users, and some images were duplicated to check for consistency in counting. Counts were then aggregated between users, noisy images with high variability in reported counts were discarded.

Methods references

1. Hespeels, B. *et al.* Gateway to genetic exchange? DNA double-strand breaks in the bdelloid rotifer Adineta vaga submitted to desiccation. *J. Evol. Biol.* **27,** 1334–1345 (2014).

2. Virtanen, P. *et al.* SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* **17,** 261–272 (2020).

3. Andrews, S. *et al. FastQC: a quality control tool for high throughput sequence data* 2010.

4. Ewels, P., Magnusson, M., Lundin, S. & Käller, M. MultiQC: summarize analysis results for multiple tools and samples in a single report. *Bioinformatics* **32,** 3047–3048 (2016).

5. Van der Auwera, G. A. *et al.* From FastQ data to high-confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinformatics* **43,** 11–10 (2013).

6. Sedlazeck, F. J. *et al.* Accurate detection of complex structural variations using single-molecule sequencing. *Nat. Methods* **15,** 461–468 (2018).

7. Danecek, P. *et al.* Twelve years of SAMtools and BCFtools. *Gigascience* **10,** giab008 (2021).

8. Tarasov, A., Vilella, A. J., Cuppen, E., Nijman, I. J. & Prins, P. Sambamba: fast processing of NGS alignment formats. *Bioinformatics* **31,** 2032–2034 (2015).

9. Yang, C., Chu, J., Warren, R. L. & Birol, I. NanoSim: nanopore sequence read simulator based on statistical characterization. *GigaScience* **6,** gix010 (2017).

Data access

Data and materials availability: Haploid genome assembly, gene annotation, and sequencing reads of the ancestor sample are available under the bioproject accession PRJNA680543. The micro-evolution sequencing reads are available under the bioproject accession: PRJNA929432. The genome repair sequencing reads are available under the bioproject accession: PRJNA930383.

Scripts and codes used to perform the analysis are available at: https://github.com/AntoineHo/genome_evolution_bdelloids_scripts

Acknowledgments

This project received funding from the Horizon 2020 research and innovation program under European Research Council (ERC) grant agreement 725998 (RHEA) and from BELSPO PRODEX for the ESA selected ILSRA-2014-0106 Project to K.V.D., from the Fédération Wallonie-Bruxelles via an "Action de Recherche Concertée" (ARC) grant to K.V.D. and B.Ha. Work in the lab of B.Ha. was supported by a grant "Action de Recherche Concertée" (vIRGIN) from the Fédération Wallonie-Bruxelles and a grant from the "fonds Spéciaux de Recherche" (RADyGen) at UCLouvain.

A.H. and A.D. are Research Fellows of the Fonds de la Recherche Scientifique—FNRS.

Special thanks to Matthieu Terwagne for his help and insights in the DNA repair mechanisms. We also thank Marc Guérineau for his help in experimental design, Jérémy Berthe for the production of desiccated rotifer patches, and Sébastien Meurant and Catherine Demazy for their help with image acquisition on the confocal microscope.

Author contributions

Conceptualization: A.D., M.L., B.Ha., and K.V.D. Data curation: A.H., A.D., and M.L. Formal analysis: A.H., and T.L. Funding acquisition: B.Ha., and K.V.D. Investigation: A.H., A.D., M.L., J.V., B. He., P.S., E.N., and T.L. Methodology: A.H., A.D., M.L., P.S., T.L., E.N., B. Ha., and K.V.D. Project administration: B.Ha., and K.V.D. Resources: B.He., E.N., J.V., A.-C.H., and K.V.D. Software: A.H., and T.L. Supervision: T.L., B.Ha., and K.V.D. Visualization: A.H., T.L., and K.V.D. Writing—parts of the original draft: A.H., P.S., and K.V.D. Writing—review, and editing: All authors. Supervising the entire article writing: A.H., B.Ha., and K.V.D

Competing interest declaration

The authors declare that they have no competing interests.

Additional information

Supplementary Information is available for this paper.