Title: Gene conversion limits the cost of asexuality in immortal worms

Authors: Simon Kershenbaum*, Manuel Jara-Espejo, Ashleigh Griffin, *Aziz Aboobaker

Author affiliations: Department of Biology, University of Oxford, Oxford, United Kingdom, OX1 3SZ

*Corresponding authors: simon.kershenbaum@new.ox.ac.uk ; aziz.aboobaker@biology.ox.ac.uk

Keywords: “Asexuality”; “Muller’s ratchet”; “Gene conversion”

Significance statement

Sexual reproduction is the most common reproductive strategy for multicellular organisms, while obligate asexuality is very rare. This rarity of asexuality is driven by a reduced ability to adapt to new challenges and an accumulation of deleterious mutations associated with the lack of sexual recombination.

Here we find that obligate asexual planarians are able to avoid the accumulation of deleterious mutations associated with asexuality by exploiting a form of recombination within the stem cells of the worm to remove deleterious mutations.

This study reinvigorates the long-standing debate on how some asexual lineages persist, by asking this question in obligate fissiparous animals and showing that asexual recombination in somatic cells followed by selection can help limit the cost of asexuality.

Abstract

Most multicellular organisms reproduce sexually despite the costs associated with sexuality[1]. This has been explained as the result of selection favouring the ability to recombine the genome. The lack of recombination in asexual species constrains their adaptability and leads to the accumulation of deleterious mutations, ultimately increasing their risk of extinction[2]. Nonetheless, successful asexual life histories persist among multicellular organisms, and explanatory mechanisms which may help limit the cost of asexuality remain enigmatic[3–6]. In search of these mechanisms, we looked at that the molecular evolutionary changes in sexual and obligate asexual strains of the planarian flatworm, *Schmidtea mediterranea*. We find that the accumulation of deleterious mutations is largely avoided in the asexual strain. We find evidence that this is achieved by somatic gene conversion in stem cells allowing for the restoration of fit alleles and the purification of deleterious mutations. Our analyses provide a potential mechanism for the maintenance of asexuality in a metazoan. Taken together, these findings suggest that gene conversion may be used by asexual lineages to unlink loci, counteracting some of the key challenges facing obligate asexual species.

Main

Most multicellular life reproduces sexually, while obligate asexuality is exceedingly rare[1]. This pattern persists despite the high costs of sex, including the “two-fold” cost of sex[8], the breaking up of beneficial haplotypes[9], and the broad costs associated with securing a mate[10]. The maintenance of sexuality in the face of such costs suggests that asexuality, the alternative reproductive strategy, is somehow fundamentally limited and is outcompeted by sexual strategies in the long term[11]. However, some ancient asexual organisms exist, which raises the question of how they avoid the limitations of asexuality[1,3–6].

The potential disadvantages of asexual reproduction are driven by the tight linkage between alleles in an asexual genome (Box 1)[2,12,13]. This linkage means that selection can not act on

individual alleles or allele combinations, but instead must act on the genomes of individuals as a whole - a problem known as "The Hill-Robertson effect"[12–14]. This reduces the adaptability of asexual populations, which has been shown to be particularly problematic in the context of changing environments and biotic competition[15–17]. Additionally, the tight linkage of alleles in asexual genomes means that deleterious mutations cannot be easily removed. This means that as the fittest genomes in the population reach fixation, so do their associated deleterious mutations, resulting in the iterative reduction of fitness in asexual populations by a process known as "Muller's ratchet"[11,18]. Taken together, the reduced ability to unlink alleles in asexual genomes is thought to drive Muller's ratchet and the Hill-Robertson effect, both suggested as explanations for why obligate asexual reproduction is rare[2].

However, some obligate asexual species persist over evolutionary time[3,6,19,20], suggesting that they may have acquired currently unknown mechanisms to limit the impact of long-term asexuality. The planarian flatworm *S. mediterranea* is a highly regenerative animal that has both a sexual and asexual species. The sexual species is a cross fertilizing hermaphrodite and the obligate asexual strain lacks reproductive organs, reproducing solely by fission followed by whole body regeneration[21,22] (Fig 1). The asexual strain is evidently able to thrive, despite the limitations of asexual reproduction, providing us with an opportunity to examine potential mechanisms for dealing with the problems of asexuality[6,23,24]. Regeneration after fission is mediated by somatic, immortal, pluripotent adult stem cells called neoblasts[25]. Continuous BrdU labelling suggests that these neoblasts have a generation time of 24-72 hours as they are constantly in the cell cycle, suggesting that since the origin of the asexual lineage there have been at least tens to hundreds of millions of asexual neoblast generations[26,27], dependent upon when the asexual species arose[6]. These different life histories of two sister species allows for the direct genetic comparison of sexual and asexual species, which may help us understand how asexual species might persist over evolutionary time.

Long-term asexual species are likely to have acquired adaptations to limit the cost of asexuality, however the mechanisms by which this is achieved are not well understood. One suggested mechanism to limit the cost of asexuality is gene conversion, a form of asexual recombination in which one allele overwrites another (Box 2)[7,28]. This provides a mechanism for the removal of deleterious mutations in asexual populations. If gene conversion limits the impact of Muller's ratchet, we would expect to see reduced levels of mutation accumulation in asexual species where gene conversion is present. Gene conversion is a GC-biased homogenising process (Box 2), so we would also expect to see GC-biased excess homozygosity in asexual species where gene conversion is active. In this study, we find that asexual planarians limit the impact of Muller's ratchet, and that they show GC-biased excess homozygosity indicative of gene conversion. Thus, we find evidence that gene conversion mediated removal of deleterious mutations may allow asexual planarians to persist despite their obligate asexuality.

<u>Limited impact of Muller's ratchet in asexual planarians</u>

We began by looking for evidence of deleterious mutation accumulation and Muller's ratchet in the obligate asexual strain by using dNdS analyses. We found that across the entire set of protein coding genes (Supplementary data 1), the median asexual dNdS value is 2% higher than that found in the sexual strain (Wilcoxon, W = 54,470,962, P < 0.001, Fig. 2A). High dNdS values can indicate both deleterious mutation accumulation and positive selection, thus we assessed dNdS values in genes that are unlikely to be under positive selection. We first compared the dNdS values of exons under significant negative selection and found that the difference between the strains was marginally significant, with the median asexual dNdS values being 11% higher than in the sexual strain (Wilcoxon, W = 5,019, p-value = 0.049, Fig 2B). This is in contrast to previous studies that found much higher impacts of Muller's ratchet in asexual lineages (dNdS values ~200%-~3,000% higher than in their sexual counterparts)[3,19,29–32]. When we compared the dNdS values of the highly conserved single copy BUSCO set of core metazoan genes[33], we found no significant difference between the strains (V = 4,489, P = 0.307, Fig.2C). Thus, while asexual planarians may be subject to Muller's ratchet, the accumulation of deleterious mutations is limited, and is not present in highly conserved single copy genes. This requires a mechanistic explanation in the absence of sexual recombination.

Loss of heterozygosity mechanisms in asexual may limit Muller's ratchet

Mechanisms limiting Muller's ratchet in asexual planarians may involve somatic recombination mechanisms, such as gene conversion, that can remove deleterious mutations from the genome (see Box 2). A signature of such processes would be higher than expected levels of homozygosity in asexual genes. We found that asexual variant homozygosity was higher than what would be expected assuming an absence of recombination, suggesting that some form of recombination takes place within the adult stem cells. In the absence of recombination, the expected proportion of homozygous ancestral variants (variants found in the common ancestor of the sexual and asexual strains) is 0.32 (see methods), while we find a median and mean homozygosity score of 1 and 0.79 respectively in ancestral asexual variants (Wilcoxon, V = 132,691,491, P < 0.001, Fig.3A). This suggests that some process is driving excess homozygosity in the asexual strain. Furthermore, this process is ongoing and not purely ancestral as we observe excess homozygosity in variants that arose after the sexual-asexual split which would be expected to be all heterozygous (homozygosity score; median = 0.05, mean = 0.42, Wilcoxon, V = 0, p-value < 0.001, Fig.3A). Additionally, ancestral variants were more homozygous than variants that arose specifically in the asexual lineage, suggesting the possibility that asexual variants become homozygous over time (Wilcoxon, W = 112,733,042, P < 0.001, Fig.3B, supplementary Fig2). These results suggest a genome wide mechanism in driving the loss of heterozygosity in asexual planarians, that could limit mutation accumulation in this asexual lineage.

GC biased Gene conversion in asexual planarians

Among the potential mechanisms for the observed loss of heterozygosity, gene conversion seemed a likely candidate. We hypothesised that since gene conversion has a GC bias[28], homozygous variants should display this signature if they were indeed generated by gene

conversion. We find that homozygous variants are GC enriched compared to heterozygous variants (Chi-sqr = 541.55, df = 1, P < 0.001, Fig.4A), suggesting a GC bias in the mechanism driving high asexual homozygosity. Additionally, we find that homozygosity is a significant predictor of GC-AT identity, with more homozygous variants being more likely to have a GC identity in the asexual (binomial GLM, P < 0.001, Fig.4B, supplementary table 3). Furthermore, we show that AT to GC mutations have a median homozygosity score 24.29x higher than that of GC to AT mutations and mean score 1.45x higher (Wilcoxon, W = 191,040,629, P < 0.001, Fig.4C). This GC bias is not driven by AT to GC mutations across the transcriptome (supplementary figure 3). Finally, we find that variants with higher allele frequency are more likely to have a GC identity, and that this bias is stronger in the asexual strain compared to the sexual (binomial GLM, P < 0.001, Fig.4D, supplementary table 3). Taken together, this suggests that GC biased gene conversion takes place in the asexual planarian, leading to loss of heterozygosity which could limit the impact of Muller's ratchet (Box 2).

<u>Selection on gene conversion events removes deleterious alleles</u>

Since gene conversion has the potential to generate homozygotes for both beneficial and deleterious variants, selection must be able to remove deleterious gene conversion events for gene conversion to be a viable solution to Muller's ratchet. We find that mutations that are likely to be more deleterious due to their effect on predicted proteins (see methods), are less likely to be homozygous, suggesting that deleterious gene conversion events are selected against (Fig5, supplementary Fig4). We find that indels in coding regions are less homozygous compared to coding SNPs, which are less homozygous compared to noncoding mutations (Pairwise Wilcoxson, P < 0.001, Fig.5A, supplementary table 3). Additionally, we find that coding indels have lower allele frequencies than coding SNPs, which have lower allele frequencies than non-coding variants (Pairwise Wilcoxson, P < 0.001, Fig.5B). We also find similar results when comparing the homozygosity of coding vs noncoding mutations (Wilcoxon, W = 329,188,179, P < 0.001), synonymous vs nonsynonymous mutations (Wilcoxon, W = 266,753,706, P < 0.001), and frameshifting vs non-frameshifting indels (Wilcoxon, W = 51,669, P = 0.003), such that in every case more deleterious mutations are less homozygous (Fig.5B). Similarly, we find that the more deleterious mutations have lower allele frequencies in the population, including coding vs noncoding (Wilcoxon, W = 207,113,417, P < 0.001), nonsynonymous vs synonymous (Wilcoxon, W = 239,588,546, P < 0.001), and frameshifting vs non-frameshifting indels (Wilcoxon, W = 28,861, P < 0.001).We propose that when gene conversion events lead to a homozygous deleterious mutation in a cell, the homozygous deleterious clone is eliminated by intra-organismal and inter-organismal selection, leading to the lower observed homozygosity in variants that are predicted to be more deleterious. The separation of alleles via gene conversion followed by subsequent selection against unfit stem cell clones may help limit the impact of both Muller's ratchet and the Hill-Robertson effect.

Discussion

Mechanisms explaining how some asexual species persist despite the severe limitations of asexuality have remained elusive. Our results support a role for gene conversion in unlinking

loci in the asexual genome of *S. mediteranea*. Combined with selection against deleterious homozygotes this allows these asexual planarians to limit the impact of Muller's ratchet and the Hill-Robertson effect, potentially explaining how this obligate asexual can persist over evolutionary time.

Our results suggest that the impact of Muller's ratchet is relatively mild in asexual planarians compared to other asexual lineages. Many asexual lineages, including relatively recent asexual species, have been shown to have dNdS values ~200% higher than their sexual counterparts, and some asexual bdelloid rotifers have been suggested to have dNdS values ~200%-~3,000% higher than their sexual equivalents[3,19,29–32]. Our results showing that asexual planarians have dNdS values only 11% higher than their sexual counterparts indicate that that the impact of Muller's ratchet might be limited by gene conversion and subsequent selection in asexual planarians. This is further supported by the complete absence of Muller's ratchet in highly conserved single copy genes, where selection is likely to be strongest.

Gene conversion has the potential to remove deleterious mutations in some members of the stem cell lineage. Stem cells whose gene conversion events removed deleterious mutations may then be selected for while those that had their beneficial allele converted to a deleterious variant may be selected against, limiting the impact of Muller's ratchet and allowing deleterious mutations to be purged from the population (See Box 2). Gene conversion as a mechanism to limit the impact of Muller's ratchet is available to most diploid (or higher ploidy) asexual lineages, and in its absence or reduction, we would predict higher levels of mutation accumulation. Supporting this are studies on *Timema* stick insects, which found that asexual populations had lower rates of gene conversion compared to their sexual relatives, and dNdS values 360%-1,340% higher compared to their sexual counterparts[3,29]. Asexual recombination by gene conversion may also reduce the impact of Hill-Robertson interference, since beneficial and deleterious mutations can be unlinked and selected upon as homozygous variants in different lineages. Thus, asexual recombination via gene conversion may solve some of the fundamental limitations of obligate asexuality by unlinking beneficial and deleterious alleles in the genome.

However, while gene conversion may be important in asexual lineages, it cannot solve all the problems of asexuality. For example, it cannot solve the problem of clonal interference, since beneficial mutations that arise in separate lineages cannot be merged[34]. Thus, ineffective selection due to clonal interference, may still limit the long-term viability of asexual lineages and potentially explain their rarity in nature even when gene conversion is present. One reason why asexual planarians may be able to endure these additional limitations of clonal interference is that their obligate fissiparous lifestyle (and thus obligately immortal soma) may allow intra-organismal somatic variability to be selected upon over evolutionary time[35]. This increase in selectable variation may help asexual planarians deal with the reduced efficacy of selection due to clonal interference, which is not addressed by gene conversion.

Obligate asexual reproduction in multicellular organisms is rare and thought to be an evolutionarily ephemeral phenomenon due to the fundamental limitations of asexuality. When we examine *S. mediterrannea*, which adopts this strategy successfully, we are able to assess how it limits the costs of obligate asexuality. Our results suggest that it likely uses asexual recombination in the form of gene conversion, to unlink loci and thus limit the impact of Muller's ratchet and Hill-Robertson interference. While such mechanisms may be available

to most asexual lineages, the opportunity for selection to act on a population of somatic stem cells may be key in allowing asexual planarians to persist.

Methods

Collecting available RNAseq data

RNAseq data was collected from previous studies (supplementary table 1), with a total of 169 libraries from 7 studies for the asexual and 35 libraries from a single study for the sexual[36–43]. When choosing the studies, we only used those that used illumina sequencing and had clear annotations of the strains used. RNAseq data was then trimmed and quality filtered with trimmomatic and FASTQC using default settings[44]. STAR was then used to map the RNAseq reads to the dd_Smes_g4 genome[45,46].

Variant calling

For variant calling we used the GATK software following the best practice (Supplementary data 2A-D)[47]. The only modification being that we did not use machine learning to recalibrate BAM base quality before variant calling. We did not use machine learning to recalibrate BAM base quality scores since it requires a known set of variants (which doesn't exist for *S.mediterranea*) and since it may mask high levels of variation which could exist between independently evolving asexual lineages.

Variant filtration

We used the suggested parameters for variant filtration (supplementary table 2), except that we allowed for variants that had no reads supporting the reference (these are usually filtered out since some of the quality measures require a comparison of reads supporting the reference and alternative alleles). We then filtered out any mutations found at a 2bp distance from one another. This was done to allow for simple interpretation of the impact of the variants in coding regions (supplementary data 3).

Defining orthologues

We ran Orthofinder using a proteome from a new planarian annotation to isolate the orthologues of each predicted gene[48,49]. To reduce redundancy, we first ran CD-hits to collapse all peptides with 99% identity to a single transcript ID (supplementary data 4).

dNdS analysis

We used the dNdScv R package for dNdS analysis and annotation of mutation impact in coding regions[50]. dNdScv uses an expansion on the Poisson framework to obtain dNdS values and a full tri-nucleotide model to avoid systematic bias due to sequence context[50]. To get the list of genes for dNdS analysis we used the list included in the OrthoFinder output, and further filtered for the longest transcript per gene. We then ran dNdScv with no covariates and no limits on the number of variants per CDS. When comparing dNdS values between the strains in genes under purifying selection, we used the local model from dNdScv, since the overdispersion value for the sexual dataset was lower than 1. To compare dNdS values of BUSCO gene exons between the strains, we isolated single copy BUSCO genes in planarians using BUSCO (metazoan gene list). We then isolated BUSCO exons that had both

synonymous and nonsynonymous mutations within them (to allow for meaningful dNdS assessment using the dNdScv local model) and compared their dNdS values provided by dNdScv.

Analysis of zygosity

To analyse the zygosity of different variants between the strains, we generated a homozygosity score from the GATK data. The homozygosity score is simply determined by the proportion of libraries supporting a homozygous mutation out of the number of libraries supporting the mutation. Thus, if a mutation is homozygous in all libraries supporting the mutation, it will have a homozygosity score of 1 and vice versa. To verify that the homozygosity score does not represent biased gene expression rather than homozygosity, we compared SNPs found in both RNAseq data to available ATAC (genomic) data[49]. This showed that when filtering for mutations supported by over 3 libraries, there is over 95% agreement between RNAseq and ATAC based methods of calling purely homozygous variants (supplementary figure 1). Additionally, we filtered for variants that had an average coverage of at least 10 to remove falsely called homozygous mutations (supplementary figure 1). Average coverage was calculated by total depth divided by number of called samples (DP / NCALLED). Note that under the assumption of asexuality, the measure of homozygosity from a group of individuals should be very similar to the measure of homozygosity within a single individual, since asexual worms from the same population should have negligent differences in loci zygosity in the absence of recombination.

Of the variants arising by mutation in the asexual lineage, the expected proportion of homozygous variants is 0 (assuming no recombination and negligible chance of an independent identical mutation in the homologous chromosome). Of the ancestral variants in the asexual lineage (variants shared also by the sexual strain), the expected proportion of homozygous variants depends on the mean frequency of the variants in the ancestral population. This is because they would have originated by a meiotic process originally, and we know that all the variants in question have had at least one copy transmitted to the original zygote (a priori, since we can detect it). The mean allele frequency of the shared ancestral variants in the sexual population is 0.32, thus if we use this frequency as a proxy for the allele frequency in the ancestral population, the expected proportion of ancestral homozygous variants is 0.32 (assuming no subsequent recombination).

We also used the allele frequency from GATK to validate our homozygosity scores and their interpretation. Allele frequency values represent the frequency of the variant across all libraries, such that in a bi-allelic site, the allele frequency of a variant is 1 – (allele frequency of reference variant). We interpret an allele frequency of 1 as a homozygous variant (since there is no evidence for other variants at the same locus). This is used to compare sexual and asexual site polymorphism.

GC bias analysis

To assess whether GC variants are more likely to be homozygous (annotated as homozygous by GATK in all supporting libraries) than heterozygous (annotated as heterozygous by GATK in all supporting libraries), we used a chi-square test. Unless explicitly stated otherwise, we used only synonymous and noncoding mutations for this analysis. We also assessed whether GC SNP proportion in each gene correlates with its heterozygosity score and allele frequency using binomial GLMs (glm() in R) (supplementary table 3).

Assessing evidence of effective selection

To assess whether effective selection is taking place in the asexual population, we wanted to see whether more harmful mutations are less likely to be fixed as homozygous. The rational is that if gene conversion allows for the removal of deleterious mutations, cells in which gene conversion led to homozygosity of a highly deleterious variant should be selected against (Fig 5). Out of all the variants we assessed, we considered noncoding variants to be the least harmful on average, followed by coding synonymous mutations, followed by coding nonsynonymous point mutations, with the most harmful variants assessed being frameshifting indels in coding regions. This ranking is based on the amount of change the mutation is likely to have on the protein sequence.

Supplementary Data:

Supplementary data 1A – dNdS per exon, asexual strain

Supplementary data 1B – dNdS per exon, sexual strain

Supplementary data 2A – Asexual SNPs called, unfiltered

Supplementary data 2B - Sexual SNPs called, unfiltered

Supplementary data 2C – Asexual INDELs called, unfiltered

Supplementary data 2D - Sexual INDELs called, unfiltered

Supplementary data 3 – Master list of variants, sexual and asexual after filtration

Supplementary data 4 – OrthoFinder output

Supplementary table 1 – All the RNAseq libraries used

Supplementary table 2 – The filtration parameters for variants

Supplementary table 3 – GLMs, coefficients values

Acknowledgements: We would like to thank Tim Barraclough and Chris Wilson for their helpful comments on the manuscript.

Bibliography

1. Judson, O. P. & Normark, B. B. Ancient asexual scandals. *Trends Ecol. Evol.* **11**, 41–46 (1996).

340 2. Felsenstein, J. THE EVOLUTIONARY ADVANTAGE OF RECOMBINATION. *Genetics* **78**, 737–756
341 (1974).
342 3. Bast, J. *et al.* Consequences of Asexuality in Natural Populations: Insights from Stick Insects. *Mol.*
343 *Biol. Evol.* **35**, 1668–1677 (2018).
344 4. Wilson, C. G., Nowell, R. W. & Barraclough, T. G. Cross-Contamination Explains "Inter and
345 Intraspecific Horizontal Genetic Transfers" between Asexual Bdelloid Rotifers. *Curr. Biol.* **28**,
346 2436-2444.e14 (2018).
347 5. Butlin, R., Schön, I. & Martens, K. Asexual reproduction in nonmarine ostracods. *Heredity* **81**,
348 473–480 (1998).
349 6. Lázaro, E. M. *et al.* Schmidtea mediterraneaphylogeography: an old species surviving on a few
350 Mediterranean islands? *BMC Evol. Biol.* **11**, 274 (2011).
351 7. Weir, W. *et al.* Population genomics reveals the origin and asexual evolution of human infective
352 trypanosomes. *eLife* **5**, e11473 (2016).
353 8. Gibson, A. K., Delph, L. F. & Lively, C. M. The two-fold cost of sex: Experimental evidence from a
354 natural system. *Evol. Lett.* **1**, 6–15 (2017).
355 9. Charlesworth, B. & Barton, N. H. Recombination load associated with selection for increased
356 recombination. *Genet. Res.* **67**, 27–41 (1996).
357 10. Lehtonen, J., Jennions, M. D. & Kokko, H. The many costs of sex. *Trends Ecol. Evol.* **27**, 172–178
358 (2012).
359 11. Lynch, M., Bürger, R., Butcher, D. & Gabriel, W. The Mutational Meltdown in Asexual
360 Populations. *J. Hered.* **84**, 339–344 (1993).
361 12. Hill, W. G. & Robertson, A. The effect of linkage on limits to artificial selection. *Genet. Res.* **8**,
362 269–294 (1966).
363 13. Comeron, J. M., Williford, A. & Kliman, R. M. The Hill–Robertson effect: evolutionary
364 consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19–31 (2008).

365 14. de Visser, J. A. G. M. & Elena, S. F. The evolution of sex: empirical insights into the roles of
366 epistasis and drift. *Nat. Rev. Genet.* **8**, 139–149 (2007).
367 15. Lively, C. M. & Morran, L. T. The ecology of sexual reproduction. *J. Evol. Biol.* **27**, 1292–1303
368 (2014).
369 16. Vergara, D., Jokela, J., Lively, C. M. & Bronstein, E. J. L. Infection Dynamics in Coexisting Sexual
370 and Asexual Host Populations: Support for the Red Queen Hypothesis. *Am. Nat.* **184**, S22–S30
371 (2014).
372 17. Hamilton, W. D., Axelrod, R. & Tanese, R. Sexual reproduction as an adaptation to resist
373 parasites (a review). *Proc. Natl. Acad. Sci.* **87**, 3566–3573 (1990).
374 18. Muller, H. J. THE RELATION OF RECOMBINATION TO MUTATIONAL ADVANCE. 8.
375 19. Barraclough, T. G., Fontaneto, D., Ricci, C. & Herniou, E. A. Evidence for Inefficient Selection
376 Against Deleterious Mutations in Cytochrome Oxidase I of Asexual Bdelloid Rotifers. *Mol. Biol.*
377 *Evol.* **24**, 1952–1962 (2007).
378 20. Schön, I., Rossetti, G. & Martens, K. Darwinulid Ostracods: Ancient Asexual Scandals or
379 Scandalous Gossip? in *Lost Sex: The Evolutionary Biology of Parthenogenesis* (eds. Schön, I.,
380 Martens, K. & Dijk, P.) 217–240 (Springer Netherlands, 2009). doi:10.1007/978-90-481-2770-
381 2_11.
382 21. Saberi, A., Jamal, A., Beets, I., Schoofs, L. & Newmark, P. A. GPCRs Direct Germline Development
383 and Somatic Gonad Function in Planarians. *PLoS Biol.* **14**, e1002457 (2016).
384 22. Newmark, P. A., Wang, Y. & Chong, T. Germ Cell Specification and Regeneration in Planarians.
385 *Cold Spring Harb. Symp. Quant. Biol.* **73**, 573–581 (2008).
386 23. Tan, T. C. J. *et al.* Telomere maintenance and telomerase activity are differentially regulated in
387 asexual and sexual worms. *Proc. Natl. Acad. Sci.* **109**, 4209–4214 (2012).
388 24. Sahu, S., Dattani, A. & Aboobaker, A. A. Secrets from immortal worms: What can we learn about
389 biological ageing from the planarian model system? *Semin. Cell Dev. Biol.* **70**, 108–121 (2017).

390 25. Wagner, D. E., Wang, I. E. & Reddien, P. W. Clonogenic Neoblasts Are Pluripotent Adult Stem
391 Cells That Underlie Planarian Regeneration. *Science* **332**, 811–816 (2011).
392 26. Newmark, P. A. & Sánchez Alvarado, A. Bromodeoxyuridine Specifically Labels the Regenerative
393 Stem Cells of Planarians. *Dev. Biol.* **220**, 142–153 (2000).
394 27. Fields, C. & Levin, M. Are Planaria Individuals? What Regenerative Biology is Telling Us About the
395 Nature of Multicellularity. *Evol. Biol.* **45**, 237–247 (2018).
396 28. Chen, J.-M., Cooper, D. N., Chuzhanova, N., Férec, C. & Patrinos, G. P. Gene conversion:
397 mechanisms, evolution and human disease. *G E N E Ti C S* 14 (2007).
398 29. Henry, L., Schwander, T. & Crespi, B. J. Deleterious Mutation Accumulation in Asexual Timema
399 Stick Insects. *Mol. Biol. Evol.* **29**, 401–408 (2012).
400 30. Johnson, S. G. & Howard, R. S. Contrasting Patterns of Synonymous and Nonsynonymous
401 Sequence Evolution in Asexual and Sexual Freshwater Snail Lineages. *Evolution* **61**, 2728–2735
402 (2007).
403 31. Paland, S. & Lynch, M. Transitions to Asexuality Result in Excess Amino Acid Substitutions.
404 *Science* **311**, 990–992 (2006).
405 32. Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. & Taylor, D. R. Accelerated Mutation
406 Accumulation in Asexual Lineages of a Freshwater Snail. *Mol. Biol. Evol.* **27**, 954–963 (2010).
407 33. Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. & Zdobnov, E. M. BUSCO Update: Novel and
408 Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of
409 Eukaryotic, Prokaryotic, and Viral Genomes. *Mol. Biol. Evol.* **38**, 4647–4654 (2021).
410 34. Gerrish, P. J. & Lenski, R. E. The fate of competing beneficial mutations in an asexual population.
411 in *Mutation and Evolution* (eds. Woodruff, R. C. & Thompson, J. N.) vol. 7 127–144 (Springer
412 Netherlands, 1998).
413 35. Leria, L., Vila-Farré, M., Solà, E. & Riutort, M. Outstanding intraindividual genetic diversity in
414 fissiparous planarians (Dugesia, Platyhelminthes) with facultative sex. *BMC Evol. Biol.* **19**, 130
415 (2019).

416 36. Davies, E. L. *et al.* Embryonic origin of adult stem cells required for tissue homeostasis and
417 regeneration. *eLife* **6**, e21052 (2017).
418 37. Tu, K. C. *et al.* Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *eLife* **4**,
419 e10501 (2015).
420 38. Duncan, E. M., Chitsazan, A. D., Seidel, C. W. & Alvarado, A. S. Set1 and MLL1/2 target distinct
421 sets of functionally different genomic loci in vivo. *Cell Rep.* **13**, 2741–2755 (2015).
422 39. van Wolfswinkel, J. C., Wagner, D. E. & Reddien, P. W. Single-cell analysis reveals functionally
423 distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **15**, 326–339 (2014).
424 40. Scimone, M. L., Kravarik, K. M., Lapan, S. W. & Reddien, P. W. Neoblast Specialization in
425 Regeneration of the Planarian Schmidtea mediterranea. *Stem Cell Rep.* **3**, 339–352 (2014).
426 41. Zhu, S. J., Hallows, S. E., Currie, K. W., Xu, C. & Pearson, B. J. A mex3 homolog is required for
427 differentiation during planarian stem cell lineage development. *eLife* **4**, e07025 (2015).
428 42. Currie, K. W. & Pearson, B. J. Transcription factors lhx1/5-1 and pitx are required for the
429 maintenance and regeneration of serotonergic neurons in planarians. *Development* **140**, 3577–
430 3588 (2013).
431 43. Vogg, M. C. *et al.* Stem cell-dependent formation of a functional anterior regeneration pole in
432 planarians requires Zic and Forkhead transcription factors. *Dev. Biol.* **390**, 136–148 (2014).
433 44. Bolger, A. M., Lohse, M. & Usadel, B. Trimmomatic: a flexible trimmer for Illumina sequence
434 data. *Bioinformatics* **30**, 2114–2120 (2014).
435 45. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013).
436 46. Grohme, M. A. *et al.* The genome of Schmidtea mediterranea and the evolution of core cellular
437 mechanisms. *Nature* **554**, 56–61 (2018).
438 47. Van der Auwera, G. A. *et al.* From FastQ Data to High-Confidence Variant Calls: The Genome
439 Analysis Toolkit Best Practices Pipeline. *Curr. Protoc. Bioinforma.* **43**, 11.10.1-11.10.33 (2013).
440 48. Emms, D. M. & Kelly, S. OrthoFinder: phylogenetic orthology inference for comparative
441 genomics. *Genome Biol.* **20**, 238 (2019).

49. Neiro, J., Sridhar, D., Dattani, A. & Aboobaker, A. Identification of enhancer-like elements defines regulatory networks active in planarian adult stem cells. 2022.02.03.479047 Preprint at https://doi.org/10.1101/2022.02.03.479047 (2022).

50. Martincorena, I. *et al.* Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041.e21 (2017).

Main text figures

Figure 1- Planarian natural history. Left: Showing position of planarians, as spiralians, in a broad metazoan phylogeny. Green circles show groups with representatives able to reproduce by fission. Right: obligate asexual planarian reproduction is solely based on fission and lacks a single cell bottleneck.

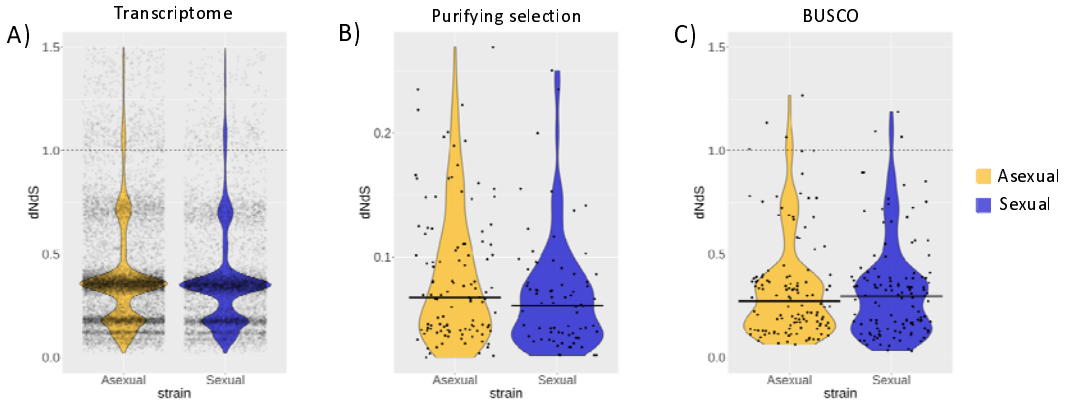


Figure 2- Limited impact of Muller's ratchet in asexual planarians. A) dNdS values across the entire transcriptomes. Median asexual dNdS = 0.353; Median sexual dNdS = 0.345. B) dNdS values of exons under purifying selection (solid black bars represent the medians). Median asexual dNdS = 0.067; Median sexual dNdS = 0.061. C) dNdS values of exons belonging to single copy planarian BUSCO genes. Median asexual dNdS = 0.287; Median sexual dNdS = 0.297.

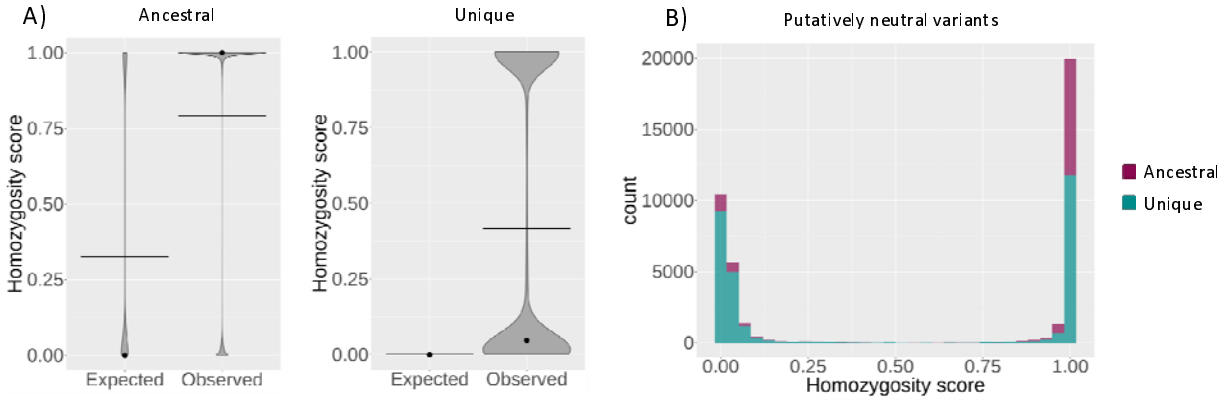


Figure 3 – High levels of asexual variant homozygosity. A) Ancestral and unique variants. Black dots indicate the median, while the bars indicate the mean. For the ancestral variants, the expected distribution was generated using rbinom in R, with a probability 0.32 for the ancestral homozygous variants (see methods). Homozygosity scores of both ancestral and unique asexual variants are higher than expected assuming no recombination. B) Histogram showing that the ancestral variants are more homozygous than variants arising uniquely in the asexual lineage.

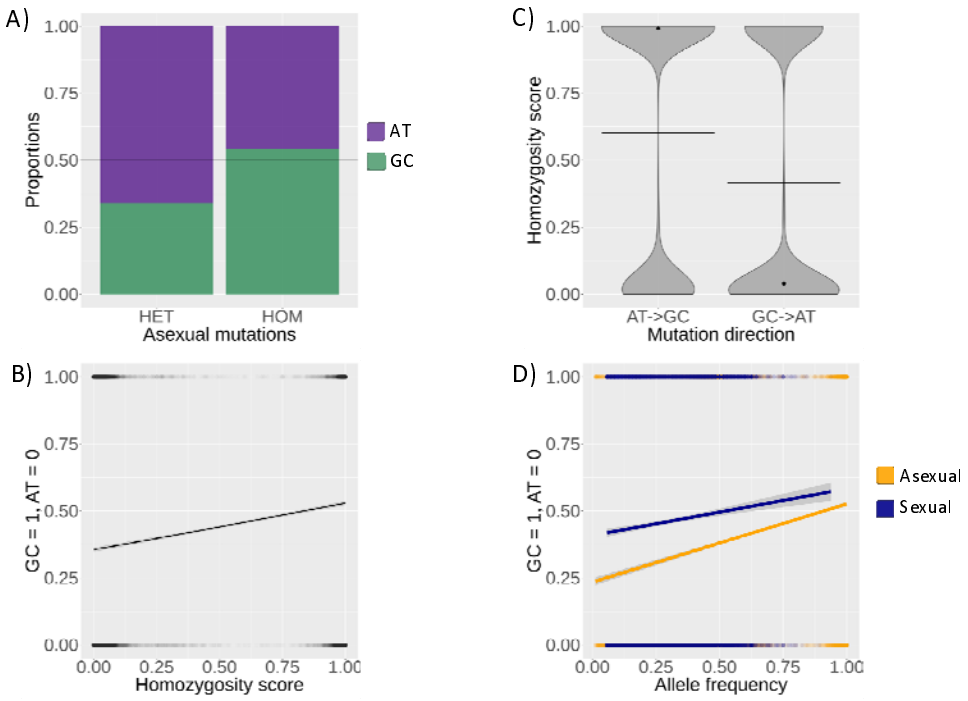


Figure 4 – Evidence for GC biased gene conversion. A) Showing that homozygous variants have a higher proportion of GC sites compared to heterozygous variants. HET GC proportion = 0.337; HOM GC proportion = 0.535. The black line marks the 0.5 line. B) Correlation between homozygosity scores and GC/AT identity. C) AT to GC mutations have higher homozygosity scores compared to GC to AT mutations. The directionality was defined by reference identity vs variant identity (reference to variant mutation direction). AT to GC HOM score = 0.988; GC to AT HOM score = 0.154. D) Correlation between alternative allele frequency and GC/AT identity.

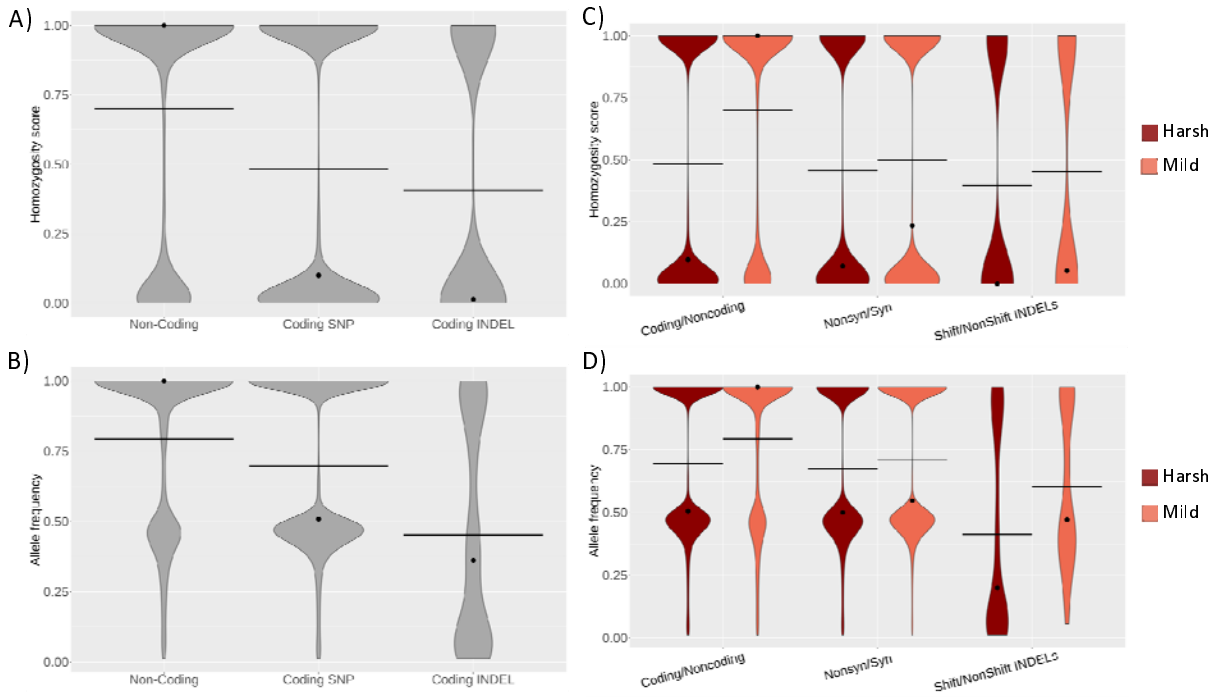
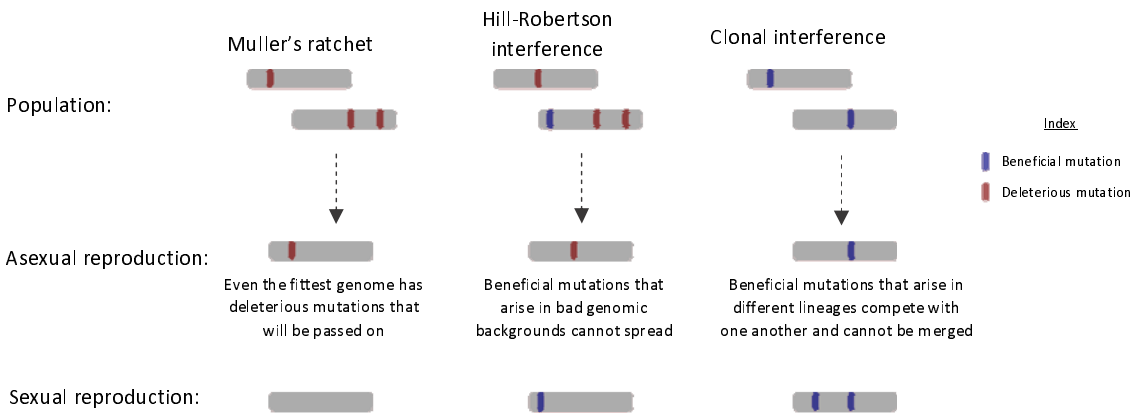
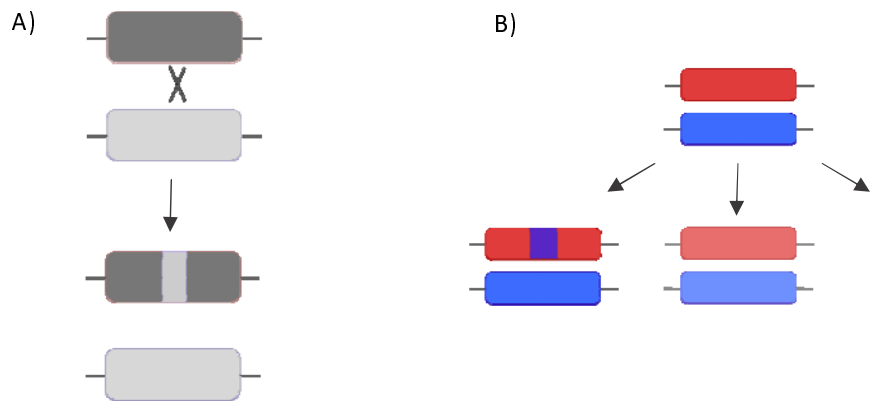


Figure 5 – Predicted deleterious mutations are less likely to be homozygous. Dots are medians, crossbars are means. A) Non-coding mutations more homozygous compared to coding SNPs which are more homozygous compared to coding INDELs. Median Non-Coding HOM score = 1; Median Coding SNP HOM score = 0.1; Median Coding INDEL HOM score = 0.014 B) More deleterious mutations have lower allele frequencies. Median Non-Coding AF = 1; Median Coding SNP AF = 0.509; Median Coding INDEL AF = 0.362 C) Non-coding mutations (predicted mild impact) are more homozygous (Median HOM score = 1) than coding mutations (predicted harsh impact) (Median HOM score = 0.097). Synonymous mutations (predicted mild impact) are more homozygous (Median HOM score = 0.234) than nonsynonymous mutations (predicted harsh impact) (Median HOM score = 0.071). Non-frameshifting indels (predicted mild impact) are more homozygous (Median HOM score = 0.053) compared to frameshifting indels (predicted harsh impact) (Median HOM score = 0). D) Median allele frequency is lower in harsher mutations. Coding vs noncoding (0.506 vs 1); nonsynonymous vs synonymous (0.5 vs 0.547); frameshifting vs non-frameshifting (0.2 vs 0.471).



Box1: The limitations of obligate asexual reproduction stem from the tight linkage of alleles in the genomes that lack recombination. Muller's ratchet describes how this lack of recombination leads to the accumulation of deleterious mutations over time. This is because even the fittest asexual genome will have some deleterious mutations that will be passed on to the next generation, while in sexual reproduction, genomes that only contain beneficial alleles can theoretically be generated and selected for. Hill-Robertson interference is a broad term describing the inefficient selection that occurs when alleles are tightly linked due to the lack of recombination. This means that deleterious mutations can interfere with the selection for beneficial mutations (and vice versa), while in sexual populations these mutations can be selected upon independently. Clonal interference describes the inefficacy resulting from the fact that two beneficial mutations in different lineages will compete with one another. Since asexual lineages don't recombine with one another, only the best mutation will persist, while in sexual populations a genome containing both beneficial mutations can be formed.

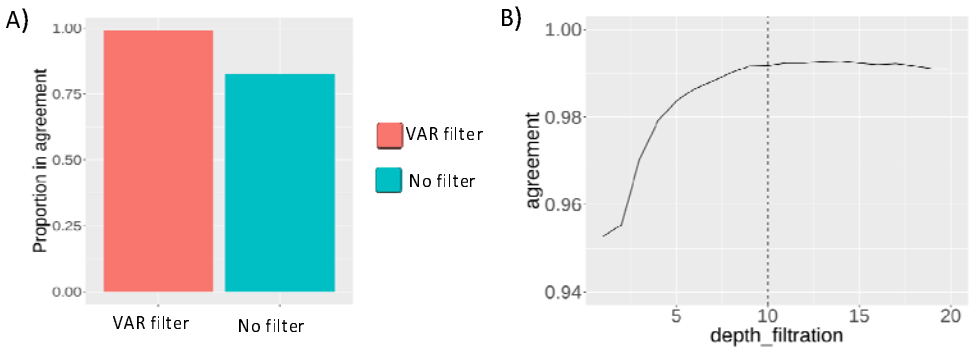
Box 1 – The limitations of asexual reproduction



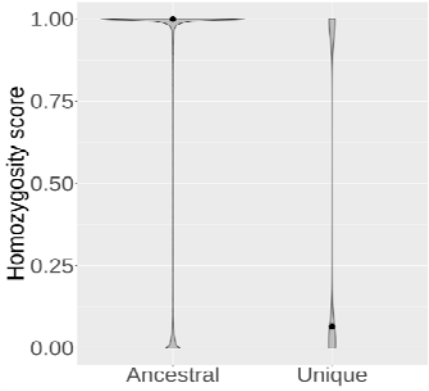
Box 2: Gene conversion and loss of heterozygosity. **A)** Diagram of gene conversion leading to a loss of heterozygosity. Gene conversion may occur during DNA damage repair if a homologous sequence is used as a template for repair. If there is a mismatch between the template and the region being repaired, this mismatch gets resolved by converting one sequence into the other. The resolution process is GC-biased meaning that mismatches will tend to be resolved in favour of GC nucleotides overwriting AT nucleotides. **B)** Diagram of how gene conversion can limit the progression of Muller's ratchet by removing deleterious mutations from some of the progeny. A heterozygous diploid is depicted by homologous regions of red and blue chromosomes. The red chromosome contains a region with a deleterious mutation. The progeny containing a gene conversion event that removes the deleterious mutation in the red chromosome has the highest fitness (left) and can be selected for, while gene conversion events leading to a duplication of the deleterious mutation (right) will be selected against.

Box 2: Gene conversion and loss of heterozygosity

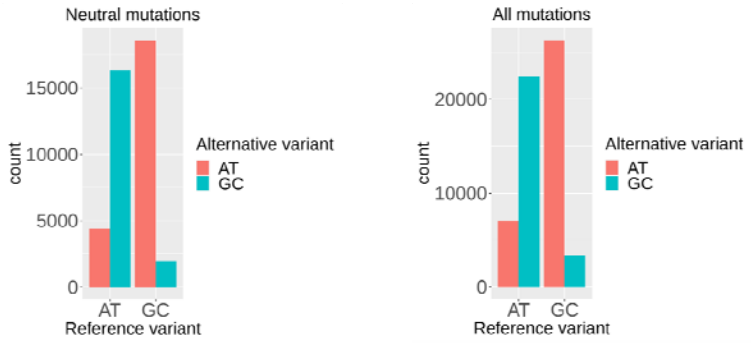
Supplementary



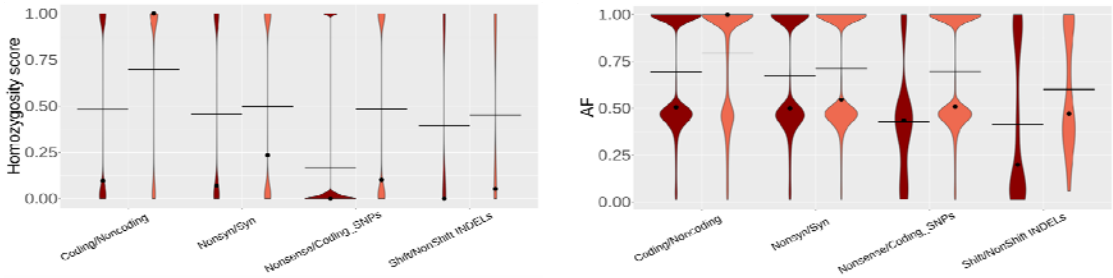
Supplementary Figure 1 – high agreement between ATAC and RNAseq results. A) Showing that when filtering for variants supported by at least 3 libraries, we can get an agreement of over 95%. B) Agreement between ATAC and RNAseq datasets as a function of the extent of depth filtration. At an average depth of 10 or above, there is over 99% agreement and more severe filtration does not drastically improve agreement.



Supplementary Figure 2 – Putatively neutral ancestral variants are more homozygous than putatively neutral uniquely asexual variants. Dots are median homozygosity scores.



Supplementary Figure 3 –There are more AT to GC mutations than vice versa. This suggests that the higher homozygosity of AT->GC mutations compared to GC->AT mutations is not driven by mutational patterns.



Supplementary Figure 4 – Nonsense mutations show a similar trend to that described in figure 5. However, we removed them from the main figure due to low sample size compared to the rest of the coding SNPs (46 vs 64,879).