Evolutionary implications and challenges of somatic immortality in planarians



Simon Kershenbaum

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2024

Declaration

I declare that this thesis was composed by myself and that the work contained herein is my own except where explicitly stated below. This work has not been submitted for any degree or professional qualification except as specified.

Details of work done by others:

Chapter 2:

- Sample preparation and DNA sequencing was done by Dr Danielle Ireland and Professor Eva-Maria Collins.
- Figure 1 – assessment of mortality and reproductive rates were done by Dr Danielle Ireland.
- Supplementary figure 4 - Karyotyping was done by the Collins laboratory

Simon Kershenbaum

Acknowledgments

First, I would like to express my gratitude to my supervisors, Aziz Aboobaker and Ashleigh Griffin. Your patience, guidance, and support have been invaluable throughout this PhD.

Thank you Manuel Jara-Espejo for your friendship, mentorship, and patient teaching.

Thank you to Danielle Ireland and Eva-Maria Collins, for your collaboration and comments on my second chapter.

Thank you to Tim Barraclough, Chris Wilson, and Tymek Pieszko. You all provided good advice and thoughtful comments on my projects.

Thank you to Peter Mulhair for being so supportive and helpful.

Finally, thank you to all members of my multiple lab groups for providing such a pleasant environment both professionally and socially. Special thanks to Juliet, Ming, Manuel, Zohar, Dianali, Abir, Cheng, Yian, Jakke, Ava, Stu and Casper, who supported me throughout, suffered my rants, and made this a memorable experience.

Finally, I am grateful to have had Juliet and Casper who have made my time here more meaningful.

Abstract

All transitions to obligate multicellularity evolved sexual reproduction *via* a single-cell bottleneck, suggesting that this strategy is important for a multicellular life. Sexual reproduction allows for effective selection, and single cell bottlenecks maintain the organism as the main level of selection. However, some species reproduce solely by fission, lacking both sex and single cell bottlenecks. So how do these fissiparous species persist in absence of sex and reproductive single cell bottlenecks? In this thesis, I use asexual planarians that reproduce solely by fission to explore the evolutionary and developmental implications of the fissiparous lifestyle. I start by asking how they overcome some of the challenges of asexuality such as Muller's ratchet and find strong signatures of gene conversion which may limit mutation accumulation in their genomes. I then ask how diversity is generated in asexual planarians and whether it accumulates in their soma over evolutionary time due to the lack of a zygotic bottleneck. By assessing the genomes of individual planarians from different populations I describe diversity in asexual planarians, with two genetically distinct strains present in laboratories across the world and find evidence that different strains contain different levels of somatic diversity. Finally, I ask what transcriptional changes follow the transition to a fissiparous lifestyle and whether these can be used to isolate key genes in sexual germline development. By isolating and knocking down transcription factors whose expression differs consistently between sexual and asexual strains of different planarian species, I find two homeobox genes with novel roles in germline regulation in planarians. Overall, I find that planarians that reproduce by fission are able to bypass some of the predicted limitations of this strategy, and that its examination is useful for both questions in evolutionary and developmental biology.

Contents

Introduction 1

Chapter 1- Gene conversion limits the cost of asexuality in somatically immortal worms 20

Chapter 2 - Diversity within and between individual planarians in the absence of sex and a single cell bottleneck. 58

Chapter 3 - Comparative approach identifies two novel male germline regulators in planarians 96

Discussion 128

Introduction

"The ship on which Theseus sailed… was preserved by the Athenians… They took away the old timbers as they decayed, and put new and sound ones in their places, so that the vessel became a standing illustration for the philosophers in the logical question of things that grow; one side holding that the ship remained the same, and the other contending that it was not the same."

—Plutarch, Life of Theseus 23.1

**An analogy to the concept of individuality in fissiparous organisms. When does one individual end and another begin in the absence of a zygotic single cell bottleneck?*

Most multicellular life can reproduce sexually *via* a zygotic single cell bottleneck (Judson and Normark, 1996). This true for all independent transitions to obligate multicellularity, representing highly convergent strategies (Fisher et al., 2013; Knoll, 2011). Such a consistent reliance on sex and a single cell bottleneck suggests that these two strategies play crucial roles in the evolution or maintenance of multicellular species. However, there are some exceptions to the rule which can be studied to better understand why sex and a zygotic single cell bottleneck are so common, and what happens when they are both absent in a species.

Sexual reproduction is a costly, raising the question of why it is so pervasive (Colegrave, 2012; Lehtonen et al., 2012). For example, sexual recombination breaks up beneficial haplotypes that are well adapted to their environment (Charlesworth and Barton, 1996). Additionally, sex often involves males which do not directly convert resources into offspring, while in asexual populations all individuals reproduce (Gibson et al., 2017). Finally, sex

requires a mate, with mate acquisition often being very costly (Lehtonen et al., 2012). Yet sex is pervasive while obligate asexuality is rare, with most species reproducing via facultative or obligate sexual reproduction and very few via purely asexual means, suggesting that despite these costs, sexual reproduction is beneficial in the long run. This is thought to be in part due to the intrinsic value of recombination has in increasing the efficacy of selection (Felsenstein, 1974). For example, in the absence of sexual recombination, beneficial mutations arising in different individuals cannot be merged into a single genome, and instead compete with one another, leading to a process called clonal interference which reduces the adaptability of asexual populations (Miralles et al., 1999). Additionally, in the absence of recombination the probability of beneficial mutations reaching fixation, and conversely deleterious mutations being purged, depends on the genetic background in which it occurred (de Visser and Elena, 2007; Felsenstein, 1974). This increases the role of genetic drift and reduces the efficacy of selection in asexual populations, a problem called the Hill-Robertson effect (Comeron et al., 2008; Hill and Robertson, 1966). Ultimately, the lack of recombination may lead to the accumulation of deleterious mutations over time in a process called "Muller's ratchet", because only back mutation can purge fixed deleterious mutations from the population, a relatively slow process of removing deleterious mutations from the genome (Lynch et al., 1993; Muller, 1964). Taken together, it is thought that the lack of recombination in asexual lineages reduces their adaptability and increases their mutational load, explaining the pervasiveness of sex and rarity of obligate asexual reproduction in nature.

That said, some asexual lineages persist and even diversify over long periods of time, presumably limiting the cost of asexuality (Brandt et al., 2017; Flot et al., 2013; Judson and Normark, 1996; Schön et al., 2009; Swanstrom et al., 2011). For example, bdelloid rotifers are thought to have persisted as asexuals for over 50 million years and have diversified into over 460 species (Segers, 2007; Tang et al., 2014). This has led people to ask how these so

called “ancient asexuals” have been able to persist despite the limitations of asexual reproduction. Perhaps the simplest explanation put forward is that they are not obligately asexual and partake in rare cryptic sex, which has been shown in species previously thought to be asexual (Boyer et al., 2021; Freitas et al., 2023). Some evidence of allele sharing was even suggested to imply cryptic sex in the bdelloid *Macrotrachela quadricornifera*, however, other mechanisms can also explain these patterns and so may not be conclusive evidence of sexual reproduction (Laine et al., 2022; Wilson et al., 2024). Beyond the possibility of rare sex, other suggestions for mechanisms allowing asexual lineages to persist have been put forward. For example, both bdelloid rotifers and darwinulid ostracods exhibit excess homozygosity that is suggested to be the result of gene conversion, a form of asexual recombination which may allow asexual lineages to more efficiently purge deleterious mutations form their genome (Flot et al., 2013; Schön et al., 2009). However, such mechanisms do not solve all the challenges associated with asexuality, such as clonal interference. Thus, despite significant effort, there seems to be no consensus about the adaptations which allow long-term asexual lineages to persist.

Gene conversion occurs when a double strand break is repaired using a homologous chromosome which may contain alternative alleles to those being repaired (Birdsell, 2002; Chen et al., 2010; Marais, 2003). This can lead to base mismatches which are then repaired by the base mismatch repair mechanisms which tend to be GC biased, such that AT alleles are more likely to be removed than GC alleles (Birdsell, 2002). GC biased gene conversion was shown to take place in both mitotic and meiotic events, and can be detected by GC biased losses of heterozygosity (Birdsell, 2002; Chen et al., 2010; Jia et al., 2021). For example, in mitotic yeast cells, it was shown that base mismatch repair led to the repair of heteromismatches into GC alleles 57.9% of the time compared to AT alleles 42.1% of the times (Birdsell, 2002). Since gene conversion is ubiquitous and does not require sex to unlink

loci and remove deleterious mutations from the population, it has been proposed as a potential mechanism by which asexual species may limit the cost of asexuality (Flot et al., 2013; Khakhlova and Bock, 2006; Wang et al., 2022). However, relatively few asexual species have been assessed for their gene conversion rates, and in some cases high gene conversion rates do not correlate with reduced rates of mutation accumulation (Omilian et al., 2006; Tucker et al., 2013), thus there is no consensus around whether gene conversion is an effective mechanism to limit the cost of asexuality and more research is needed to assess this.

The majority of research into the persistence of asexual lineages has focused on parthenogenic species, with very little attention being given to species that reproduce solely by fission. However, reproduction by fission and other somatic mechanisms differ significantly from parthenogenesis. For example, some parthenogenic species can reproduce via automictic process which involve large scale asexual recombination which may help purge deleterious mutations, while in fissiparous species this mechanism is not available (Engelstädter, 2017). Additionally, fission, budding, and other forms of somatic reproduction, often involve abundant pluripotent stem cells that are inherited from one generation to the next, potentially increasing the available genetic variation in the population (Bosch et al., 2010; Martínez, 1998; Wagner et al., 2011). This may help fissiparous species deal with the reduced efficacy of selection associated with asexuality, but since very few studies have been conducted on obligately fissiparous asexuals it is currently unknown. Finally, comparing radically different forms of asexuality may be helpful for isolating the fundamental mechanisms that help asexuals persist, rather than lineage specific solutions. This all suggests that we should focus on understanding how different forms of asexuality persist rather than focusing only on parthenogenic asexuals.

Additionally, reproduction by fission presents challenges that are not directly related to the lack of recombination. For example, in animals that reproduce by fission, somatic cells can be inherited by the next generation (Domazet-Lošo et al., 2014). This provides more time for somatic cells to acquire selfish adaptations which can result in an increased load of cancerous cells in the next generation (Aktipis et al., 2015; Howe et al., 2022). More generally, the lack of frequent single cell bottlenecks could result in increased somatic diversity, which over evolutionary time might select for reduced cooperation between cells (Howe et al., 2022). While this has been shown in multiple aggregative multicellular species, including the ascidian *Botryllus*, it has not been investigated in fissiparous animals (Buss, 1982; Stoner et al., 1999). Presumably these challenges are significant since all independent transitions to obligate multicellularity involve periodic single cell bottlenecks (Fisher et al., 2013; Knoll, 2011). Despite their oddity, obligately fissiparous species have not yet been studied in the context of how they limit the accumulation of intra-organismal diversity, and the potential for accumulating selfish cell adaptations. Thus, this understudied mode of reproduction needs to be better investigated to understand how fissiparous species deal with the absence of frequent single cell bottlenecks.

These challenges associated with fissiparous reproduction may apply to a diverse set of reproductive strategies ranging from budding to vegetative growth which are found abundantly across the tree of life (Edwards et al., 1990; Howe et al., 2022; Leria et al., 2019; Reusch et al., 2021; Simberloff and Leppanen, 2019). For example, many plants reproduce vegetatively without passing through frequent single cell bottlenecks, leading to the accumulation of somatic diversity. In some cases, this somatic diversity has been shown to be beneficial in plants as it can generate variation that may drive pest resistance, leading to locally increased fitness around where beneficial somatic mutations occur (Edwards et al., 1990; Simberloff and Leppanen, 2019). One striking example of such mosaic resistance in

plants is found in yellow box trees (*Eucalyptus melliodora*) where somatic variation is thought underly branch specific variation in Christmas beetle (*Anoplognathus montanus*) herbivory (Edwards et al., 1990). However, the implications of somatic variation may fundamentally differ between plants and animals due to the potential for cell migration and mixing (Aktipis et al., 2015). This is because in plants a deleterious somatic mutation will harm the local area where it occurred, and is unlikely to spread to the rest of the soma due to the lack of widespread cell migration in plants (Aktipis et al., 2015; Simberloff and Leppanen, 2019). On the other hand, in animals cell migration means that cells with deleterious somatic mutations may migrate to other tissues, and more easily be transmitted to all offspring (Aktipis et al., 2015; Domazet-Lošo et al., 2014). This has been document with cancer cells in hydra being transmitted to the next generation through the bud (Domazet-Lošo et al., 2014). Despite the differences in how somatic variation may impact fitness in plants and animals, very little work has been done to understand the extent and implication of somatic variation in fissiparous animals. Previous work where the somatic diversity of planarians with evidence of fissiparous reproduction was compared to those with no evidence of fission has shown increased levels of somatic diversity in fissiparous populations (Leria et al., 2019). However, the implications of this increased diversity are not understood, and no work has been done on animal species suggested to reproduce solely by fission over evolutionary time. Thus, despite our understanding of how vegetative growth drives somatic diversity in plants, very little is understood about how this occurs in animals and what the implications of increased somatic diversity are on organismal fitness.

My aim in this thesis is to enhance our understanding of reproductive strategies through an examination of fissiparous species. Planarian flatworms are a diverse group of platyhelminths that contain species with obligately asexual fissiparous strains making them ideal models to study the implications of reproduction by fission. *Schmidtea mediterranea* is widely studied

for its regenerative abilities and contains both a hermaphroditic sexual strain and an asexual strain that reproduces solely by fission. The asexual strain is suggested to have a crown age of around 600,000 years, and so is likely to have been reproducing asexually for millions of generations (Lázaro et al., 2011). In both the sexual and asexual strains of *Schmidtea mediterranea*, adult pluripotent stem cells called neoblasts are responsible for replenishing all the cells in the body, and a single asexual neoblast has been experimentally shown to be sufficient for whole body regeneration of the asexual strain (Wagner et al., 2011). In the asexual strain, these neoblasts replace the germline as the only heritable information from one generation to the next, and due to the lack of a natural single cell bottleneck it is possible for multiple neoblast clones to be transmitted to the next generation (Solana, 2013). This has led some to suggest that obligately asexual planarians may accumulate diversity in their body, and that intra-organismal selection may drive the evolution of selfish adaptations in neoblasts (Fields and Levin, 2018; Howe et al., 2022). Additionally, these asexual planarians can be thought of as somatically immortal, since they must maintain a functional soma continuously over evolutionary time, raising philosophical questions about the individuality of these worms (Fields and Levin, 2018; Solana, 2013). While Field and Levin 2018, sate that neoblasts should be considered the main unit of selection in planarians, since the survival of neoblasts depends on the fitness of the worm in which they are found, it is unclear how well neoblast-level adaptations will be able to spread in the population if they reduce the fitness of the planarian as a whole. Thus more empirical, modelling, and conceptual work is needed to understand the degree to which neoblasts can be considered the main unit of selection in planarians. Taken together, asexual planarians provide us with the opportunity to better understand how obligately fissiparous species deal with the classical challenges of asexuality, as well as the unique difficulties of maintaining an immortal soma over evolutionary time.

In addition to these evolutionary questions, asexual planarians provide us with an opportunity to better understand the genes involved in sexual reproduction (Iii et al., 2010; Issigonis and Newmark, 2019; Wang et al., 2010; Zayas et al., 2005). Asexual strains can be compared to their sexual counterparts to isolate genes that are more highly expressed in sexual strains, which may underly various aspects of sexual reproduction (Iii et al., 2010; Wang et al., 2010; Zayas et al., 2005). Planarians are particularly important species for understanding the genetic mechanisms by which animal germlines were ancestrally developed (Issigonis and Newmark, 2019). This is because the ancestral metazoan is thought to have developed its germlines by means of cell-cell communication, or "induction", which is similar to the way planarians develop their germlines today (Donoughe et al., 2014; Extavour and Akam, 2003; Issigonis and Newmark, 2019; Johnson et al., 2003). However, most of what we know about germline development comes form model organisms such as *Drosophila*, *Xenopus*, *C.elegans*, and others, which all develop their germline by a derived process using maternal determinants (Seervai and Wessel, 2013). While these models taught us a lot about the molecular biology of germ cells, we have relied mostly on mice for our knowledge about the mechanisms of germline induction, which may or may not be broadly conserved (Issigonis and Newmark, 2019; Seervai and Wessel, 2013). The recent emergence of planarians as key models in inductive germline development, together with other invertebrate species like crickets, is now allowing us to better understand the genes that ancestrally regulated inductive germline development in the ancestral metazoan (Donoughe et al., 2014; Handberg-Thorsager and Saló, 2007; Sato et al., 2006; Wang et al., 2007) (for more information see Chapter 3).

In summary, asexual planarians represent an understudied mode of reproduction which raises many questions including:

i) How do they maintain asexuality in the face of Muller's ratchet and other classical challenges of asexuality?

ii) To what degree do clonal planarians generate diversity in their population without sex? And how do they avoid the accumulation of somatic diversity given the lack of a single cell bottleneck?

iii) Are there conserved genes whose expression differs consistently between sexual and asexual strains? And can this teach us about what genes may have deeply conserved roles in germline induction in animals?

In this thesis, I aim to address these questions and lay the groundwork for further investigation into this peculiar mode of reproduction. Specifically, I will:

i) Test whether asexual planarians accumulate deleterious mutations as predicted by the reduced efficacy of selection in asexual populations and assess how they might avoid this. (Chapter 1)

ii) Test how much diversity is present in asexual populations of planarians, and whether they accumulate significant levels of diversity within their body. (Chapter 2)

iii) Identify transcription factors whose expression differs between sexual and asexual strains in two planarian species and knock them down to better understand their role in germline development. (Chapter 3)

Bibliography

Aktipis, C. A., Boddy, A. M., Jansen, G., Hibner, U., Hochberg, M. E., Maley, C. C. and Wilkinson, G. S. (2015). Cancer across the tree of life: cooperation and cheating in multicellularity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140219.

Birdsell, J. A. (2002). Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution* **19**, 1181–1197.

Bosch, T. C. G., Anton-Erxleben, F., Hemmrich, G. and Khalturin, K. (2010). The Hydra polyp: Nothing but an active stem cell community. *Development, Growth & Differentiation* **52**, 15–25.

Boyer, L., Jabbour-Zahab, R., Mosna, M., Haag, C. R. and Lenormand, T. (2021). Not so clonal asexuals: Unraveling the secret sex life of Artemia parthenogenetica. *Evol Lett* **5**, 164–174.

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. and Bast, J. (2017). Effective purifying selection in ancient asexual oribatid mites. *Nat Commun* **8**, 873.

Buss, L. W. (1982). Somatic cell parasitism and the evolution of somatic tissue compatibility. *Proceedings of the National Academy of Sciences* **79**, 5337–5341.

Charlesworth, B. and Barton, N. H. (1996). Recombination load associated with selection for increased recombination. *Genetics Research* **67**, 27–41.

Chen, J.-M., Férec, C. and Cooper, D. N. (2010). Gene Conversion in Human Genetic Disease. *Genes (Basel)* **1**, 550–563.

Colegrave, N. (2012). The evolutionary success of sex. *EMBO Rep* **13**, 774–778.

Comeron, J. M., Williford, A. and Kliman, R. M. (2008). The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19–31.

de Visser, J. A. G. M. and Elena, S. F. (2007). The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* **8**, 139–149.

Domazet-Lošo, T., Klimovich, A., Anokhin, B., Anton-Erxleben, F., Hamm, M. J., Lange, C. and Bosch, T. C. G. (2014). Naturally occurring tumours in the basal metazoan Hydra. *Nat Commun* **5**, 4222.

Donoughe, S., Nakamura, T., Ewen-Campen, B., Green, D. A., Henderson, L. and Extavour, C. G. (2014). BMP signaling is required for the generation of primordial germ cells in an insect. *Proceedings of the National Academy of Sciences* **111**, 4133–4138.

Edwards, P. B., Wanjura, W. J., Brown, W. V. and Dearn, J. M. (1990). Mosaic resistance in plants. *Nature* **347**, 434–434.

Engelstädter, J. (2017). Asexual but Not Clonal: Evolutionary Processes in Automictic Populations. *Genetics* **206**, 993–1009.

Extavour, C. G. and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**, 5869–5884.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* **78**, 737–756.

Fields, C. and Levin, M. (2018). Are planaria individuals? What regenerative biology is telling us about the nature of multicellularity. *Evol Biol* **45**, 237–247.

Fisher, R. M., Cornwallis, C. K. and West, S. A. (2013). Group formation, relatedness, and the evolution of multicellularity. *Current Biology* **23**, 1120–1125.

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G. J., Hejnol, A., Henrissat, B., Koszul, R., Aury, J.-M., et al. (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer Adineta vaga. *Nature* **500**, 453–457.

Freitas, S., Parker, D. J., Labédan, M., Dumas, Z. and Schwander, T. (2023). Evidence for cryptic sex in parthenogenetic stick insects of the genus Timema. *Proceedings of the Royal Society B: Biological Sciences* **290**, 20230404.

Gibson, A. K., Delph, L. F. and Lively, C. M. (2017). The two-fold cost of sex: Experimental evidence from a natural system. *Evol Lett* **1**, 6–15.

Handberg-Thorsager, M. and Saló, E. (2007). The planarian nanos-like gene Smednos is expressed in germline and eye precursor cells during development and regeneration. *Dev Genes Evol* **217**, 403–411.

Hill, W. G. and Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research* **8**, 269–294.

Howe, J., Rink, J. C., Wang, B. and Griffin, A. S. (2022). Multicellularity in animals: The potential for within-organism conflict. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120457119.

Iii, J. J. C., Hou, X., Romanova, E. V., Lambrus, B. G., Miller, C. M., Saberi, A., Sweedler, J. V. and Newmark, P. A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLOS Biology* **8**, e1000509.

Issigonis, M. and Newmark, P. A. (2019). Chapter four - From worm to germ: Germ cell development and regeneration in planarians. In *Current Topics in Developmental Biology* (ed. Lehmann, R.), pp. 127–153. Academic Press.

Jia, X., Zhang, Q., Jiang, M., Huang, J., Yu, L., Traw, M. B., Tian, D., Hurst, L. D. and Yang, S. (2021). Mitotic gene conversion can be as important as meiotic conversion in driving genetic variability in plants and other species without early germline segregation. *PLOS Biology* **19**, e3001164.

Johnson, A. D., Crother, B., White, M. E., Patient, R., Bachvarova, R. F., Drum, M. and Masi, T. (2003). Regulative germ cell specification in axolotl embryos: a primitive trait conserved in the mammalian lineage. *Phil. Trans. R. Soc. Lond. B* **358**, 1371–1379.

Judson, O. P. and Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution* **11**, 41–46.

Khakhlova, O. and Bock, R. (2006). Elimination of deleterious mutations in plastid genomes by gene conversion. *The Plant Journal* **46**, 85–94.

Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet. Sci.* **39**, 217–239.

Laine, V. N., Sackton, T. B. and Meselson, M. (2022). Genomic signature of sexual reproduction in the bdelloid rotifer *Macrotrachella quadricornifera*. *Genetics* **220**, iyab221.

Lázaro, E. M., Harrath, A. H., Stocchino, G. A., Pala, M., Baguñà, J. and Riutort, M. (2011). *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evolutionary Biology* **11**, 274.

Lehtonen, J., Jennions, M. D. and Kokko, H. (2012). The many costs of sex. *Trends in Ecology & Evolution* **27**, 172–178.

Leria, L., Vila-Farré, M., Solà, E. and Riutort, M. (2019). Outstanding intraindividual genetic diversity in fissiparous planarians (Dugesia, Platyhelminthes) with facultative sex. *BMC Evolutionary Biology* **19**, 130.

Lynch, M., Bürger, R., Butcher, D. and Gabriel, W. (1993). The Mutational meltdown in asexual populations. *Journal of Heredity* **84**, 339–344.

Marais, G. (2003). Biased gene conversion: implications for genome and sex evolution. *Trends in Genetics* **19**, 330–338.

Martínez, D. E. (1998). Mortality patterns suggest lack of senescence in hydra. *Exp Gerontol* **33**, 217–225.

Miralles, R., Gerrish, P. J., Moya, A. and Elena, S. F. (1999). Clonal Interference and the Evolution of RNA Viruses. *Science* **285**, 1745–1747.

Muller, H. J. (1964). The relation of recombination to mutational advance. 8.

Omilian, A. R., Cristescu, M. E. A., Dudycha, J. L. and Lynch, M. (2006). Ameiotic recombination in asexual lineages of Daphnia. *Proceedings of the National Academy of Sciences* **103**, 18638–18643.

Reusch, T. B. H., Baums, I. B. and Werner, B. (2021). Evolution via somatic genetic variation in modular species. *Trends in Ecology & Evolution* **36**, 1083–1092.

Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S. and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed

by the expression of nanos-related gene in planarians. *Development, Growth & Differentiation* **48**, 615–628.

Schön, I., Rossetti, G. and Martens, K. (2009). Darwinulid ostracods: Ancient asexual scandals or scandalous gossip? In l*ost sex: The evolutionary biology of parthenogenesis* (ed. Schön, I.), Martens, K.), and Dijk, P.), pp. 217–240. Dordrecht: Springer Netherlands.

Seervai, R. N. H. and Wessel, G. M. (2013). Lessons for Inductive Germline Determination. *Mol Reprod Dev* **80**, 590–609.

Segers, H. (2007). Annotated checklist of the rotifers (Phylum Rotifera), with notes on nomenclature, taxonomy and distribution. *Zootaxa* **1564**, 1–104.

Simberloff, D. and Leppanen, C. (2019). Plant somatic mutations in nature conferring insect and herbicide resistance. *Pest Management Science* **75**, 14–17.

Solana, J. (2013). Closing the circle of germline and stem cells: the primordial stem cell hypothesis. *EvoDevo* **4**, 2.

Stoner, D. S., Rinkevich, B. and Weissman, I. L. (1999). Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. *Proceedings of the National Academy of Sciences* **96**, 9148–9153.

Swanstrom, J., Chen, K., Castillo, K., Barraclough, T. G. and Fontaneto, D. (2011). Testing for evidence of inefficient selection in bdelloid rotifers: do sample size and habitat differences matter? *Hydrobiologia* **662**, 19–25.

Tang, C. Q., Obertegger, U., Fontaneto, D. and Barraclough, T. G. (2014). Sexual species are separated by larger genetic gaps than asexual species in rotifers. *Evolution* **68**, 2901–2916.

Tucker, A. E., Ackerman, M. S., Eads, B. D., Xu, S. and Lynch, M. (2013). Population-genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. *Proc Natl Acad Sci U S A* **110**, 15740–15745.

Wagner, D. E., Wang, I. E. and Reddien, P. W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811–816.

Wang, Y., Zayas, R. M., Guo, T. and Newmark, P. A. (2007). Nanos function is essential for development and regeneration of planarian germ cells. *Proc Natl Acad Sci U S A* **104**, 5901–5906.

Wang, Y., Stary, J. M., Wilhelm, J. E. and Newmark, P. A. (2010). A functional genomic screen in planarians identifies novel regulators of germ cell development. *Genes Dev* **24**, 2081–2092.

Wang, Y., Li, X.-Y., Xu, W.-J., Wang, K., Wu, B., Xu, M., Chen, Y., Miao, L.-J., Wang, Z.-W., Li, Z., et al. (2022). Comparative genome anatomy reveals evolutionary insights into a unique amphitriploid fish. *Nat Ecol Evol* **6**, 1354–1366.

Wilson, C. G., Pieszko, T., Nowell, R. W. and Barraclough, T. G. (2024). Recombination in bdelloid rotifer genomes: asexuality, transfer and stress. *Trends in Genetics* **40**, 422–436.

Zayas, R. M., Hernández, A., Habermann, B., Wang, Y., Stary, J. M. and Newmark, P. A. (2005). The planarian *Schmidtea mediterranea* as a model for epigenetic germ cell specification: Analysis of ESTs from the hermaphroditic strain. *Proc Natl Acad Sci U S A* **102**, 18491–18496.

Chapter 1- Gene conversion limits the cost of asexuality in somatically immortal worms

Abstract

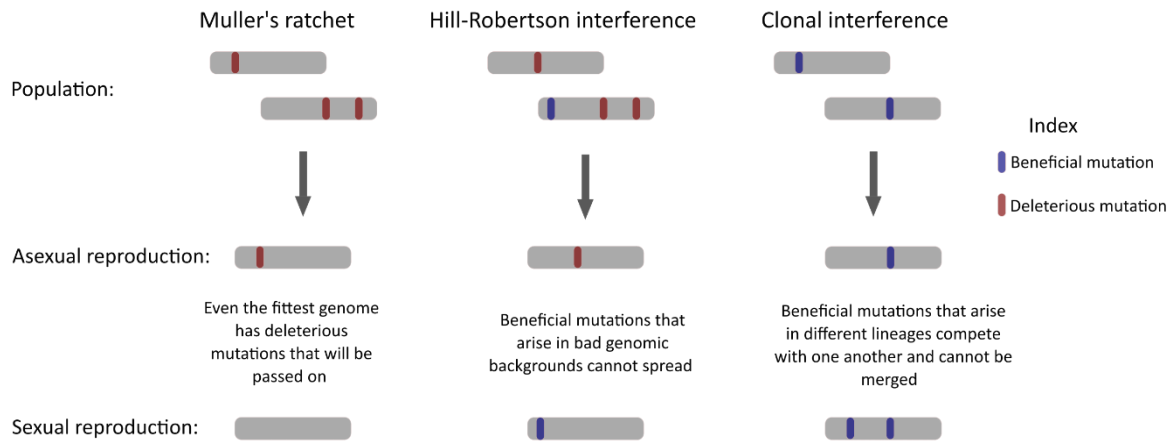
Most multicellular organisms reproduce sexually despite the costs associated with sexuality. This has been explained as the result of selection favouring the ability to recombine the genome. The lack of recombination in asexual species constrains their adaptability and leads to the accumulation of deleterious mutations, ultimately increasing their risk of extinction. Nonetheless, successful asexual life histories persist among multicellular organisms, and explanatory mechanisms which may help limit the cost of asexuality remain enigmatic. In search of these mechanisms, we looked at that the molecular evolutionary changes in the sexual and obligately asexual strains of the planarian flatworm, *Schmidtea mediterranea*. We find that the accumulation of deleterious mutations in highly conserved genes 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 beneficial alleles and purification of deleterious mutations. Taken together, our analysis identifies gene conversion as a mechanism which may contribute to the maintenance of asexuality in an obligately fissiparous metazoan.

Introduction

Most multicellular life reproduces sexually, while obligate asexuality is exceedingly rare (Judson and Normark, 1996). This pattern persists despite the high costs of sex, including the "two-fold" cost of sex (Gibson et al., 2017), the breaking up of beneficial haplotypes (Charlesworth and Barton, 1996), and the broad costs associated with securing a mate (Lehtonen et al., 2012). 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 (Lynch et al., 1993). However, some ancient asexual organisms exist, which raises the question of how they avoid the limitations of asexuality (Bast et al., 2018; Butlin et al., 1998; Judson and Normark, 1996; Lázaro et al., 2011; Wilson et al., 2018).

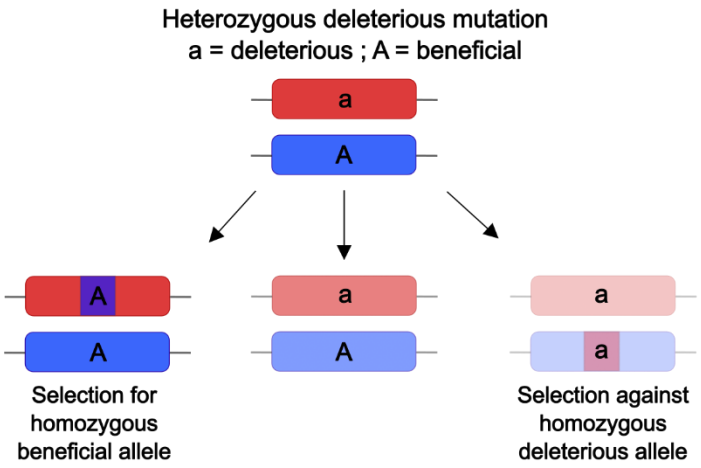
The potential disadvantages of asexual reproduction are driven by the tight linkage between alleles in an asexual genome (Box. 1) (Comeron et al., 2008; Felsenstein, 1974; Hill and Robertson, 1966). This linkage means that selection cannot act on individual alleles or allele combinations, but instead must act on the genomes of individuals as a whole - a problem known as "The Hill-Robertson effect" (Comeron et al., 2008; de Visser and Elena, 2007; Hill and Robertson, 1966). This reduces the adaptability of asexual populations, which has been shown to be particularly problematic in the context of changing environments and biotic competition (Hamilton et al., 1990; Lively and Morran, 2014; Vergara et al., 2014). Additionally, the inability to remove deleterious mutations from the next generation through recombination results in the accumulation of deleterious mutations over time in a process called "Muller's ratchet" (Lynch et al., 1993; Muller, 1964). This occurs since even the fittest individual in an asexual population contains some deleterious mutations, and as this fittest genome reaches fixation, so do its deleterious mutations, leading to a ratchet-like process of mutation accumulation (Lynch et al., 1993; Muller, 1964). 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 (Felsenstein, 1974).



Box. 1: 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 fittest variant will persist, while in sexual populations a genome containing both variants can be formed.

However, some obligate asexual species persist over evolutionary time (Barraclough et al., 2007; Bast et al., 2018; Lázaro et al., 2011; Schön et al., 2009), suggesting that they may have evolved mechanisms to limit the impact of long-term asexuality. Such ancient asexuals include bdelloid rotifers, oribatid mites, darwinulid ostracods and more, all of which have likely been asexual for millions of generations (Brandt et al., 2017; Brandt et al., 2021; Flot et al., 2013; Mark Welch and Meselson, 2000; Schön et al., 2003; Schön et al., 2009). To date, there is no consensus regarding the methods by which they avoid the costs associated with asexuality. One suggested mechanism is gene conversion (Box. 2): a form of mitotic recombination in which one allele overwrites another (Butlin et al., 1998; Flot et al., 2013; Schön et al., 2009). While deleterious and beneficial alleles are equally likely to get overwritten in the process of gene conversion, selection may be able to act to remove the deleterious gene conversion events and drive the spread of beneficial alleles (Box. 2). Ancient asexuals such as bdelloid rotifers (Flot et al., 2013) and darwinulid ostracods (Schön et al., 2009) exhibit excess homozygosity, which has been proposed to be the result of gene conversion, leading the authors to suggest a role for gene conversion in the long term survival of these asexuals. However, evidence from other species suggests that gene conversion may not be sufficient to limit mutation accumulation, since young asexual lineages of Daphnia show relatively high rates of gene conversion while still showing a significant accumulation of deleterious mutations (Omilian et al., 2006; Tucker et al., 2013). In neither case has it been shown that selection acts upon gene conversion events to remove deleterious mutations and fix beneficial alleles in the population. Taken together, there is currently no consensus about the mechanisms that limit the costs of asexuality in ancient asexuals, and we need to assess more groups of long-term asexuals with diverse modes of asexual reproduction to better understand how ancient asexuals might overcome the limitations of asexuality.



Box. 2: Gene conversion as a method to remove deleterious mutations. 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 which 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 homozygous deleterious mutation (right) will be selected against. The most deleterious genome (a/a) is indicated by high transparency, followed by a/A, and A/A respectively.

The planarian flatworm *S. mediterranea* is a highly regenerative animal that has both a sexual and an asexual strain (Tan et al., 2012). The sexual strain is a cross fertilizing hermaphrodite and the obligate asexual strain lacks reproductive organs, reproducing solely by fission followed by whole body regeneration (Newmark et al., 2008; Saberi et al., 2016) (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 limitations of asexuality (Lázaro et al., 2011; Sahu et al., 2017; Tan et al., 2012). The fissiparous nature of asexual planarian reproduction is based on whole body regeneration that is mediated by somatic, immortal, pluripotent adult stem cells called neoblasts (Wagner et al., 2011). 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 (Fields and Levin, 2018; Newmark and Sánchez Alvarado, 2000), dependent upon when the asexual strain arose (Lázaro et al., 2011). Such a high number of asexual generations is equivalent to some of the previously studied ancient asexuals but represents an understudied mode of asexual reproduction in metazoans. This fissiparous lifestyle and the presence of a closely related sexual strain which allows for the direct genetic comparison of sexual and asexual strains, makes *S. mediterranea* a valuable species for understanding how such asexual lineages might persist over evolutionary time.

In this study, we investigate how asexual planarians may limit the costs of asexuality. Specifically, we test whether planarian flatworms are able to escape the costs of asexuality through gene conversion. We find that asexual planarians limit the impact of Muller's ratchet, and that they show GC-biased excess homozygosity indicative of gene conversion. Furthermore, we find that mutations that are predicted to be more deleterious are less homozygous, suggesting that gene conversion events can be acted upon by selection to fix the beneficial allele in the population and remove deleterious mutations. Thus, we find evidence that gene conversion mediated removal of deleterious mutations may allow asexual planarians to limit Muller's ratchet and persist despite their obligate asexuality.

Materials and Methods

Collecting available RNAseq data

RNAseq data was collected from previous studies with a total of 169 libraries from 7 studies for the asexual strain and 46 libraries from 3 different sexual populations (SI1) (Currie and Pearson, 2013; Davies et al., 2017; Duncan et al., 2015; Khan and Newmark, 2022; Scimone et al., 2014; Tu et al., 2015; van Wolfswinkel et al., 2014; Vogg et al., 2014; Zhu et al., 2015). The Oxford sexual *S. mediterranea* population data was acquired by extracting RNA from whole, mature sexual worms using a standard TRIzol extraction (Liu and Rink, 2018). Extracted RNA was then sequenced with Ilumina sequencing of 150bp paired end reads.

Variant calling and filtration

RNAseq data was then trimmed and quality filtered with trimmomatic and FASTQC using default settings (Bolger et al., 2014). STAR was then used to map the RNAseq reads to the dd_Smes_g4 genome (Dobin et al., 2013; Grohme et al., 2018). Thus, all alternative variants will be called relative to the sexual genome that is represented by a single haplotype. For variant calling we used the GATK 4.2.2.0 software following best practice (Van der Auwera et al., 2013). We used the suggested parameters for variant filtration (SI2), providing us with a list of variants in each population (SI3).

Analysis of population structure

To ensure that our collected data really corresponds to distinct strains and that no one library is biasing our results, we analysed the population structure and mean library allele frequency (Fig. S1). Data from the sexual Davies et al population (Davies et al., 2017) was enriched for

rare alleles, so all relevant analyses were preformed both with and without these data with no difference being found (Fig. S2). Plink was used to run a Principal Component Analysis (PCA), which confirmed that the populations were distinct strains (Fig. S1) (Purcell et al., 2007).

Defining Orthologues

We ran Orthofinder using a proteome from the Neiro et al annotation to isolate the orthologues of each predicted gene (Emms and Kelly, 2019; Neiro et al., 2022). To reduce redundancy, we first ran CD-hits to collapse all peptides with 99% identity to a single transcript ID.

dN/dS analysis

We used the dNdScv R package for dN/dS analysis and annotation of mutation impact in coding regions (Martincorena et al., 2017). To get the list of genes for dN/dS 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 coding sequence (SI4-SI5). When comparing dN/dS 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 dN/dS values of BUSCO gene exons between the strains, we isolated single copy BUSCO genes in planarians using BUSCO (metazoan gene list) (Manni et al., 2021). We then isolated BUSCO exons that had both synonymous and nonsynonymous mutations within them (to allow for meaningful dN/dS assessment using the dNdScv local model) and compared their dNdS values provided by dNdScv (SI6). We also compared dN/dS ratios in exons that were under negative selection in

both strains (SI7). These were defined by having a qmis_loc (adjusted P value) < 0.05 and wmis_loc (dN/dS) < 1 in the dN/dS output. Finally, we used the global model to isolate genes under positive selection in asexual planarians (SI8-SI9), and top blast hits were taken from planmine (Rozanski et al., 2019).

Analysis of zygosity and allele frequency

To assess allele frequency and homozygosity levels accurately, we only used information about variants from samples that had a minimum depth of 10 at the site (SI10). Allele frequency in each sample was calculated using VCFtools (0.1.16) (Danecek et al., 2011). This resulted in each variant obtaining a 0, 0.5 or 1 value for each sample, with the numbers corresponding to homozygous reference, heterozygous, and homozygous alternative respectively. Note that since we are using gene expression data, there is ambiguity about whether unbalanced haplotypes would indicate intra-organismal diversity or biased expression patterns, which is why we do not look at continuous values of allele frequency within individuals (for intra-organismal diversity assessments see chapter 2). We could then use this information to calculate allele frequency in the population for each variant (total number of alternative alleles in the population / total number of alleles in the population). We also calculated the proportion of libraries that were homozygous for each variant (number of libraries homozygous for alternative variant / number of libraries containing the variant). Note that assuming a complete lack of recombination in asexuals, the measure of homozygosity from a group of individuals should be very similar to the measure of homozygosity within a single individual.

Expected values for mean proportion of homozygosity

To test whether there is excess homozygosity in the asexual strain, we compared the observed and expected proportions of homozygosity in the asexual population. When testing for excess homozygosity, or GC bias in homozygous variants, only synonymous mutations were used to limit the impact of selection. To calculate the expected proportion of homozygous ancestral variants (those shared between the sexual and asexual strains), we assumed that the alternative and reference alleles were equally common in the ancestral population that gave rise to the asexual strain and ignored the possibility of back mutations. Under these assumptions and in the absence of recombination, ~33% of the ancestral variants are expected to be homozygous in the asexual strain (Fig. S3). This is because if the alternative and reference allele were equally common in the population that gave rise to the asexual lineage, the probability of inheriting a ref/ref, ref/alt, alt/ref, or alt/alt genotype is equal (P = 0.25). However, since only genotypes with at least one alternative allele are included in our analysis, the possibility of a ref/ref genotype needs to be excluded. This means that the probability of observing an alt/alt genotype is 0.33 (P(alt/alt) / (P(alt/alt) + P(alt/ref) + P(ref/alt)) = 0.25 / 0.75 = 0.33). Ancestral variants were defined as being alternative variants relative to the sexual genome (represented by a single haplotype), that are found in both the sexual and asexual strains. Thus, homozygous alternative variants [1,1] and heterozygous alternative variants [1,0] found in the asexual strain would be considered ancestral if the same alternative variant was found in the sexual strain (irrespective of whether it was found as a homozygous or heterozygous in the sexual strain). However, since we are only assessing alternative variants relative to the sexual genome that is represented by a single haplotype, we are not assessing shared ancestral alleles that are homozygous reference in the asexual [0,0] and homozygous [1,1] or heterozygous [1,0] for the alternative allele in the sexual strain. Thus, our results related to the ancestral variants are based on a subset of ancestral variants that are arbitrarily defined by the allele used as the reference allele in the sexual genome. The

expected proportion of homozygous asexual specific variants cannot be precisely known because it depends on how many of them were formed by mutation after the sexual-asexual split, and how many of them were inherited from a sexual ancestor and subsequently lost from the sexual population. However, we can place lower and upper limits for the expected homozygosity. The lower limit corresponds to a scenario in which all asexual specific variants arose by mutation after the sexual-asexual split. In the absence of recombination, we would expect this to lead to no homozygous variants (since the de-novo mutations should be heterozygous). The upper limit corresponds to a scenario where all asexual specific variants were inherited from a sexual ancestor and so the proportion of homozygous variants depends on Mendelian principles similar to the expected homozygosity of ancestral variants. To reiterate, in this case there are 3 possible ways that the alternative alleles could have been inherited, alt/alt, alt/ref, and ref/alt, meaning that if all 3 options are equally likely (assuming that the alternative and reference alleles were equally common in the population), we would expect one third of all asexual specific variants to be homozygous (Fig. S3).

Assessing evidence for effective selection on gene conversion events

To assess whether there is effective selection to remove deleterious gene conversion events, we wanted to test whether more harmful mutations are less likely to be homozygous. Since we cannot know the precise impact of each mutation on fitness, we broadly ranked variants based on the amount of change the mutation is likely to have on the protein sequence. In our analysis we assumed that coding indels are more deleterious than coding SNPs, frameshifting indels are more deleterious than non-frameshifting indels, nonsynonymous mutations are more deleterious than synonymous mutations, and nonsense mutations are more deleterious than nonsynonymous mutations.

It would also be reasonable to assume that that noncoding mutations are less deleterious than coding mutations. However, noncoding regions are often rich in repeats which can complicate the assessment of homozygosity and allele frequency (Nishant et al., 2009; Patil and Vijay, 2021). To test if this was the case in our data, we ran RepeatMasker 4.1.2-p1 on the dd_Smes_g4 genome and found that noncoding regions were rich in repetitive sequences, making them inappropriate comparisons to variants in coding regions (Fig. S4).

Results

Limited impact of Muller's ratchet in asexual planarians

We began by looking for evidence of deleterious mutation accumulation in the obligate asexual strain by using dN/dS analyses. We found that across the entire set of protein coding genes (Neiro et al., 2022) (SI11), the median asexual dN/dS value is 3.6% higher than that found in the sexual strain (Wilcoxon rank sum; $N_{Asexual}$ = 16,445, N_{Sexual} = 10,700, W = 91,647,379, $P < 0.001$, Fig. 1A). The distribution observed with the majority of values clustered around discrete values is driven by discrete values of dS due to an abundance of exons with low numbers of synonymous mutations (Fig. S6). High dN/dS values can indicate both deleterious mutation accumulation and positive selection, thus we assessed dN/dS values in genes that are unlikely to be under positive selection. Specifically, we compared the dN/dS values of exons belonging to highly conserved single copy BUSCO genes (SI6) (Manni et al., 2021), and found no significant difference between the strains (Wilcoxon signed rank; N_{pairs} = 160, V = 7,069, P = 0.284, Fig. 1B). However, when we compared the dN/dS values of exons under significant negative selection in both strains (SI7), we found that the asexual dN/dS values were 27% higher than in the sexual strain (Wilcoxon signed rank; Npairs = 41, V = 665, p-value = 0.002, Fig. 1C). This is in contrast to some previous studies that found

higher impacts of Muller's ratchet in asexual lineages (dN/dS values ~200%-~3,000% higher than in their sexual counterparts) (Barraclough et al., 2007; Bast et al., 2018; Henry et al., 2012; Johnson and Howard, 2007; Neiman et al., 2010; Paland and Lynch, 2006). Additionally, we isolated genes that contain exons under positive selection in the asexual population (Fig. 1D, SI9). These included multiple genes related to GLIPR1-like genes (see methods), which may suggest that these genes evolved new functions in the asexual population. Taken together, while asexual planarians may be subject to Muller's ratchet, the accumulation of deleterious mutations is somehow limited relative to other asexual taxa, and is not present in highly conserved single copy genes. This requires a mechanistic explanation in the absence of sexual recombination.

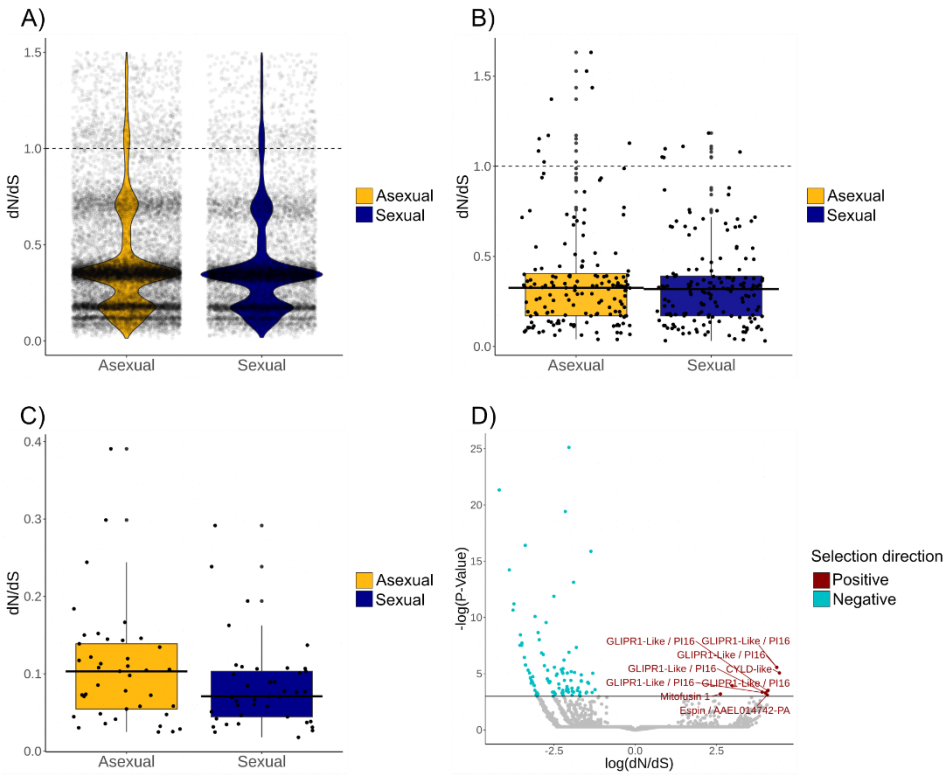


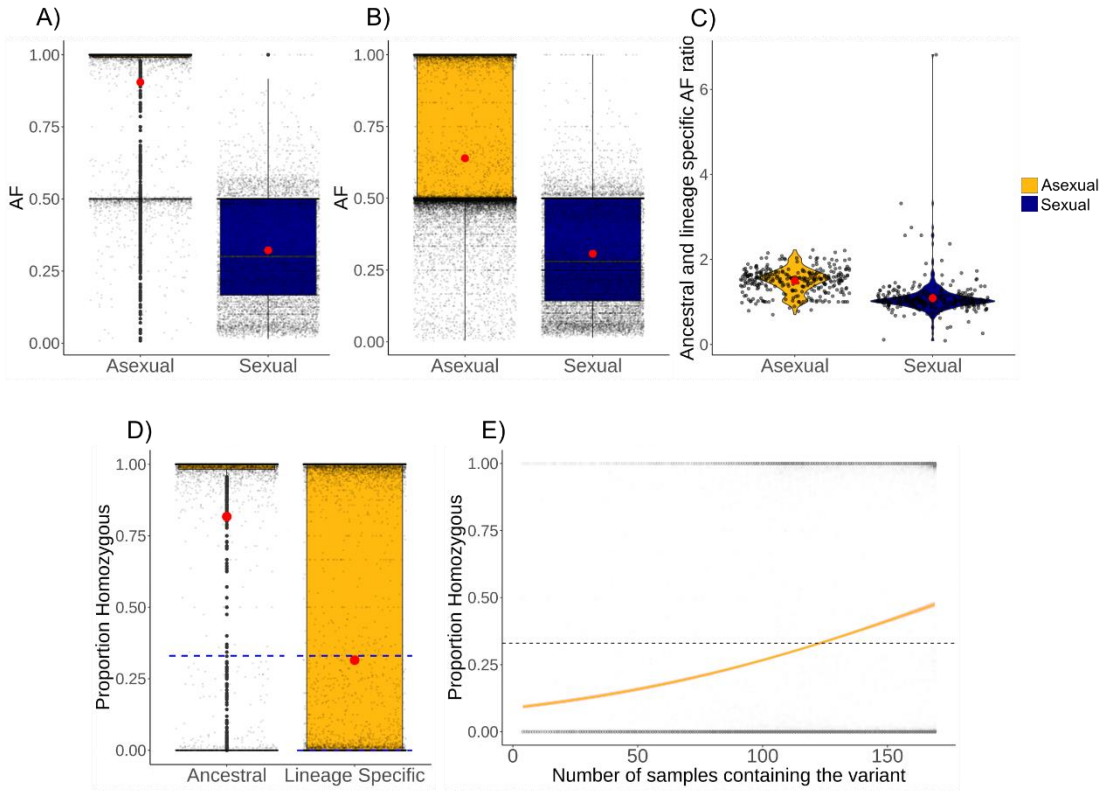
Figure 1 - dN/dS analysis in sexual and asexual populations. A) Violin plot showing the dN/dS values of all exons for both sexual and asexual populations. Sexual mean dN/dS = 0.444, asexual mean dN/dS = 0.46. ; B) dN/dS of all BUSCO exons that contained both

synonymous and nonsynonymous variants in both sexual and asexual population. Sexual mean dN/dS = 0.318, asexual mean dN/dS = 0.325; C) dN/dS of exons under negative selection in both sexual and asexual populations. Sexual mean dN/dS = 0.084, asexual mean dN/dS = 0.107; D) Volcano plot showing exons under positive selection in the asexual population.

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

Mechanisms limiting Muller's ratchet in asexual planarians may involve somatic recombination mechanisms, such as gene conversion, that in combination with selection can remove deleterious mutations (Box. 2). A signature of such a process would be homogenisation of the genome, leading to low polymorphism and high homozygosity of alternative alleles. We compared allele frequencies of synonymous alternative variants in asexual and sexual populations and found that asexual alternative allele frequency is significantly higher than in the sexual population for both shared ancestral variants and lineage specific variants suggesting that there is lower polymorphism in the asexual population (Wilcoxon signed $rank_{ancestral}$; N_{pairs} = 9,206, V = 40,369,398, $P_{ancestral} < 0.001$, Fig. 2A; Wilcoxon rank $sum_{lineage\text{-}specific}$, $N_{asexual}$ = 30,463, N_{sexual} = 13,428, W = 338,332,851, $P_{lineage\text{-}specific} < 0.001$, Fig. 2B). Additionally, in asexuals, the alternative allele frequencies of ancestral variants were 45% higher than lineage specific variants, while in the sexual population the alternative allele frequencies of ancestral variants were only 4% higher than lineage specific variants (Wilcoxon rank sum; $N_{asexual}$ = 309, N_{sexual} = 305, W = 82,102, P < 0.001, Fig. 2C). The increased allele frequency in ancestral variants suggests a process of homogenisation over time, which is more effective in the asexual genome.

High alternative allele frequencies in asexuals may be explained by the clonal nature of asexual planarians, so we tested whether they also show excess homozygosity which would be indicative of a homogenising process that cannot be accounted for by clonality. We find that the oldest variants in our dataset, those shared by both sexual and asexual populations, showed significant excess homozygosity in the asexual population, with a mean homozygosity proportion of 0.82 (Wilcoxon signed rank; $N_{asexual\text{-}ancestral} = 9{,}206$, $V = 29{,}633{,}451$, $P < 0.001$, Fig. 2D). Additionally, we found that in asexual specific variants, the proportion of homozygosity correlated with the number of libraries supporting the variant, suggesting that older variants are more homozygous (Binomial GLM, $N = 30{,}463$, $P < 0.001$, Fig. 2E). The oldest asexual specific variants, those shared by all asexual populations, had a mean homozygosity of proportion of 0.58. This suggests that there is excess homozygosity even under the extreme assumption that all asexual specific variants were originally inherited from a sexual ancestor (Wilcoxon signed rank; $N_{ancestral\ asexual\ specific} = 2{,}275$, $V = 2{,}121{,}916$, $P < 0.001$, Fig. 2E). Taken together, this suggests that some homogenising process such as gene conversion may be impacting the asexual genome over time.



.

Figure 2 - Excess homozygosity of alternative alleles in the asexual population. Red dots indicate means. A) Alternative allele frequencies of ancestral variants shared by sexual and asexual populations. Mean sexual allele frequency = 0.32. Mean asexual allele frequency = 0.9; B) Alternative allele frequencies of lineage specific. Mean sexual allele frequency = 0.31. Mean asexual allele frequency = 0.64; C) Violin plot showing the ratio of average alternative allele frequencies per scaffold of ancestral vs lineage specific variants. Mean sexual ratio = 1.038. Mean asexual ratio= 1.446; D) Boxplot showing the proportion of homozygous libraries per alternative variant in ancestral (shared between the sexual and asexual populations) and lineage specific alternative alleles in the asexual population. Dashed blue lines indicate the expected proportion of homozygosity in the absence of recombination; E) Binomial GLM showing that the proportion of homozygosity in asexual lineage specific

variants increases with the number of libraries containing the variant. When all libraries contain the variant the mean proportion of homozygosity is 0.58. The dashed line indicates the expected homozygosity assuming that all asexual specific variants were inherited from sexual ancestors.

GC-biased gene conversion in asexual planarians

We hypothesised that since gene conversion has a GC bias (Chen et al., 2007), 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, with 53% of homozygous variants being composed of GC nucleotides compared to 35% in heterozygous variants (Chi-sqr = 1,328.6, df = 1, N_{HOM} = 17,487, N_{HET} = 22,517, $P < 0.001$, Fig. 3A), 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 (Binomial GLM, N = 44,669, $P < 0.001$, Fig. 3B). Furthermore, we show that AT to GC mutations are 1.5 times more homozygous than GC to AT mutations, with mean homozygosity proportions of 0.57 and 0.38 respectively (Wilcoxon rank sum, $N_{AT->GC}$ = 17,984, $N_{GC->AT}$ = 21,483 W = 232,985,919, $P < 0.001$, Fig. 3C). This GC bias is not driven by AT to GC mutations across the transcriptome generally (Fig. S5). 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. 3D). Taken together, this

suggests that GC biased gene conversion takes place in the asexual planarian, leading to a loss of heterozygosity which could limit the impact of Muller's ratchet (Box. 2).

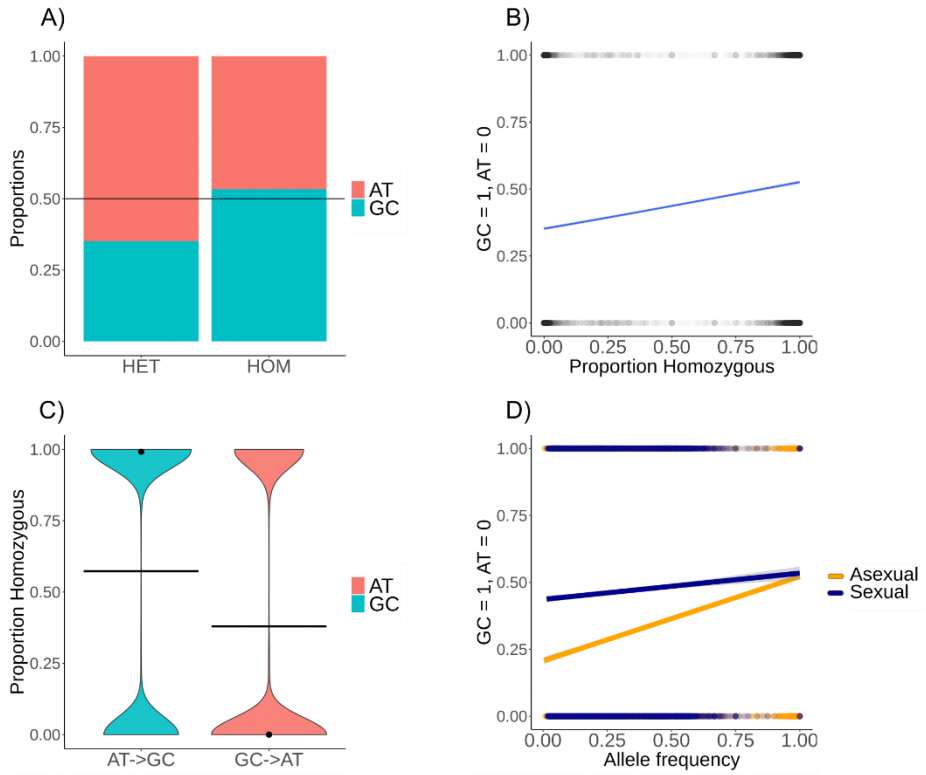


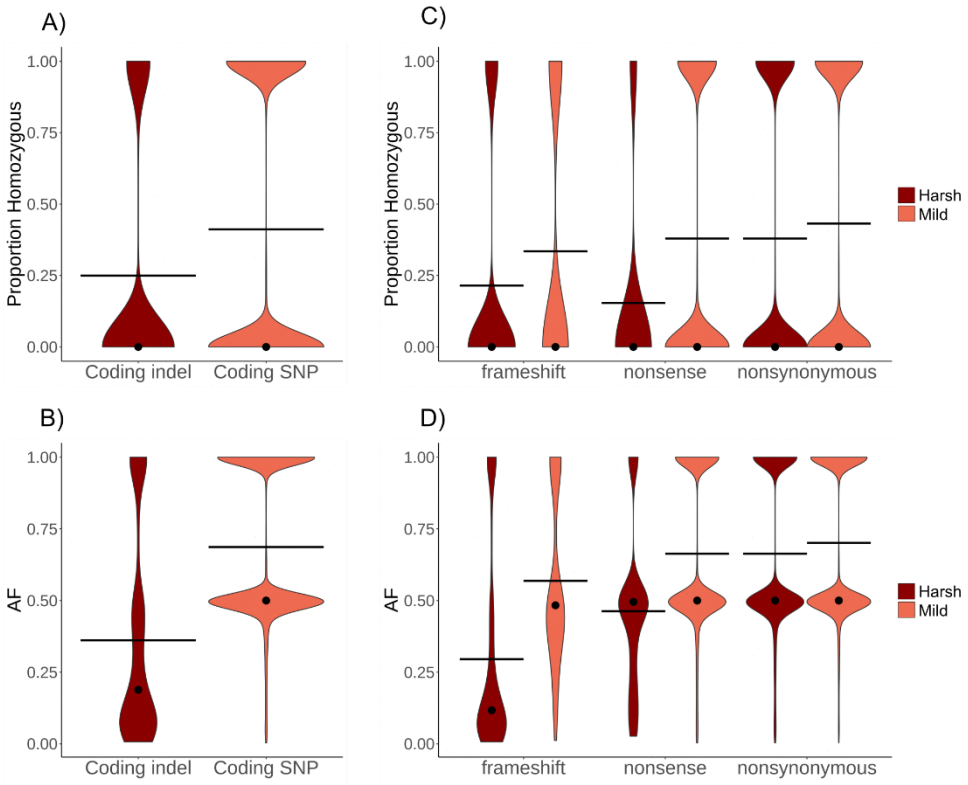
Figure 3 - GC bias in homozygous variants. A) Bar plot showing that the proportion of GC is higher in homozygous variants compared to heterozygous variants. Black line indicates the 50% line; B) Binomial GLM showing that variants with a higher proportion of homozygosity are more likely to be of GC identity in asexual planarians; C) Violin plot showing that AT to GC mutations are more homozygous than GC to AT mutations. Black dots indicate medians and black cross bars indicate means; D) Binomial GLM showing that variants with higher allele frequencies are more likely to be of GC identity in both sexuals and asexual, but in asexuals the effect is larger.

Selection on gene conversion events removes deleterious mutations

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. We find that coding SNPs have a mean homozygosity proportion of 0.41 and are more homozygous than coding indels which have a mean homozygosity proportion of 0.25 (Wilcoxon rank sum; $N_{indels} = 883$, $N_{SNPs} = 64{,}220$, $W = 23{,}118{,}081$, $P < 0.001$, Fig. 4A). Additionally, coding SNPs have higher allele frequencies than coding indels, with mean allele frequencies of 0.69 and 0.36 respectively (Wilcoxon rank sum; $N_{indels} = 883$, $N_{SNPs} = 64{,}220$, $W = 11{,}652{,}733$, $P < 0.001$, Fig. 4B). These patterns could be driven by other factors such as error rates in each category of mutation, so we also compared the homozygosity and allele frequencies of mutations within each category. We found that the homozygosity and allele frequencies were higher in synonymous vs nonsynonymous SNPs (Wilcoxon rank sum; $N_{Synonymous} = 39{,}669$, $N_{Nonsynonymous}$ 24,471=, $W_{Hom} = 462{,}855{,}666$, $P_{Hom} < 0.001$, $W_{AF} = 452{,}432{,}400$, $P_{AF} < 0.001$, Fig. 4C-D), non-framshifting indels vs framshiting indels (Wilcoxon rank sum; $N_{Non\text{-}frameshifting} = 154$, $N_{Frameshifting} = 664$, $W_{Hom} = 44{,}439$, $P_{Hom} < 0.001$, $W_{AF} = 23{,}168$, $P_{AF} < 0.001$, Fig. 4C-D), and nonsynonymous SNPs vs nonsense SNPs (Wilcoxon rank sum; $N_{Nonsynonymous} = 24{,}471$, $N_{Nonsense} = 65$, $W_{Hom} = 613{,}538$, $P_{Hom} < 0.001$, $W_{AF} = 494097$, $P_{AF} < 0.001$ Fig. 4C-D). This suggests that mutations that are more likely to be deleterious are less likely to be homozygous and are rarer in the population. 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.

Figure 4 - Predicted deleterious mutations are less likely to be homozygous. Dots indicate medians, crossbars indicate means. A) Mean proportions of homozygosity are lower in coding indels than coding SNPs; B) Mean allele frequencies are lower in coding indels than coding SNPs; C) X-axis text refers to the more deleterious mutation in each category. Frameshifting indels have lower homozygosity than non-frameshifting indels (mean proportion homozygous of 0.215 and 0.335 respectively). Nonsense variants have lower



homozygosity than nonsynonymous variants (mean proportion homozygous of 0.15 and 0.38 respectively). Nonsynonymous variants have lower homozygosity than synonymous variants (mean proportion homozygous of 0.38 and 0.43 respectively); D) X-axis text refers to the more deleterious mutation in each category. Frameshifting indels have lower allele

frequencies than non-frameshifting indels (mean allele frequencies of 0.3 and 0.57 respectively). Nonsense variants have lower allele frequencies than nonsynonymous variants (mean allele frequencies of 0.46 and 0.66 respectively). Nonsynonymous variants have lower allele frequencies than synonymous variants (mean allele frequencies of 0.66 and 0.7 respectively).

Discussion

Mechanisms explaining how some asexual species persist despite the potential costs 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 Hill-Robertson effects, potentially explaining how this obligate asexual can persist over evolutionary time.

Our results suggest that the impact of Muller's ratchet is relatively limited in asexual planarians. This is surprising since it has been shown both empirically and by simulation that Muller's ratchet should drive the rapid accumulation of deleterious mutations even in recent asexual lineages (Loewe and Lamatsch, 2008; Lynch and Gabriel, 1990; Neiman et al., 2010; Tucker et al., 2013). For example, *Daphnia pulex*, an asexual lineage that is estimated to be only around 1,000 years old, is already showing signs of deleterious mutation accumulation (Tucker et al., 2013). Additionally, *Potamopyrgus antipodarum*, a freshwater snail in which the sexual form coexists with a recent obligate asexual form, shows higher rates of deleterious mutations in the asexual strain (Johnson and Howard, 2007; Neiman et al., 2010). This suggests that Muller's ratchet can act rapidly to drive mutation accumulation in asexuals. Thus, one might expect asexual planarians to show clear signs of Muller's ratchet,

given that the asexual stem cells that make up their soma have been asexual for at least tens of millions of generations. However, they seem to only show an 27% increase of dN/dS in exons under negative selection, and this may be driven by the higher number of variants in the asexual dataset (Fig. S1), increasing the statistical power for identifying exons under negative selection, including those with higher dN/dS ratios. Furthermore, there is no evidence for elevated dN/dS ratios in highly conserved BUSCO genes, where nonsynonymous mutations are likely to be deleterious.

This lack of high mutation accumulation in asexual planarians could explain how they have persisted as asexuals and is similar to the patterns seen in some other long term asexual lineages (Brandt et al., 2017; Swanstrom et al., 2011). For example, asexual oribatid mites have been shown to have lower dN/dS ratios compared to their sexual counterparts, suggesting that, in this case, purifying selection might be stronger in the asexual species (Brandt et al., 2017). In bdelloid rotifers nuclear genes do not show signs of Mullers ratchet (Birky et al., 2005; Welch and Meselson, 2001). Others have shown an accumulation of deleterious mutations in bdelloid mitochondrial genes (Barraclough et al., 2007), but have later suggested that this may be due to other factors which when taken into account suggest no signs of Muller's ratchet in the mitochondrial genes of bdelloid rotifers (Swanstrom et al., 2011). The lack of strong impacts of Muller's ratchet on some long term asexuals suggests that they may possess mechanisms to limit the accumulation of deleterious mutations.

Some authors have previously suggested gene conversion as a mechanism to limit Muller's ratchet in asexual lineages (Flot et al., 2013; Loewe and Lamatsch, 2008; Schön et al., 2009). The idea is that deleterious alleles can be overwritten by beneficial alleles, and that selection will favour these events, ultimately removing deleterious mutations from the population (Box. 2). However, beneficial alleles can equally be overwritten by deleterious alleles in the gene conversion process, and so selection must be able to act against these events to

effectively counteract Muller's ratchet. This might explain why some asexuals show high rates of gene conversion without limiting Mullers ratchet (Omilian et al., 2006; Tucker et al., 2013). Furthermore, for gene conversion to stop the progression of Mullers ratchet, it may need to occur at a rate higher than the rate of point mutations. Otherwise, mutations may accumulate at a rate faster than gene conversion can remove them. This might be the case in some asexuals such as *Timema* stick insects, which show lower rates of gene conversion compared to their sexual relatives, and dN/dS values 3.6-13.4 times higher compared to their sexual counterparts (Bast et al., 2018; Henry et al., 2012). Despite these caveats, gene conversion remains a potential mechanism for limiting the cost of asexuality. This is in part because it is a byproduct of DNA damage repair mechanisms that are available to diploid (or higher ploidy) organisms (Chen et al., 2007) and so, it does not need to evolve *de novo* as a solution to asexuality. It is worth pointing out that key components of the machinery that could be used to drive gene conversion through the DNA mismatch repair pathway have been shown to be enriched in neoblasts, and indeed are likely to be expressed in most dividing cells that are subject to cell division related double strand breaks (Birdsell, 2002; Solana et al., 2012 (additional file 13)). However, further research would be useful to understand whether all the machinery necessary for gene conversion is present in neoblasts, and whether their knockdown stops gene conversion in response to DNA damage. Taken together these shared signatures of gene conversion between bdelloid rotifers (Flot et al., 2013), darwinulid ostracods (Butlin et al., 1998), and now *S.mediterranea*.

While gene conversion may be important in both fissiparous and parthenogenic asexuality, we hypothesise that it might be significantly more effective in fissiparous asexuality. This is due to the variation within the soma being heritable in fissiparous organisms, increasing the amount of variation subject to selection (Howe et al., 2022; Leria et al., 2019). This is particularly relevant in *S.mediterranea* whose immortal stem cells are extremely abundant

and represent a considerable proportion of the soma, and previous work in other planarian species has shown that a fissiparous lifestyle increases somatic diversity (Leria et al., 2019; van Wolfswinkel et al., 2014). These high numbers of stem cells are important since each stem cell will experience unique gene conversion events, and with a large enough population of stem cells, it is likely that some of them will overwrite deleterious alleles with their beneficial counterparts. Selection may then act both within individuals (cell-cell competition) and between individual worms, selecting for those that have purged deleterious mutations, ultimately removing deleterious mutations from the population. Thus, all else being equal, fissiparous organisms may be more efficient than parthenogenetic asexuals at removing deleterious mutations from the population, since gene conversion within their soma can also be selected upon to remove deleterious mutations.

However, it is important to point out that there is always the possibility that asexual planarians have rare cryptic sex that we are simply not aware of. For example, the brine shrimp, *Artemia parthenogenetica*, was once thought to be asexual but was later shown to undergo rare sex in a mass crossing experiment (Boyer et al., 2021). Additionally, even the presumed ancient asexual bdelloid rotifer *Macrotrachela quadricornifera* has had its asexuality contested with evidence supporting allele sharing between different isolates (Laine et al., 2022). However, similar patterns can potentially be explained without sex by ancestral duplications followed by independent allele losses (Wilson et al., 2024). Either way, it is entirely possible that in the case of *S. mediterrannea* rare cryptic sex might occur and might explain how they avoid the accumulation of deleterious mutations. However, it would not explain the strong signatures of gene conversion, which is not as clear in the genome of sexual planarians (Fig. 2C, Fig. 3D). Finally, even if rare sex occurs, high levels of gene conversion may still dramatically impact the genome (Hartfield et al., 2016), and potentially be a useful mechanism to deal with the cost of highly reduced levels of sexual recombination.

In summary, 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.

Supplementary Data:

SI1 – List of all libraries used in study (Sexual RNAseq from Oxford is unpublished).

SI2 – variant filtration criteria.

S3 – List of variants

SI4 - Asexual dN/dS local model.

SI5 – Sexual dN/dS local model.

SI6 – List of all BUSCO exons with synonymous and nonsynonymous mutations in both sexual and asexual strains.

SI7 – List of all exons under negative selection in both sexual and asexual strains

SI8 – Asexual dN/dS global model.

SI9 - List of exons under positive selection in asexual strain using the global dNdScv model.

S10 – List of variants where homozygosity and allele frequency information only come from libraries with at least a depth of 10 at the variant site.

SI11 – List of all exons used in study.

Author contribution

Simon Kershenbaum (Conceptualisation, data analysis, writing); Manuel Jara-Espejo (Data analysis); Ashleigh S. Griffin (Conceptualisation, writing); Aziz Aboobaker (Conceptualisation, writing)

Acknowledgements

We would like to thank Tim Barraclough, Chris Wilson, and Tymoteusz Pieszko for their helpful comments on the manuscript.

Conflict of interest statement

The authors declare no conflicts of interest.

Funding

The project was funded by the BBSRC (BB/T008784/1) and MRC (MR/T028165/1)

Data availability

Supplementary information and data can be found at:

https://github.com/SimonOX-Gene/Smed_GeneConversion

Additional information is available upon request.

Bibliography

Barraclough, T. G., Fontaneto, D., Ricci, C. and Herniou, E. A. (2007). Evidence for inefficient selection against deleterious mutations in cytochrome oxidase I of asexual bdelloid rotifers. *Molecular Biology and Evolution* **24**, 1952–1962.

Bast, J., Parker, D. J., Dumas, Z., Jalvingh, K. M., Tran Van, P., Jaron, K. S., Figuet, E., Brandt, A., Galtier, N. and Schwander, T. (2018). Consequences of asexuality in natural populations: Insights from stick insects. *Molecular Biology and Evolution* **35**, 1668–1677.

Birdsell, J. A. (2002). Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution* **19**, 1181–1197.

Birky, C. W., Wolf, C., Maughan, H., Herbertson, L. and Henry, E. (2005). Speciation and selection without sex. *Hydrobiologia* **546**, 29–45.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Boyer, L., Jabbour-Zahab, R., Mosna, M., Haag, C. R. and Lenormand, T. (2021). Not so clonal asexuals: Unraveling the secret sex life of *Artemia parthenogenetica*. *Evol Lett* **5**, 164–174.

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. and Bast, J. (2017). Effective purifying selection in ancient asexual oribatid mites. *Nat Commun* **8**, 873.

Brandt, A., Tran Van, P., Bluhm, C., Anselmetti, Y., Dumas, Z., Figuet, E., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., et al. (2021). Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proceedings of the National Academy of Sciences* **118**, e2101485118.

Butlin, R., Schön, I. and Martens, K. (1998). Asexual reproduction in nonmarine ostracods. *Heredity* **81**, 473–480.

Charlesworth, B. and Barton, N. H. (1996). Recombination load associated with selection for increased recombination. *Genetics Research* **67**, 27–41.

Chen, J.-M., Cooper, D. N., Chuzhanova, N., Férec, C. and Patrinos, G. P. (2007). Gene conversion: mechanisms, evolution and human disease. *g e n e ti c s* 14.

Comeron, J. M., Williford, A. and Kliman, R. M. (2008). The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19–31.

Currie, K. W. and Pearson, B. J. (2013). Transcription factors lhx1/5-1 and pitx are required for the maintenance and regeneration of serotonergic neurons in planarians. *Development* **140**, 3577–3588.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., et al. (2011). The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158.

Davies, E. L., Lei, K., Seidel, C. W., Kroesen, A. E., McKinney, S. A., Guo, L., Robb, S. M., Ross, E. J., Gotting, K. and Alvarado, A. S. (2017). Embryonic origin of adult stem cells required for tissue homeostasis and regeneration. *eLife* **6**, e21052.

de Visser, J. A. G. M. and Elena, S. F. (2007). The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* **8**, 139–149.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Duncan, E. M., Chitsazan, A. D., Seidel, C. W. and Alvarado, A. S. (2015). Set1 and MLL1/2 target distinct sets of functionally different genomic loci in vivo. *Cell Rep* **13**, 2741–2755.

Emms, D. M. and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 238.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* **78**, 737–756.

Fields, C. and Levin, M. (2018). Are planaria individuals? What regenerative biology is telling us about the nature of multicellularity. *Evol Biol* **45**, 237–247.

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G. J., Hejnol, A., Henrissat, B., Koszul, R., Aury, J.-M., et al. (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* **500**, 453–457.

Gibson, A. K., Delph, L. F. and Lively, C. M. (2017). The two-fold cost of sex: Experimental evidence from a natural system. *Evol Lett* **1**, 6–15.

Grohme, M. A., Schloissnig, S., Rozanski, A., Pippel, M., Young, G. R., Winkler, S., Brandl, H., Henry, I., Dahl, A., Powell, S., et al. (2018). The genome of Schmidtea mediterranea and the evolution of core cellular mechanisms. *Nature* **554**, 56–61.

Hamilton, W. D., Axelrod, R. and Tanese, R. (1990). Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences* **87**, 3566–3573.

Hartfield, M., Wright, S. I. and Agrawal, A. F. (2016). Coalescent times and patterns of genetic diversity in species with facultative sex: Effects of gene conversion, population structure, and heterogeneity. *Genetics* **202**, 297–312.

Henry, L., Schwander, T. and Crespi, B. J. (2012). Deleterious mutation accumulation in asexual timema stick Insects. *Molecular Biology and Evolution* **29**, 401–408.

Hill, W. G. and Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research* **8**, 269–294.

Howe, J., Rink, J. C., Wang, B. and Griffin, A. S. (2022). Multicellularity in animals: The potential for within-organism conflict. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120457119.

Johnson, S. G. and Howard, R. S. (2007). Contrasting patterns of synonymous and nonsynonymous sequence evolution in ssexual and sexual freshwater snail lineages. *Evolution* **61**, 2728–2735.

Judson, O. P. and Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution* **11**, 41–46.

Khan, U. W. and Newmark, P. A. (2022). Somatic regulation of female germ cell regeneration and development in planarians. *Cell Reports* **38**, 110525.

Laine, V. N., Sackton, T. B. and Meselson, M. (2022). Genomic signature of sexual reproduction in the bdelloid rotifer *Macrotrachella quadricornifera*. *Genetics* **220**, iyab221.

Lázaro, E. M., Harrath, A. H., Stocchino, G. A., Pala, M., Baguñà, J. and Riutort, M. (2011). *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evolutionary Biology* **11**, 274.

Lehtonen, J., Jennions, M. D. and Kokko, H. (2012). The many costs of sex. *Trends in Ecology & Evolution* **27**, 172–178.

Leria, L., Vila-Farré, M., Solà, E. and Riutort, M. (2019). Outstanding intraindividual genetic diversity in fissiparous planarians (Dugesia, Platyhelminthes) with facultative sex. *BMC Evolutionary Biology* **19**, 130.

Liu, S.-Y. and Rink, J. C. (2018). Total RNA isolation from planarian tissues. In *planarian regeneration: Methods and protocols* (ed. Rink, J. C.), pp. 259–265. New York, NY: Springer.

Lively, C. M. and Morran, L. T. (2014). The ecology of sexual reproduction. *Journal of Evolutionary Biology* **27**, 1292–1303.

Loewe, L. and Lamatsch, D. K. (2008). Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (*Poecilia formosa*). *BMC Evolutionary Biology* **8**, 88.

Lynch, M. and Gabriel, W. (1990). Mutation load and the survival of small populations. *Evolution* **44**, 1725–1737.

Lynch, M., Bürger, R., Butcher, D. and Gabriel, W. (1993). The mutational meltdown in asexual populations. *Journal of Heredity* **84**, 339–344.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. and Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Molecular Biology and Evolution* **38**, 4647–4654.

Mark Welch, D. B. and Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers Without sexual reproduction or genetic exchange. *Science* **288**, 1211–1215.

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Loo, P. V., Davies, H., Stratton, M. R. and Campbell, P. J. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029-1041.e21.

Muller, H. J. (1964). The relation of recombination to mutational advance. 8.

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. and Taylor, D. R. (2010). Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol Biol Evol* **27**, 954–963.

Neiro, J., Sridhar, D., Dattani, A. and Aboobaker, A. (2022). Identification of putative enhancer-like elements predicts regulatory networks active in planarian adult stem cells. *eLife* **11**, e79675.

Newmark, P. A. and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Developmental Biology* **220**, 142–153.

Newmark, P. A., Wang, Y. and Chong, T. (2008). Germ cell specification and regeneration in planarians. *Cold Spring Harb Symp Quant Biol* **73**, 573–581.

Nishant, K. T., Singh, N. D. and Alani, E. (2009). Genomic mutation rates: What high-throughput methods can tell us. *Bioessays* **31**, 912–920.

Omilian, A. R., Cristescu, M. E. A., Dudycha, J. L. and Lynch, M. (2006). Ameiotic recombination in asexual lineages of *Daphnia*. *Proceedings of the National Academy of Sciences* **103**, 18638–18643.

Paland, S. and Lynch, M. (2006). Transitions to asexuality result in excess amino acid substitutions. *Science* **311**, 990–992.

Patil, A. B. and Vijay, N. (2021). Repetitive genomic regions and the inference of demographic history. *Heredity (Edinb)* **127**, 151–166.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.

Rozanski, A., Moon, H., Brandl, H., Martín-Durán, J. M., Grohme, M. A., Hüttner, K., Bartscherer, K., Henry, I. and Rink, J. C. (2019). PlanMine 3.0—improvements to a mineable resource of flatworm biology and biodiversity. *Nucleic Acids Research* **47**, D812–D820.

Saberi, A., Jamal, A., Beets, I., Schoofs, L. and Newmark, P. A. (2016). GPCRs Direct Germline Development and Somatic Gonad Function in Planarians. *PLoS Biol* **14**, e1002457.

Sahu, S., Dattani, A. and Aboobaker, A. A. (2017). Secrets from immortal worms: What can we learn about biological ageing from the planarian model system? *Seminars in Cell & Developmental Biology* **70**, 108–121.

Schön, I., Martens, K., Van Doninck, K. and Butlin, R. K. (2003). Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: *Ostracoda*): Molecular evolution in sexual and asexual ostracods. *Biological Journal of the Linnean Society* **79**, 93–100.

Schön, I., Rossetti, G. and Martens, K. (2009). Darwinulid ostracods: Ancient asexual scandals or scandalous gossip? In *lost sex: The evolutionary biology of parthenogenesis* (ed. Schön, I.), Martens, K.), and Dijk, P.), pp. 217–240. Dordrecht: Springer Netherlands.

Scimone, M. L., Kravarik, K. M., Lapan, S. W. and Reddien, P. W. (2014). Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. *Stem Cell Reports* **3**, 339–352.

Solana, J., Kao, D., Mihaylova, Y., Jaber-Hijazi, F., Malla, S., Wilson, R. and Aboobaker, A. (2012). Defining the molecular profile of planarian pluripotent stem cells using a

combinatorial RNA-seq, RNA interference and irradiation approach. *Genome Biology* **13**, R19.

Swanstrom, J., Chen, K., Castillo, K., Barraclough, T. G. and Fontaneto, D. (2011). Testing for evidence of inefficient selection in bdelloid rotifers: do sample size and habitat differences matter? *Hydrobiologia* **662**, 19–25.

Tan, T. C. J., Rahman, R., Jaber-Hijazi, F., Felix, D. A., Chen, C., Louis, E. J. and Aboobaker, A. (2012). Telomere maintenance and telomerase activity are differentially regulated in asexual and sexual worms. *Proceedings of the National Academy of Sciences* **109**, 4209–4214.

Tu, K. C., Cheng, L.-C., TK Vu, H., Lange, J. J., McKinney, S. A., Seidel, C. W. and Sánchez Alvarado, A. (2015). Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *eLife* **4**, e10501.

Tucker, A. E., Ackerman, M. S., Eads, B. D., Xu, S. and Lynch, M. (2013). Population-genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. *Proc Natl Acad Sci U S A* **110**, 15740–15745.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high-confidence variant calls: The genome analysis toolkit best practices pipeline. *Current Protocols in Bioinformatics* **43**, 11.10.1-11.10.33.

van Wolfswinkel, J. C., Wagner, D. E. and Reddien, P. W. (2014). Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **15**, 326–339.

Vergara, D., Jokela, J., Lively, C. M. and Bronstein, E. J. L. (2014). Infection dynamics in coexisting sexual and asexual host populations: Support for the red queen hypothesis. *The American Naturalist* **184**, S22–S30.

Vogg, M. C., Owlarn, S., Pérez Rico, Y. A., Xie, J., Suzuki, Y., Gentile, L., Wu, W. and Bartscherer, K. (2014). Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. *Developmental Biology* **390**, 136–148.

Wagner, D. E., Wang, I. E. and Reddien, P. W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811–816.

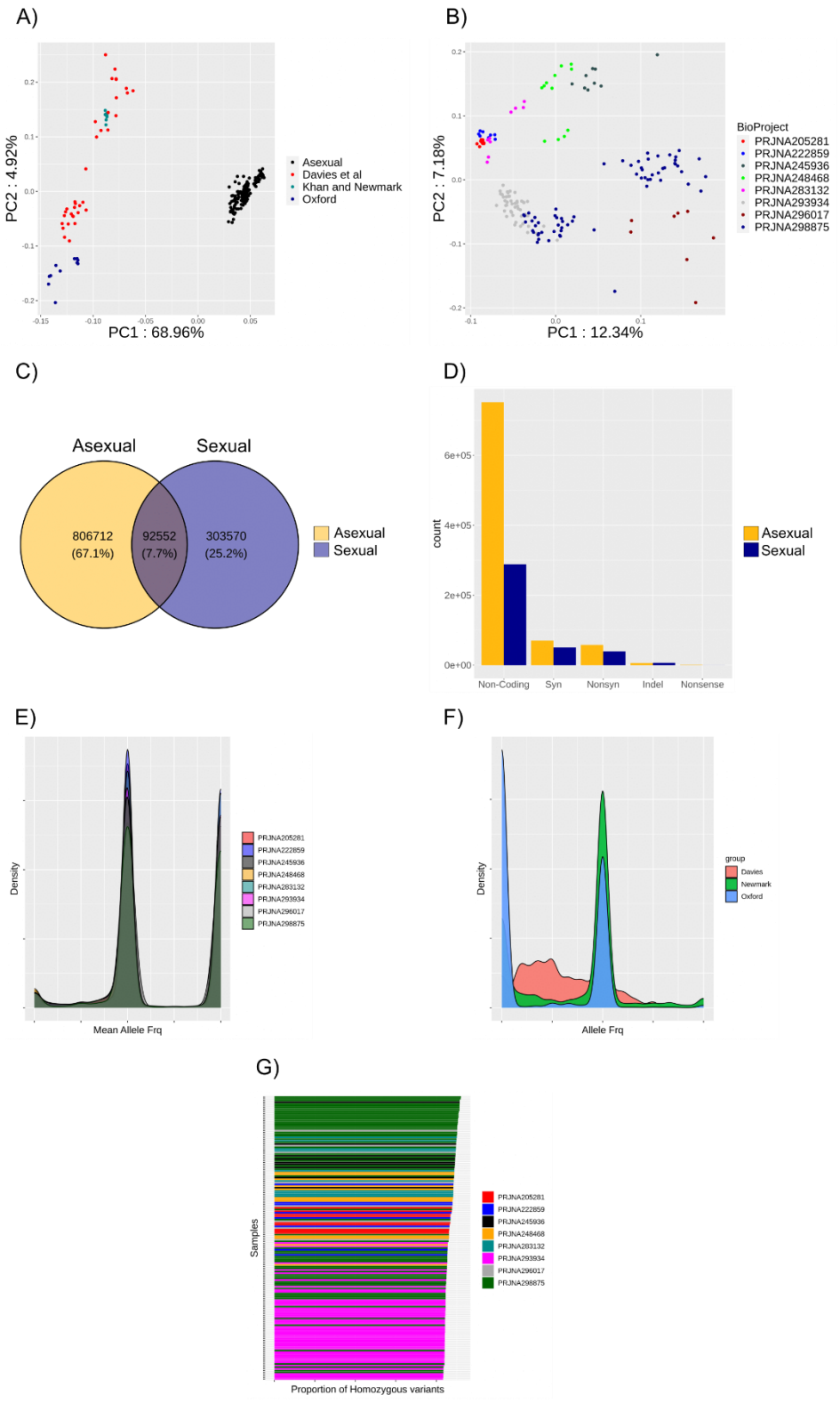
Welch, D. B. M. and Meselson, M. S. (2001). Rates of nucleotide substitution in sexual and anciently asexual rotifers. *Proceedings of the National Academy of Sciences* **98**, 6720–6724.

Wilson, C. G., Nowell, R. W. and Barraclough, T. G. (2018). Cross-contamination explains “Inter and intraspecific horizontal genetic transfers” between asexual bdelloid rotifers. *Current Biology* **28**, 2436-2444.e14.

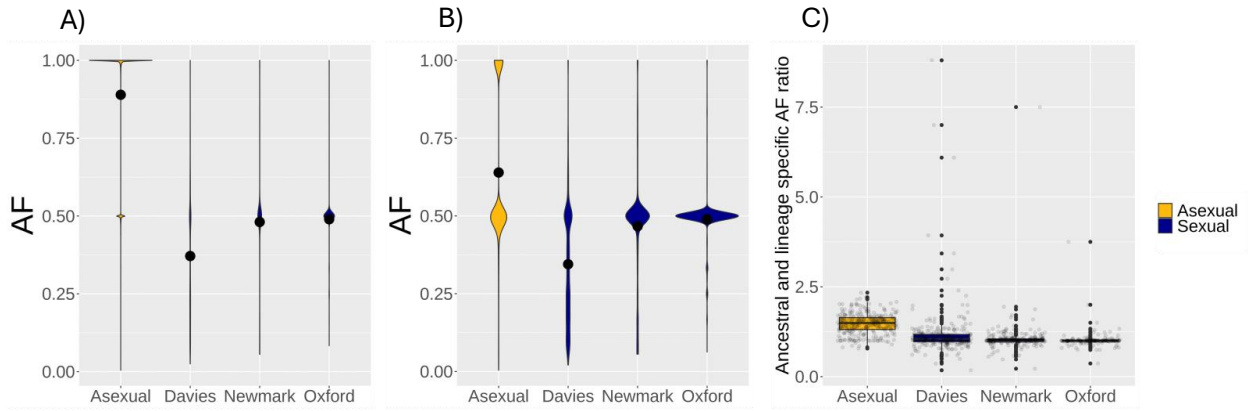
Wilson, C. G., Pieszko, T., Nowell, R. W. and Barraclough, T. G. (2024). Recombination in bdelloid rotifer genomes: asexuality, transfer and stress. *Trends in Genetics* **40**, 422–436.

Zhu, S. J., Hallows, S. E., Currie, K. W., Xu, C. and Pearson, B. J. (2015). A mex3 homolog is required for differentiation during planarian stem cell lineage development. *eLife* **4**, e07025.

Supplementary figures

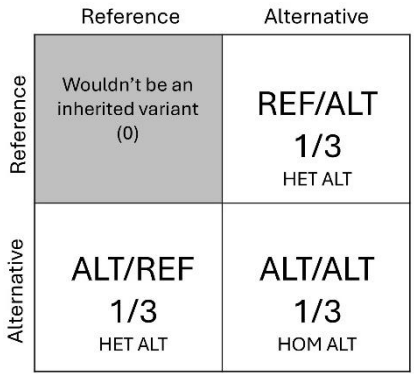

A)
PC2 : 4.92%
PC1 : 68.96%
Asexual
Davies et al
Khan and Newmark
Oxford
B)
PC2 : 7.18%
PC1 : 12.34%
BioProject
PRJNA205281
PRJNA222859
PRJNA245936
PRJNA248468
PRJNA283132
PRJNA293934
PRJNA296017
PRJNA298875
C)
Asexual
Sexual
806712
(67.1%)
92552
(7.7%)
303570
(25.2%)
Asexual
Sexual
D)
count
Non-Coding
Syn
Nonsyn
Indel
Nonsense
Asexual
Sexual
E)
Density
Mean Allele Frq
F)
Density
Allele Frq
group
Davies
Newmark
Oxford
G)
Samples
Proportion of Homozygous variants

Supplementary figure 1 – Description of population structure and library features. A) PCA of all libraries in study; B) PCA of asexual libraries in study coloured by BioProject ID; C) Venn Diagram showing the percentage of shared and unique variants between sexual and asexual populations. D) Barplot showing the numbers of different types of variants in sexual and asexual populations. E) Distribution of allele frequencies in asexual BioProjects seem broadly similar. F) Distribution of allele frequencies in sexual populations. The sexual population from the Davies et al 2017 seems to have an abundance of rare alleles. G) Proportion of homozygous variants in asexual libraries by BioProject.

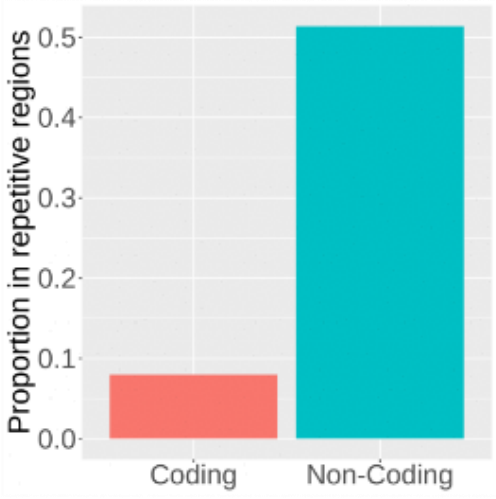


Supplementary figure 2 – Allele frequency comparison with the sexual populations separated. A) Ancestral variants have higher allele frequencies in the asexual compared to each of the sexual populations populations (Wilcoxon rank sum; $N_{Asexual}$ = 14,206 , $N_{Daviesl}$ = 9,625, $N_{Newmark}$ = 7,898, N_{Oxford} = 4,368, $W_{Asexua\text{-}Daviesl}$ = 126,240,405, $W_{Asexual\text{-}Newmark}$ = 97,403,170, $W_{Asexual\text{-}Oxford}$ = 53,920,171, $P < 0.001$). Note that they don't all have the same number of shared ancestral SNPs because they have different numbers of SNPs supported by sufficient depth (see methods); B) Lineage specific variants have higher allele frequencies in the asexual compared to each of the sexual populations (Wilcoxon rank sum; $N_{Asexual}$ =

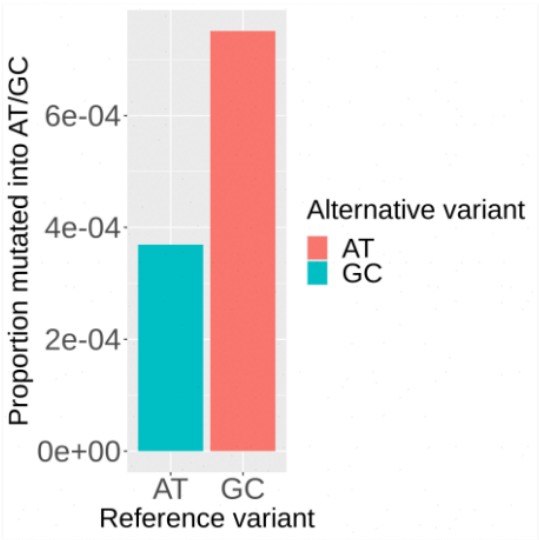
30,463, $N_{Daviesl}$ = 11,438 , $N_{Newmark}$ = 8,978, N_{Oxford} = 5,049, $W_{Asexua\text{-}Daviesl}$ = 277,930,015, $W_{Asexual\text{-}Newmark}$ = 165,802,180 , $W_{Asexual\text{-}Oxford}$ = 86,457,384, $P < 0.001$); C) Average allele frequencies per scaffold of ancestral vs lineage specific variants (split sexual groups) (Wilcoxon rank sum; $N_{Asexual}$ = 363, $N_{Daviesl}$ = 337 , $N_{Newmark}$ = 332, N_{Oxford} = 246, $W_{Asexua\text{-}Daviesl}$ = 82,516, $W_{Asexual\text{-}Newmark}$ = 86,862 , $W_{Asexual\text{-}Oxford}$ =63,322, $P < 0.001$).



Supplementary figure 3 – Punnet square explaining the expected proportion of homozygosity assuming all lineage specific asexual variants were inherited from sexual worms only to be subsequently lost in the sexual population. One third of the potential variants would be expected to be inherited as homozygous.

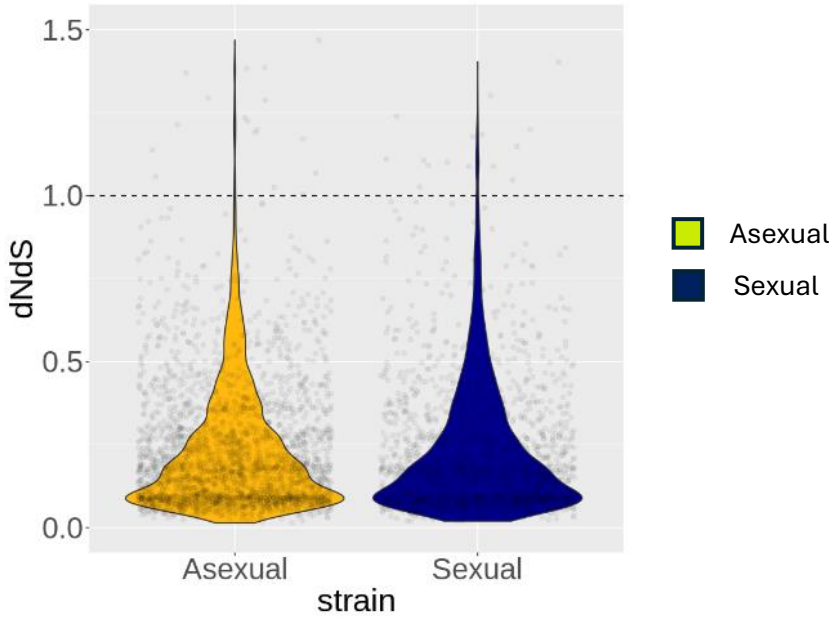


Supplementary figure 4 – The majority of noncoding variants are found within repetitive regions.



Supplementary figure 5 – Proportion of AT variants that mutate into GC and vice versa using synonymous variants. Calculated by the number of synonymous AT to GC mutations divided by the total number of AT nucleotides in CDS regions (and vice versa for GC to AT mutations). This shows that GC nucleotides are more likely to mutate into AT nucleotides in

the coding regions of the planarian genome, meaning that AT to GC mutations cannot explain the GC bias in homozygosity.



Supplementary figure 6 – dN/dS ratios of all genes when only 3 exons with more than 3 synonymous mutations are used. This removes the discrete peaks in the distribution and suggests that they are driven by most exons having very few synonymous mutations.

Chapter 2 - Diversity within and between individual planarians in the absence of sex and a single cell bottleneck.

Abstract

Sexual reproduction *via* a single celled bottleneck evolved independently in all transitions to obligate multicellularity, suggesting that both sex and periodic single cell bottlenecks are important for a multicellular lifestyle. However, some species reproduce solely by fission, thus avoiding both sex and reproductive single cell bottlenecks. This raises two problems; A) How do they generate diversity in the absence of sex; B) How do they avoid the accumulation of somatic diversity in the absence of reproductive single-cell bottlenecks? In this study, we use asexual *Schmidtea mediterranea* that reproduce solely by fission to estimate how much diversity is produced in their population. Additionally, we ask whether the lack of single cell bottlenecks drives the accumulation of somatic diversity, potentially introducing conflict between cells within an organism. By tracking the reproductive behaviour of thousands of individuals and genotyping a subset of worms, we found two previously undescribed strains within a widely used clonal planarian population. One strain reproduced more slowly and suffered higher mortality rates, suggesting that individuals from this strain had lower fitness. The strain with lower fitness showed evidence of increased somatic diversity, consistent with the hypothesis that high somatic diversity may interfere with cellular cooperation over time. Finally, we estimate the per generation mutation rate to be over 700X higher than the rate seen in vertebrates, which may help provide genetic diversity to this population of asexual worms. Overall, we describe unexpected diversity in asexual planarian populations and provide initial evidence for a link between somatic diversity and reduced organismal fitness in an animal that lacks periodic single cell bottlenecks.

Introduction

Obligate multicellularity evolved at least five times, with sexual reproduction *via* a zygotic single cell bottleneck evolving independently in all cases, suggesting that this reproductive strategy is important for multicellular life (Knoll, 2011; West et al., 2015). However, some species reproduce only asexually, either by parthenogenesis or by fission-like processes that do not involve single cell bottlenecks (Brandt et al., 2017; Fields and Levin, 2018; Judson and Normark, 1996; Lázaro et al., 2011; Leria et al., 2019; Wilson et al., 2024). While the evolutionary implications of asexuality in parthenogenic species is well studied, the implications of fission-based reproduction are not well understood and may differ significantly due to the lack of periodic single cell bottlenecks (Howe et al., 2022).

Some challenges, such as generating and maintaining population diversity in the absence of sex, impact both parthenogenic and fissiparous species. Asexual populations cannot generate different allele combinations from across the population, limiting the amount of genetic diversity available (Colegrave, 2012). This is important since reduced diversity limits the adaptability of asexual populations, making them susceptible to environmental change and evolutionary arms races (Hamilton et al., 1990; Lively and Morran, 2014; Vergara et al., 2014). Additionally, sexual populations can combine beneficial alleles from multiple individuals into a single offspring, while in asexual populations beneficial mutations must outcompete and eliminate others to reach fixation, a challenge known as clonal interference (Miralles et al., 1999). Thus, both parthenogenic and fissiparous species are limited in their ability to generate genetic diversity compared to sexual populations and may need to acquire mechanisms to maintain genetic diversity to persist over time.

However, fissiparous species have additional challenges due to the lack of a reproductive single cell bottleneck. In most species, the presence of frequent single cell bottlenecks is

thought to limit the accumulation of selfish cell-level adaptions (Howe et al., 2022). While selection acts between individual organisms it also acts between cells in the body, selecting for cells with mutations that allow them to outcompete their neighbours, potentially at the expense of the organism. This is thought to be one of the reasons that cancer is found in almost all multicellular species, since it is the product of fundamental selective pressures within the soma (Aktipis et al., 2015). Reproduction *via* a single cell bottleneck deals with these problems by resetting clonality and ensuring that the individuals in the next generation consist either entirely of cooperative or selfish cells. Since organisms produced from a mutant selfish cell are unlikely to develop, survive and reproduce, these selfish cell adaptations are periodically purged from the population. There are some interesting exceptions such as transmissible cancers that have transcended the single cell bottleneck, but such events are thought to be rare (Ní Leathlobhair and Lenski, 2022; Stammnitz et al., 2023; Yonemitsu et al., 2019). However, in the absence of frequent single cell bottlenecks, such selfish adaptions may be more common since clonality is not necessarily maintained and any selfish cells that arise can be transmitted to the next generation. This has been shown in aggregative multicellular species of animals, fungi, and slime moulds, whose lifecycle involves the fusing of multiple individuals which can result in high somatic diversity (Buss, 1982; Grum-Grzhimaylo et al., 2021; Stoner et al., 1999). However, little work has been done on fissiparous species to see if they are likely to suffer similar problems due to an accumulation of somatic diversity.

Asexual *Schmidtea mediterranea* are freshwater planarians that reproduce solely by fission, and so lack reproductive single cell bottlenecks (Issigonis and Newmark, 2019; Lázaro et al., 2011). Their body contains abundant pluripotent stem cells that underlie the regenerative abilities for which planarians are often studied (Rink, 2013). Fission results in a small tail fragment and a large head fragment, both of which undergo whole body regeneration, giving

rise to two fully functional planarians (Arnold et al., 2021). Additionally, these planarians also have the ability to fragment, a process by which multiple rounds of fission occur before whole body regeneration is complete, resulting in the production of many offspring in a short period of time (Goel et al., 2021). Asexual *S. mediterranea* are suggested to have a crown age of ~600,000 years, implying that they have been reproducing asexually without a single cell bottleneck for potentially millions of generations (Lázaro et al., 2011). This makes asexual planarians an ideal model to study the evolutionary implications of asexual reproduction in the absence of a single cell bottleneck, allowing us to ask how much diversity they are able to generate, and how they limit diversity from accumulating in the soma.

In this study, we measured the phenotypic diversity in reproductive behaviour and mortality rates across thousands of asexual planarians and genotyped a subset of individuals to understand how phenotypic and genetic diversity arises in clonal planarians. We found that there was significant variation in fitness between planarian populations, and that these differences correlated with genetically distinct asexual strains. We also found that the strain with reduced fitness contained higher somatic diversity, which is consistent with the hypothesis that high levels of somatic diversity may allow selfish cells to arise in the soma. Finally, we assess the mutation rates in these worms and show that they exhibit a per generation mutation rate over ~700X higher than what is typically seen in vertebrates, which is likely driven by the lack of a single cell bottleneck and may help these asexual populations produce diversity in the absence of sex.

Methods

Planarian maintenance

CIW4 asexual *S. mediterranea* planarians were used for all experiments and maintained according to published protocols (Dunkel et al., 2011).

Experimental design

To understand how diversity is generated and distributed in asexual planarian populations we took two approaches. The first was to compare four populations of wildtype asexual planarians, each generated by a single founder and allowed to grow for two and a half years. During this time the reproductive and mortality rates of each population was measured, and a subset of individuals from each population were used for genome sequencing. We also generated lines that were exposed to a cellular bottleneck (with a single irradiated individual rescued by cell injection used as a founder for a whole population – see below), which allowed us to assess the degree of somatic diversity within a single individual. The second experimental approach was to sequence the head offspring of individuals over multiple generations to quantify the mutation rate and assess how much new somatic diversity is transmitted to the next generation via the tail (Fig. 4B). By assessing the number of newly detected mutations in each generation, and whether they are transmitted *via* the tail to the next generation or not we could provide insight into the mutation rate and somatic heritability of mutations in fissiparous planarians.

Generation of bottlenecked planarians

To generate the bottlenecked asexual *S. meditteranea*, sexual *S. mediterranea* planarians were lethally irradiated with 93 Grays (9,300 rads) on a Gammacell 40 Exactor irradiator (Best Theratronics, Springfield, VA). After irradiation, sexual planarians were kept in 10 cm plastic petri dishes (Simport Scientific) in 50 µg/ml gentamycin (Omega Scientific, Tarzana, CA, USA) in 1x Montjuic Salts. Two days following irradiation, irradiated hosts were injected with ~50 cells/planarian from an asexual *S. mediterranea* donor. The donor had been fasted for at least 7 days before dissociation into individual cells. The donor planarian was placed in a 1.7

ml microcentrifuge tube with 1 ml Calcium/Magnesium-Free Medium ("CMF") + 1% BSA (CMFB) (Wang et al., 2018) and by pipetting up and down with a P1000 for 5 min. The dissociated cells were filtered with a 40 µm cell strainer (Fisher Scientific, Hampton, NH) and subsequently centrifuged at 290 x*g* for 5 min at 4 °C. The pellet was resuspended in 0.5 ml CMFB. Roughly 20 µl of the cell suspension was spotted onto an uncoated glass coverslip and inspected using a Nikon Eclipse C*i* compound microscope at 10x magnification. While avoiding cell clumps, individual cells from a region of high density on the coverslip were mouth pipetted using a borosilicate glass needle. The needle was prepared from a capillary tube (World Precision Instruments, product #1B100-4 1.0 mm OD, 0.58 mm ID) using a Sutter Instrument P-1000 micropipette puller and were pretreated with Sigmacote for 5 min before use. The number of cells injected (~50) was determined manually by counting the number of cells entering the needle during the aspiration, using the microscope, however, the proportion of neoblasts is unknown since no markers for neoblast identity were used. The cells were injected by sticking the needle through the tail stripe on the ventral side of the sexual hosts and pointing the needle towards the parenchymal space using 1.5-2.5 psi pressure on an injection rig using a WPI PV 820 Pneumatic PicoPump.

Two out of 180 irradiated sexual planarians that were injected with asexual neoblasts survived. These two rescued planarians (B1 and B2) began showing signs of regeneration approximately 43 days after the rescue and divided 82 days after the rescue. They were confirmed to have been converted to asexual using karyotype analysis. Tail snips from the injected planarians were used for chromosome preparation as described in (Guo et al., 2018). Slides with chromosome spreads were stained with DAPI (Molecular Probes, Thermo Fisher Scientific, Waltham, MA, USA) and imaged using a ZEISS Axio Observer Z1 epifluorescent microscope equipped with a ZEISS Axiocam 702 mono camera with a ×100 oil objective. Ten chromosome spreads from each tail snip were analyzed for the presence of the asexual chromosome

translocation. Tail snips from one asexual and one sexual *S. mediterranea* were used as controls.

Planarian individual maintenance

The progeny of the two rescued individuals were maintained and tracked. Individual planarians were maintained in 10 cm petri dishes (Simport Scientific, QC, CA) in approximately 25 ml 1x Montjuic Salts per worm in the same incubator in the dark, according to previously described protocols (Dunkel et al., 2011). Individual planarians were fed and cleaned following the same schedule as bulk planarians. The planarians were checked every 2-3 days for fission events. Divisions were recorded in the SAPling database (Thomas and Schötz, 2011) and offspring separated into new dishes, as previously described. Each planarian container had a unique QR-code identifier (Thomas and Schötz, 2011). Certain individuals (SI1) were flash frozen and used for whole genome sequencing.

Genomic DNA extraction

Individual planarians were flash frozen in liquid nitrogen and stored at -80°C prior to genomic DNA extraction, which was done using a Wizard® Genomic DNA Purification Kit (Promega, Madison, WI). The protocol for Isolation of Genomic DNA from Animal Tissue and Tissue Culture Cells was used, including sample lysis for 30 minutes at 55°C with shaking. Genomic DNA yield was measured using a Qubit dsDNA BR assay (Thermo Fisher Scientific, Waltham, MA).

Generation of 'head lines'

For a subset of the progeny from the two rescued planarians (B1 and B2), we froze the head offspring while allowing the tail offspring to regenerate and reproduce (SI1). We chose to retain

the tails because they are the smaller offspring and must regenerate more complicated structures, including the brain, thus requiring more stem cell divisions. This provided 3-4 consecutive head offspring for each tail line, resulting in 5 separate lineages for a total of 16 samples. One outlier replicate was removed due to high levels of missing genotype data (Fig. S1).

Non-bottlenecked planarian samples

We used existing frozen worms from previously produced planarian families that were started from arbitrarily selected founding planarians (Carter et al., 2015; Dunkel et al., 2011; Quinodoz et al., 2011; Thomas and Schötz, 2011; Thomas et al., 2012). In total, we used samples from four of these randomly founded families: Fritz (F), Sophie (S), Peter (P) and Hans (H). To allow for direct comparisons between populations, corresponding individuals from the same generational history were used for sequencing (SI1). For each generational timepoint, samples were compared from two different non-bottlenecked families and from the two bottlenecked families.

Next generation whole genome sequencing and data analysis

Individual planarians were flash frozen in liquid nitrogen and stored at -80°C prior to genomic DNA extraction using a Wizard® Genomic DNA Purification Kit (Promega, Madison, WI). Library preparation and next generation whole genome sequencing was performed by Admera Health using a KAPA HyperPrep Kit (Roche, Indianapolis, IN) and Illumina sequencer (paired 150 bp reads, ~200M paired end reads per sample), respectively.

Variant calling

Sequencing data was uploaded to the Galaxy web platform and the public server usegalaxy.org was used to create the trimmed and mapped reads (Afgan et al., 2022). The quality of the reads was checked using fastp (Chen et al., 2018), which also performed adaptor trimming. Reads were mapped to the refence sexual *S. mediterannea* genome (Grohme et al., 2018) using BWA-MEM (Li, 2013). The quality of the mapped reads was checked with ValidateSamFile and duplicates were marked using MarkDuplicates. Variant detection was performed using GATK (Van der Auwera and O'Conner, 2020), focusing on all gene bodies described in the Neiro et al annotation (S2-3) (Neiro et al., 2022). Variants were quality filtered using GATK suggested parameter values (SI4), and we only included variants with a depth between 0.5 and 2 times the mean depth. Repetitive regions and transposable elements were identified in the dd_Smes_g4 genome and removed from the analysis using RepeatModeler and RepeatMasker unless explicitly stated in figure legend (SI5). PLINK v1.9 was used to prune the data and run Principal Component Analysis (Purcell et al., 2007). Pruning was done with a sliding window of 50 SNPs, a step size of 10 SNPs, and an r^2 threshold of 0.1 (SI4). Outliers were identified in the PCA and upon inspection were removed from the analysis since they contained high levels of uncovered variants (Fig. S1).

FST score assessment

We used VCFtools to assess the mean fixation index scores (FST) between each pair of families, with non-overlapping window sizes of 10,000 bps. We also compared FST scores between the strains using similar settings.

Nuclear diversity tests

We used VCFtools v0.1.16 with a non-overlapping sliding window of 10,000 bps to assess the nucleotide diversity within families. When comparing the nucleotide diversity of populations

derived from bottlenecked and non-bottlenecked worms we ran pairwise nucleotide diversity tests between each pair of individuals within each family. The mean nucleotide diversity per scaffold was calculated using the within family nucleotide diversity tests. Similar settings were used when calculating nucleotide diversity across the genome and comparing between strains.

Runs of homozygosity

Runs of homozygosity were called using PLINK, searching for regions that are at least 50Kb long, containing at least 25 SNPs, with neighbouring SNPs being at most 50Kb from one another (SI4, SI6). Bedtools nuc was used to calculate the GC content within and outside runs of homozygosity (Quinlan and Hall, 2010). To assess whether the FS-specific runs of homozygosity are found in other laboratories across the world, we used previously published data about the variants present in RNAseq libraries from different BioProjects (Kershenbaum et al., 2024). Each variant in this dataset has an associated allele frequency within its BioProject, so we plotted the mean allele frequency of each variant in our region of interest to see if there is a deviation from the expected allele frequencies (1 or 0.5), which would indicate mixed zygosity in the population.

dNdScv

To assess the impact of variants on protein sequence and isolate genes and isolate genes containing non-synonymous mutations between the FS and PH strains we ran dNdScv (SI7) (Martincorena et al., 2017).

Somatic allele frequency

To assess the frequency of detected mutations in the soma, we divided the allele depth (AD) supporting the variant by the total allele depth at the site (AD for variant / (AD for variant +

AD for reference)). For example, a heterozygous variant that is found in all cells should have an allele frequency of 0.5.

Newly detected mutations

To quantify the number of new mutations that arose in each generation, we identified all SNPs that were present in a daughter worm, but not in its mother. We then excluded all SNPs that were found in other libraries, excluding any of the progeny of the daughter worm (since a *de novo* mutation may be passed on to the next generation). We also removed any variants that were found in other datasets, including the family dataset and those found in other publications (Kershenbaum et al., 2024). By plotting the cumulative number of discarded putative *de novo* mutations per cross referenced library, we show that after cross referencing, most remaining putative *de novo* mutations likely to be truly unique mutations (Fig. S2). Additionally, we only considered newly detected mutations whose position was covered by all libraries to make sure that it was truly absent from other libraries. We compared the somatic allele frequency of newly detected mutations to old mutations. In this comparison we defined old mutations as those that are also found in all individuals from the "family" dataset (and therefore also found in both PH and FS strains). To calculate mutation rate per generation per bp, we divided the number of newly detected mutations by the number of bps examined which amounts to 3,165,817 (non-repetitive regions within gene bodies).

Results

There is variation in fitness between clonal planarian populations

We measured the mortality and reproductive rates across different populations of asexual planarians and found significant variation in fitness between populations. We found that

mortality rate was ~2X higher in the F and S families compared to the P and H families (Fisher exact test; $N_H = 4329$, $N_P = 5304$, $N_F = 2153$, $N_S = 1604$, $Mortality_H = 8.1\%$, $Mortality_P = 7.4\%$, $Mortality_F = 15.4\%$, $Mortality_S = 16.4\%$, $P_{P\text{-}H} = 1$, $P_{F\text{-}S} = 1$, $P_{P\text{-}F} < 0.001$, $P_{P\text{-}S} < 0.001$, $P_{H\text{-}F} < 0.001$, $P_{H\text{-}S} < 0.001$, P-adjustment = Bonferroni, Fig. 1A). We also found that F and S took over 2X longer between fissions compared to the P and H populations (Kruskal-Wallis omnibus test; $N_H = 3979$, $N_P = 4910$, $N_F = 1823$, $N_S = 1341$, Median-RWT in weeks; $Median_H = 10$, $Median_P = 11$, $Median_F = 24$, $Median_S = 25$; $P_{P\text{-}H} = 1$, $P_{F\text{-}S} = 1$, $P_{P\text{-}F} < 0.001$, $P_{P\text{-}S} < 0.001$, $P_{H\text{-}F} < 0.001$, $P_{H\text{-}S} < 0.001$, P-adjustment = Bonferroni, Fig. 1B). Finally, a higher proportion of the reproductive events consisted of fragmentation in the P and H families compared to the F and S families (Fisher exact test; $N_H = 2398$, $N_P = 3047$, $N_F = 1374$, $N_S = 974$, Proportion of fragmentation events; $Fragmentation_H = 0.38$, $Fragmentation_P = 0.35$, $Fragmentation_F = 0.25$, $Fragmentation_S = 0.27$; $P_{P\text{-}H} = 0.178$, $P_{F\text{-}S} = 1$, $P_{P\text{-}F} < 0.001$, $P_{P\text{-}S} < 0.001$, $P_{H\text{-}F} < 0.001$, $P_{H\text{-}S} < 0.001$, P-adjustment = Bonferroni, Fig. 1C). Taken together, we find that the P and H families have lower mortality rates and higher reproductive potential than the F and S families.

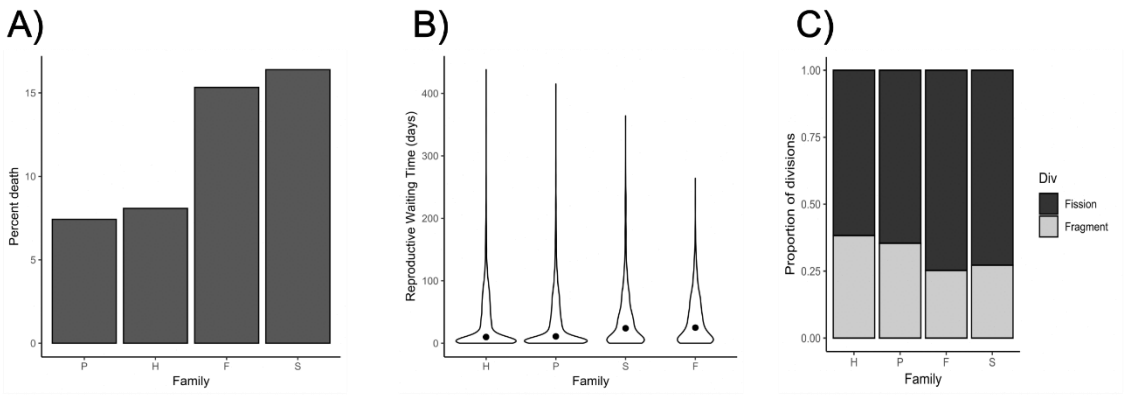


Figure 1 – The P and H families show higher fitness than the F and S families. A) The F and S families have higher mortality rates than the P and H families (Percent of fission products that died = 15.4% and 16.4% vs 8.1% and 7.4% respectively). B) The F and S families spend

more than double the time between fission events compared to P and H (Median reproductive waiting time in weeks = 24 and 25 vs 11 and 10 respectively). C) A lower proportion of the reproductive events in F and S are fragmentation events compared to P and H (Proportion of reproductive events involving fragmentation = 0.25 and 0.27 vs 0.35 and 0.38 respectively).

Population structure correlates with reproductive behaviour

We then tried to identify genetic differences between the populations which may contribute to differences in fitness. Principal Component Analysis (PCA) on the genotypes of the individuals from each family revealed one cluster containing the F and S families, and another more dispersed cluster containing the P and H families (Fig. 2A, Fig. S1). Pairwise fixation index scores (FST) also suggest that these clusters represent distinct subpopulation within our dataset, since comparisons of F/S to P/H had positive FST scores while all within cluster comparisons had negative FST scores (Fig. 2B). When comparing FST scores between the FS (F and S) and PH (P and H) subpopulations, we find that their genetic differences are mostly found at discrete peaks across the scaffolds, rather than spread equally across the genome (Fig. 2C). Overall, our results suggest that the populations can be divided into two subpopulations, the PH families, and the FS families which correspond to the more and less fit phenotypes respectively (Fig. 1-2).

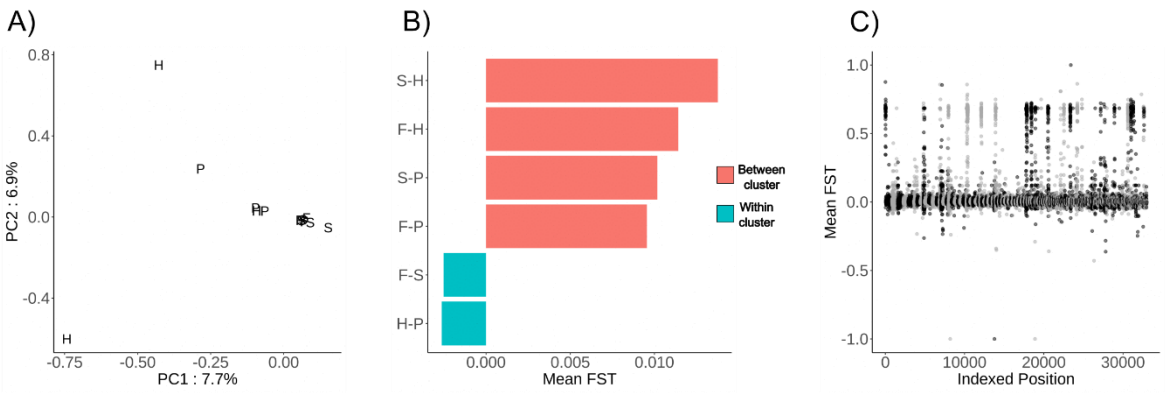


Figure 2 – Population structure analysis reveals two subpopulations in data. A) Individuals from F and S cluster tightly in a PCA, while individuals from P and H are not as tightly clustered and are distinct from F and S. B) Mean pairwise FST scores are positive for between cluster comparisons but negative for within cluster comparisons, suggesting that FS and PH represent distinct subpopulations. C) FST scores across the genome between the FS and PH subpopulations show peaks of genetic distance.

Subpopulations correspond to two distinct strains

We then characterised the differences between the subpopulations and found that they correspond to two distinct strains with different patterns of homozygosity. Most variants were found at least once in both the FS and PH subpopulations (93.2%), however, this still left 5,026 FS specific SNPs and 12,635 PH specific SNPs (Fig. 3A). When comparing the nucleotide diversity across the genome of both subpopulations, we found that PH had peaks of diversity that were absent from the FS population (Fig. 3B).

To test whether lower diversity in the FS population is due to loss of heterozygosity, we isolated the number of runs of homozygosity (long stretches of homozygous regions in the genome), and found that FS contained significantly more runs of homozygosity than PH

(Wilcoxon rank sum; W = 42, N_{PH} = 6, N_{FS} = 7, P < 0.01) (Fig. 3C). These runs of homozygosity cannot be explained by large scale deletions since there is no drop in depth within FS specific runs of homozygosity (Wilcoxon signed rank; V = 94, N = 24, $mean_{FS\text{-}roh}$ = 28, $mean_{outside\ roh}$ = 30, P = 0.114 Fig. S3). Additionally, they are unlikely to be purely the result of GC-biased gene conversion since these runs of homozygosity are AT-rich compared to the rest of the genome (Wilcoxon signed rank; V = 81, N = 35, $mean_{FS\text{-}roh}$ = 0.284, $mean_{outside\text{-}roh}$ = 0.297, P < 0.001 Fig. S3). One of the clearest examples of these FS specific runs of homozygosity can be found in the dd_Smes_g4_1 scaffold and is over 1.3 Mb in length (Fig. 3D). Using this FS-specific run of homozygosity, we found that the two subpopulations are mixed within the CIW4 populations of various labs across the world (Fig. 3E, Fig. S4), suggesting that the widely used CIW4 strain consists of two distinct strains with different patterns of homozygosity.

To understand how these differences between the strains may underlie phenotypic diversity in asexual planarians, we assessed which genes contain strain-specific nonsynonymous mutations. We identified 131 genes with FS-specific nonsynonymous mutations, of which 27 genes contained more than one FS-specific nonsynonymous mutation (SI8-9). Taken together, we describe two distinct strains of asexual *S. mediterranea* found in laboratories across the world, with significant genetic and phenotypic differences.

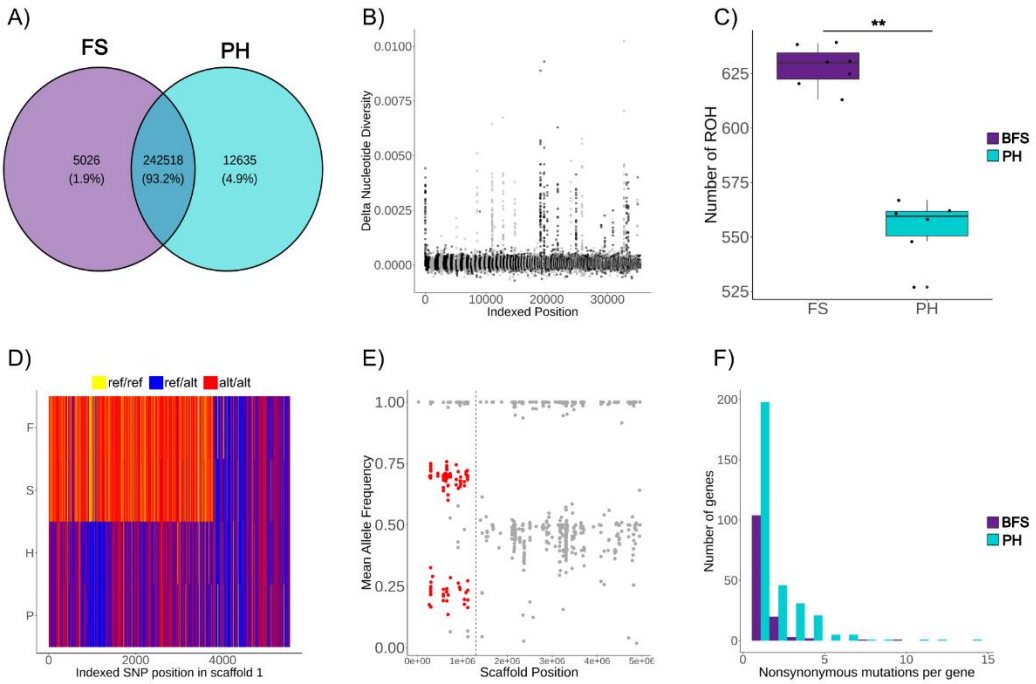


Figure 3 – PH and BFS represent two distinct strains. A) Most variants are shared, but there are still substantial numbers of variants specific to each group. B) PH has peaks of diversity that are absent from FS (but not vice versa). Included scaffolds are coloured in alternating black and grey for ease of visualisation. C) There are more long runs of homozygosity in FS compared to PH. D) Visualisation of a FS-specific run of homozygosity in scaffold dd_Smes_g4_1 (the first 2,000,000 bps). Each line represents a variant that is called in all samples, and their position is indexed based on the position in the scaffold. Note that this figure uses variants that are also found in repetitive regions or transposable elements. This is to allow us to distinguish between the possibility of independent losses of heterozygosity and an ancestral loss of heterozygosity (see Fig. S4 for a similar plot without repetitive regions) E) Mean allele frequencies within a portion of scaffold dd_Smes_g4_1 using various published BioProjects (see chapter 1). This shows inconsistency in zygosity in the region of the FS-specific run of homozygosity, suggesting that FS and PH populations are mixed in CIW4 populations across the world. Variants in FS-specific run of homozygosity are coloured in red, and the dashed line indicates the end of the run of homozygosity. Similar

to figure “D”, this also uses variants included in repetitive regions. F) Number of genes containing strain specific nonsynonymous mutations.

High levels of somatic diversity in strain with low fitness

After finding significant genetic diversity between asexual planarian populations, we tested whether there may also be significant diversity within individual worms due to the lack of reproductive single cell bottlenecks (Fields and Levin, 2018; Howe et al., 2022; Nishimura et al., 2015). Fission rate, which differs between different asexual planarian strains, may impact the potential for intra-organismal diversity, so we assessed the somatic frequency of alternative variants relative to the sexual genome in both strains. In both cases, shared alternative variants (alternative variants relative to the sexual genome that are found at least once in both PH and FS individuals) had a mean allele frequency close to 0.5 in the soma (Fig. 4A). Alternative variants specific to PH also had an average allele frequency close to 0.5. However, FS-specific alternative variants showed a much lower allele frequency in the soma, which may suggest higher levels of somatic diversity (Wilcoxon signed rank; $V = 126{,}575$, $N_{\text{strain specific}} = 7{,}162$, mean = 0.16, $P < 0.001$) (Fig. 4A). While these patterns could potentially be explained by polyploidy, widespread duplications, or strain specific sample contamination, they were all rejected as potential explanations for the allelic imbalance in FS-specific variants (Fig. S5).

To estimate intra-organismal diversity, we generated two populations that were exposed to a cellular bottleneck (see methods, Fig. 4B). Since the bottlenecked populations were of the FS strain, we compared the bottlenecked populations directly to the FS strain (Fig. S6). We found that despite a ~50 cell bottleneck, FS specific variants still showed an allelic imbalance indicative of high intra-organismal diversity (Wilcoxon signed rank; $V = 738{,}188$, $N_{\text{strain specific}}$

= 14,363, mean = 0.15, $P < 0.001$, Fig. 4C). However, strain specific variants found in all individuals of the FS strain (supported in all B1, B2, F, and S individuals) did not show an allelic imbalance (Wilcoxon signed rank; V = 5,432.5, $N_{\text{strain specific}} = 169$, mean = 0.48, P = 0.051, Fig. 4D). Additionally, we found no difference in the per scaffold nucleotide diversity of the unperturbed FS populations and the bottlenecked populations, suggesting that a ~50 cell bottleneck is sufficient to capture the intra-organismal diversity present in the FS strain (Wilcoxon signed rank; $N_{\text{Scaffolds}} = 432$, $\text{mean}_{\text{FS}} = 0.001$, $\text{mean}_{\text{B}} = 0.001$, P = 0.553, Fig.4E). However, since the proportion of neoblasts injected during the experimental rescue of the bottlenecked worms is unknown, it is possible that the 2 out of 180 successfully rescued bottlenecked individuals were successful due to being injected with a high number of neoblasts by chance (potentially close to 50). Thus, it is unclear whether the lack of a genetic bottleneck following the cellular bottleneck is due to low somatic diversity in asexual planarians, or using a founding population (50 cells) that is simply too large to act as a genetic bottleneck. That said, despite this ambiguity, we are able to put an upper limit to the somatic diversity found in individual planarians, since 50 cells (up to 50 neoblasts) are able to capture the genetic diversity of the soma (Fig. 4E). Overall, we find that the PH strain which fissions faster and has lower mortality than the FS strain shows lower levels of allelic imbalance than the FS strain, suggesting that the FS strain contains high levels of intra-organismal diversity, but that this diversity can be captured by a ~50 cell bottleneck.

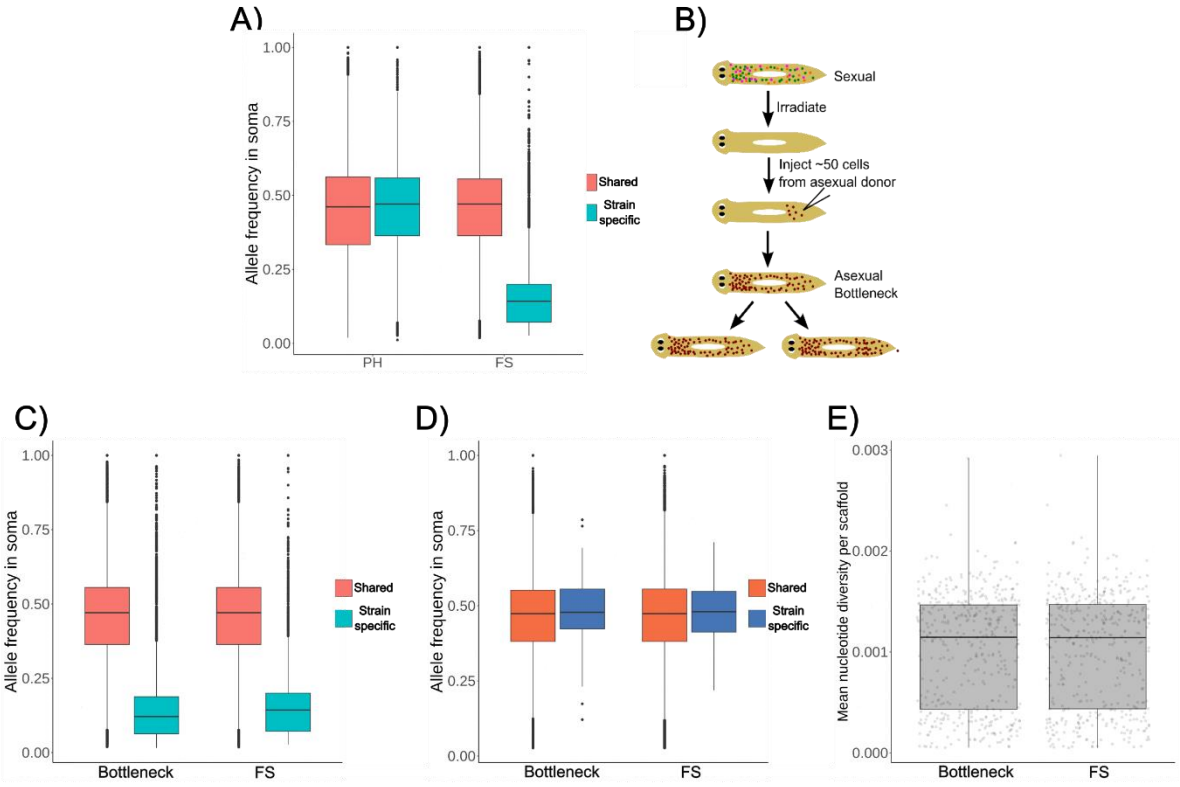


Figure 4 – Differences in intra-organismal diversity between the strains. A) In both strains the shared variants have an allele frequency of 0.5 in the body, suggesting that they are mostly heterozygous in all the cells in the soma. FS specific mutations have an allele frequency of 0.16, suggesting that they are only found in a subset of cells in the soma. B) Generating bottlenecked populations by using injections of 50 cells into a lethally irradiated host. C) Bottlenecked populations of FS strain still have a low allele frequency for FS specific variants. D) Strain specific variants found in all individuals of the FS strain (B1, B2, F, and S) do not show an allelic imbalance. “Bottleneck” refers to the B1 and B2 lineages that were exposed to a cellular bottleneck. E) There is no difference in population diversity between bottlenecked populations and FS populations.

Somatic per generation mutation rates

To estimate the rate of increasing diversity in the soma, we measured the number of new mutations in the soma per generation. This was done by sequencing the consecutive head fragments of individual worms over multiple generations (Fig. 5A). Newly detected variants had lower frequencies than old mutations that were shared by all individuals in the family dataset (Wilcoxon rank sum; W = 6,844,233, N_{new} =1,062, N_{old} = 119261, $mean_{new}$ = 0.2, $mean_{old}$ = 0.47, P < 0.001, Fig. 5B). We detected an average of 297 ± 88.5 (mean ± standard deviation) new mutations per generation, which represents a mutation rate of ~9.4×10^{-5} per bp per generation (see methods) (Fig. 5C, Fig. S). Of these newly detected mutations, only 9.1 ± 3.2% are detected again in the next generation, suggesting that fission may help act as a genetic bottleneck (Fig. 5D).

However, there are alternative explanations for only ~10% of all newly detected variants being detected in the next generation. For example, it is possible that only around 10% of the cycling cells are truly pluripotent, and that the rest of the mutations arise in transient populations of progenitor cells, meaning that over long periods of time the genotype of a planarian is only impacted by a small percentage of pluripotent neoblasts. Additionally, it is possible that only 10% are true *de novo* mutations (verified by their presence in the next generation), and that most of the rest are simply false positives due to sequencing errors or other technical limitations. However, even if we only consider *de novo* mutations that are detected again in the next generation, planarians still show a high mutation rate of ~8.5×10^{-6} per bp per generation. Taken together, we find that only ~9.1% of newly detected variants in the head fragment are found again in the next generation, and that even under conservative estimates the mutation rate per generation is high in asexual planarians.

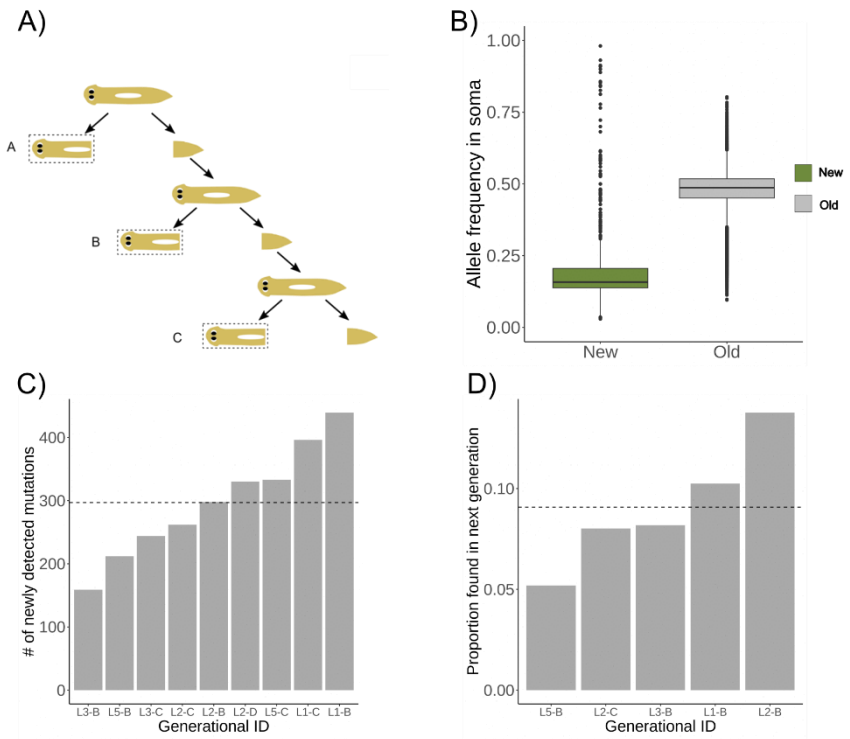


Figure 5 – Somatic mutation rates per generation in asexual planarians. A) The "head dataset" used to assess the number of *de novo* mutations per generation. The heads from natural fission products were collected and sequenced, allowing the tail to regenerate and form the next generation. B) The somatic frequency of newly detected mutations is lower than those found in all individuals from the "family dataset" which includes both PH and FS strains (and so these mutations are predicted to be relatively old). C) There is an average of 297 newly detected mutations per generation. D) Only 9.1% of these newly detected mutations are detected again in the next generation.

Discussion

We investigated the generation of diversity in clonal asexual populations of planarians, and observed how the accumulation of intra-organismal diversity may impact organismal fitness in obligately asexual planarians. We found two genetically distinct strains of asexual *S.*

mediterranea with different reproductive behaviour, mortality rates, and levels of somatic diversity. Overall, we find that obligately clonal populations of planarians are able to generate significant variation in both genotype and phenotype, and provide initial evidence for a potential link between somatic diversity and reduced fitness in animal that lack periodic single cell bottlenecks.

Why are there two distinct strains in a recent clonal population?

Our results suggest that there are two genetically distinct strains within the clonal CIW4 population, with the FS strain containing runs of homozygosity that are absent from the PH strain (Fig. 2-3). These differences can be driven both by asexual mechanisms such as automixis, GC biased gene conversion, mitotic crossing over events, or inbreeding (that requires sexual reproduction) (Ceballos et al., 2018). Given that the long runs of homozygosity can reach lengths of 1.3Mb (Fig. 3E) and that they do not show clear signs of a GC bias (Fig. S3B), we believe that gene conversion is unlikely to have been the major mechanism driving this particular loss of heterozygosity. Instead, this pattern fits better with break induced replication (BIR) which can drive long tracks of homozygosity and is not necessarily associated with a GC bias (Sakofsky and Malkova, 2017; Simion et al., 2021). Additionally, BIR has previously been implicated in the losses of heterozygosity of other asexual sexual species (Simion et al., 2021). That said, these results could also suggest some rare cryptic sex in asexual S. *mediterranea* driving runs of homozygosity by inbreeding. However, we suggest that this is unlikely for a few reasons; 1) Asexual S. *mediterranea* have never been observed to reproduce sexually or develop gonads (Handberg-Thorsager and Saló, 2007; Lázaro et al., 2011; Wang et al., 2007); 2) sexual S. *mediterranea* are cross-fertilising hermaphrodites, meaning that even if a rare asexual individual did undergo sexualisation, it would likely require another sexualised individual to reproduce (Newmark et al., 2008); 3) Since both FS and PH are members of the CIW4 population, a clonal line that was

established in 1999 (Alvarado, 2006; Alvarado et al., 2002), the observed differences between the two strains would have needed to occur since the establishment of the line. This makes rare sexual reproduction an unlikely explanation for the differences between the strains, since it would have needed to occur within laboratory cultures that have been closely monitored for 25 years. It is also worth noting that both strains cluster with all published asexual strains in a PCA, and so are highly unlikely to represent contamination from the sexual strain (Fig. S7).

Nevertheless, it is possible that the CIW4 line is not actually clonal and has been contaminated with individuals from other populations. In this case, it is conceivable that unknown environmental conditions could have sexualized the ancestral FS population, which could have driven inbreeding and resulted in runs of homozygosity that are not found in the PH strain. While we consider this unlikely due to the reasons stated above, it is a real possibility, since other asexual planarian species can be sexualized experimentally, although this has not been achieved in *S. mediterranea* (Hoshi et al., 2003; Nakagawa et al., 2018; Stocchino et al., 2019). Overall, population specific runs of homozygosity suggest the existence of two strains within the clonal CIW4 line, with the origin of these differences having significant implications for our understanding of asexual planarian biology.

Differences in intra-organismal diversity and the potential impact on fitness

In addition to the genetic differences between the strains, we found differences in reproductive behaviour and offspring mortality rates which correlated with differences in somatic diversity between the strains (Fig. 1, Fig. 4). Previous research has shown that the absence of a single cell bottleneck is associated with relatively high somatic diversity in *Dugesia subtentaculata* planarians. While both PH and FS strain reproduce solely by fission, FS had low somatic frequencies for strain specific variants (which likely arose since their split from the PH strain), while PH strain specific variants had allele frequencies close to 0.5 suggesting that they are

found in most of the cells in the body. One potential reason for this difference in somatic diversity is the rate of fission and fragmentation. If neoblasts are not perfectly mixed within the body, then fission and fragmentation events should act as genetic bottlenecks, and the more often these genetic bottlenecks occur the more they might be expected to limit somatic diversity. This is particularly interesting since different species asexual planarians reproduce at different rates with different fragment sizes, which may impact their propensity for accumulating somatic diversity over time (Carter et al., 2015; Goel et al., 2021). Thus, we hypothesise that higher rates of fission and fragmentation in the PH strain allow it to better limit the accumulation of somatic diversity compared to the FS strain.

However, we found that the bottleneck lineages, which are produced from the FS strain, also showed similar alternative allele distributions as non-bottlenecked FS (Fig. 4C). This requires an explanation since if the low alternative allele frequency is driven by clonal mosaicism, then we would expect the cellular bottleneck to drive an increase in alternative allele frequency. One potential explanation is that a 50-cell bottleneck is sufficient to capture all the diversity in the soma. However, there are other potential reasons for this retention of an allelic imbalance following the cellular bottleneck. For example, it is possible that FS contains higher numbers of transiently amplifying progenitors, whose acquired mutations will only be present in a subset of cells. This could explain the unbalanced haplotype and why it remains following a cellular bottleneck since new mutations in transiently amplifying progenitors will continue to emerge. More research will be needed to distinguish between these possibilities, with implications for both planarian stem cell biology and the social evolution in fissiparous multicellular species.

Limiting somatic diversity is important in all multicellular organisms since high levels of intra-organismal diversity may lead to the selection of selfish cell traits such as biased contribution to the germline, or even cancerous transformation (Aktipis et al., 2015; Howe et al., 2022). For

example, colonies of *Botryllus* ascidians that can fuse to form a single soma with a shared circulatory system, have been shown to have evolved strains that invest disproportionally in the germline once fused, parasitising the rest of the colony (Stoner et al., 1999). It has also been suggested that Botryllus evolved mechanisms to limit selfish cell parasitism, by only fusing with closely related individuals that share histocompatibility alleles and thus are likely to share a similarly cooperative phenotype (Laird et al., 2005). High somatic diversity in the FS strain may similarly increase susceptibility to selfish cell evolution, reducing organismal fitness. The higher mortality rate in FS compared to PH supports this hypothesis, but further research is needed to determine if selfish cells exist in the FS strain and whether they contribute to their elevated mortality rates.

Mutation rates in asexual planarians

Mutation rates are central to generating diversity in the soma and in the population, both being of utmost importance to asexual planarians that reproduce solely by fission. Previous research on asexual *Dugesia japonica* planarians has shown that they accumulate diversity rapidly when reproducing clonally, with nonsynonymous mutations being detected in 74% of genes after 20 years of clonal reproduction (Nishimura et al., 2015). Our results also support a high potential for mutation accumulation in asexual populations, with a potential mutation rate of 9.4x10^-5 per bp per generation. This mutation rate per generation is around ~8000X higher than in vertebrates (1.17×10^{-8}).

Even when only considering putative *de novo* mutations that were detected again in the next generation, asexual *S. mediterranea* show a per bp per generation mutation rate of $\sim 8.5 \times 10^{-6}$, which is ~725X higher than those reported in vertebrates and ~650X higher than those reported in sexual *Schmidtea polychroa* (1.31×10^{-8}) that reproduce via a single cell bottleneck (Póti et al., 2024). This high per generation mutation rate in asexual *S.*

mediterranea is likely driven by the lack of a single cell bottleneck allowing many cells to contribute to the next generation, which may help fissiparous species to produce diversity in the absence of sex. However, it is worth noting that these mutation rates were measured in the bottlenecked populations of the FS strain, which we later found to have high levels of intra-organismal diversity (Fig. 4). Thus, it is possible that these high mutation rates are specific to the FS strain, and that other asexual *S. mediterranea* populations may have lower per generation mutation rates. Taken together, we find that asexual planarians of the FS strain have a per generation mutation rate at least ~650X higher than those described in related sexual species, which we hypothesise may help asexual planarians generate diversity in the absence of sex.

In summary, we find that asexual planarians that reproduce solely by fission can generate and maintain significant diversity in their population despite their clonal nature and suggest that this is helped by their somatic inheritance. We also find that of the two newly described strains, the one with lower fitness contains higher levels of somatic diversity. This is consistent with the hypothesis that increased somatic diversity introduces the potential for conflict within the body and raises the question of whether selfish cell adaptations evolve within the planarian soma when intra-organismal diversity is not limited.

Supplementary information

SI1 – Details on individual worms sequenced

SI2 – Table with all SNPs from “Head dataset”

SI3 – Table with all SNPs from “Family dataset”

SI4 – Details of parameters used in code

SI5 – Bed file with transposable elements and repetitive regions to be excluded from analysis

SI6 – All runs of homozygosity

SI7 – Annotated variants (dNdScv output)

SI8 – Genes with PH specific nonsynonymous variants

SI9 – Genes with FS specific nonsynonymous variants

Acknowledgements

This work was done in collaboration with Dr Danielle Ireland and Professor Eva-Maria Collins. The Collins lab collected and sequenced the samples, karyotyped the individuals, and analysed the rates of reproduction and mortality.

Data availability

Supplementary information and data can be found at:

https://github.com/SimonOX-Gene/SmedDiversity

SI2, SI3, SI5, and SI7, are large files that will be uploaded to repositories upon publication. Additional information is available upon request.

Afgan, E., Nekrutenko, A., Grüning, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Ostrovsky, A. E., Mahmoud, A., Lonie, A. J., Syme, A., et al. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research* 50, W345–W351.

Aktipis, C. A., Boddy, A. M., Jansen, G., Hibner, U., Hochberg, M. E., Maley, C. C. and Wilkinson, G. S. (2015). Cancer across the tree of life: cooperation and cheating in multicellularity. *Philosophical Transactions of the Royal Society B: Biological Sciences* 370, 20140219.

Alvarado, A. S. (2006). Planarian regeneration: Its end is its beginning. *Cell* 124, 241–245.

Alvarado, A. S., Newmark, P. A., Robb, S. M. C. and Juste, R. (2002). The *Schmidtea mediterranea* database as a molecular resource for studying platyhelminthes, stem cells and regeneration. *Development* 129, 5659–5665.

Arnold, C. P., Lozano, A. M., Mann, F. G., Nowotarski, S. H., Haug, J. O., Lange, J. J., Seidel, C. W. and Alvarado, A. S. (2021). Hox genes regulate asexual reproductive behavior and tissue segmentation in adult animals. *Nat Commun* 12, 6706.

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. and Bast, J. (2017). Effective purifying selection in ancient asexual oribatid mites. *Nat Commun* 8, 873.

Buss, L. W. (1982). Somatic cell parasitism and the evolution of somatic tissue compatibility. *Proceedings of the National Academy of Sciences* 79, 5337–5341.

Carter, J. A., Lind, C. H., Truong, M. P. and Collins, E.-M. S. (2015). To each his own. *Journal of Statistical Physics* 161, 250–272.

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. and Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* 19, 220–234.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890.

Colegrave, N. (2012). The evolutionary success of sex. *EMBO Rep* 13, 774–778.

Dunkel, J., Talbot, J. and Schötz, E.-M. (2011). Memory and obesity affect the population dynamics of asexual freshwater planarians. *Physical biology* 8, 026003.

Fields, C. and Levin, M. (2018). Are planaria individuals? What regenerative biology is telling us about the nature of multicellularity. *Evol Biol* 45, 237–247.

Goel, T., Ireland, D., Shetty, V., Rabeler, C., Diamond, P. H. and Collins, E.-M. S. (2021). Let it rip: the mechanics of self-bisection in asexual planarians determines their population reproductive strategies. *Phys. Biol.* 19, 016002.

Grohme, M. A., Schloissnig, S., Rozanski, A., Pippel, M., Young, G. R., Winkler, S., Brandl, H., Henry, I., Dahl, A., Powell, S., et al. (2018). The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature* 554, 56–61.

Grum-Grzhimaylo, A. A., Bastiaans, E., van den Heuvel, J., Berenguer Millanes, C., Debets, A. J. M. and Aanen, D. K. (2021). Somatic deficiency causes reproductive parasitism in a fungus. *Nat Commun* 12, 783.

Guo, L., Accorsi, A., He, S., Guerrero-Hernández, C., Sivagnanam, S., McKinney, S., Gibson, M. and Sánchez Alvarado, A. (2018). An adaptable chromosome preparation methodology for use in invertebrate research organisms. *BMC Biology* 16, 1–14.

Hamilton, W. D., Axelrod, R. and Tanese, R. (1990). Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences* 87, 3566–3573.

Handberg-Thorsager, M. and Saló, E. (2007). The planarian nanos-like gene Smednos is expressed in germline and eye precursor cells during development and regeneration. *Dev Genes Evol* 217, 403–411.

Hoshi, M., Kobayashi, K., Arioka, S., Hase, S. and Matsumoto, M. (2003). Switch from asexual to sexual reproduction in the planarian *Dugesia ryukyuensis*. *Integrative and Comparative Biology* 43, 242–246.

Howe, J., Rink, J. C., Wang, B. and Griffin, A. S. (2022). Multicellularity in animals: The potential for within-organism conflict. *Proc. Natl. Acad. Sci. U.S.A.* 119, e2120457119.

Issigonis, M. and Newmark, P. A. (2019). Chapter Four - From worm to germ: Germ cell development and regeneration in planarians. In *Current Topics in Developmental Biology* (ed. Lehmann, R.), pp. 127–153. Academic Press.

Judson, O. P. and Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution* 11, 41–46.

Kershenbaum, S., Jara-Espejo, M., Griffin, A. S. and Aboobaker, A. (2024). Gene conversion limits the cost of asexuality in somatically immortal worms. 2023.03.20.533553.

Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet. Sci.* 39, 217–239.

Laird, D. J., De Tomaso, A. W. and Weissman, I. L. (2005). Stem cells are units of natural selection in a colonial ascidian. *Cell* 123, 1351–1360.

Lázaro, E. M., Harrath, A. H., Stocchino, G. A., Pala, M., Baguñà, J. and Riutort, M. (2011). *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evolutionary Biology* 11, 274.

Leria, L., Vila-Farré, M., Solà, E. and Riutort, M. (2019). Outstanding intraindividual genetic diversity in fissiparous planarians (*Dugesia*, Platyhelminthes) with facultative sex. *BMC Evolutionary Biology* 19, 130.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*.

Lively, C. M. and Morran, L. T. (2014). The ecology of sexual reproduction. *Journal of Evolutionary Biology* 27, 1292–1303.

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Loo, P. V., Davies, H., Stratton, M. R. and Campbell, P. J. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* 171, 1029-1041.e21.

Miralles, R., Gerrish, P. J., Moya, A. and Elena, S. F. (1999). Clonal interference and the evolution of RNA viruses. *Science* 285, 1745–1747.

Nakagawa, H., Sekii, K., Maezawa, T., Kitamura, M., Miyashita, S., Abukawa, M., Matsumoto, M. and Kobayashi, K. (2018). A comprehensive comparison of sex-inducing activity in asexual worms of the planarian *Dugesia ryukyuensis*: the crucial sex-inducing substance appears to be present in yolk glands in Tricladida. *Zoological Lett* 4, 14.

Neiro, J., Sridhar, D., Dattani, A. and Aboobaker, A. (2022). Identification of enhancer-like elements defines regulatory networks active in planarian adult stem cells. 2022.02.03.479047.

Newmark, P. A., Wang, Y. and Chong, T. (2008). Germ cell specification and regeneration in planarians. *Cold Spring Harb Symp Quant Biol* 73, 573–581.

Ní Leathlobhair, M. and Lenski, R. E. (2022). Population genetics of clonally transmissible cancers. *Nat Ecol Evol* 6, 1077–1089.

Nishimura, O., Hosoda, K., Kawaguchi, E., Yazawa, S., Hayashi, T., Inoue, T., Umesono, Y. and Agata, K. (2015). Unusually large number of mutations in asexually reproducing clonal planarian *Dugesia japonica*. *PLOS ONE* 10, e0143525.

Póti, Á., Szüts, D. and Vermezovic, J. (2024). Mutational profile of the regenerative process and de novo genome assembly of the planarian *Schmidtea polychroa*. *Nucleic Acids Research* 52, 1779–1792.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 81, 559–575.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26, 841–842.

Quinodoz, S., Thomas, M. A., Dunkel, J. and Schötz, E. M. (2011). The more the merrier?: Entropy and statistics of asexual reproduction in freshwater planarians. *Journal of Statistical Physics* 142, 1324–1336.

Rink, J. C. (2013). Stem cell systems and regeneration in planaria. *Dev Genes Evol* 223, 67–84.

Sakofsky, C. J. and Malkova, A. (2017). Break induced replication in eukaryotes: mechanisms, functions, and consequences. *Crit Rev Biochem Mol Biol* 52, 395–413.

Simion, P., Narayan, J., Houtain, A., Derzelle, A., Baudry, L., Nicolas, E., Arora, R., Cariou, M., Cruaud, C., Gaudray, F. R., et al. (2021). Chromosome-level genome assembly reveals homologous chromosomes and recombination in asexual rotifer *Adineta vaga*. *Science Advances* 7, eabg4216.

Stammnitz, M. R., Gori, K., Kwon, Y. M., Harry, E., Martin, F. J., Billis, K., Cheng, Y., Baez-Ortega, A., Chow, W., Comte, S., et al. (2023). The evolution of two transmissible cancers in Tasmanian devils. *Science* 380, 283–293.

Stocchino, G. A., Sluys, R., Harrath, A. H., Mansour, L. and Manconi, R. (2019). The invasive alien freshwater flatworm *Girardia tigrina* (Girard, 1850) (Platyhelminthes, *Tricladida*) in Western Europe: new insights into its morphology, karyology and reproductive biology.

Stoner, D. S., Rinkevich, B. and Weissman, I. L. (1999). Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. *Proceedings of the National Academy of Sciences* 96, 9148–9153.

Thomas, M. A. and Schötz, E.-M. (2011). SAPling: a Scan-Add-Print barcoding database system to label and track asexual organisms. *The Journal of experimental biology* 214, 3518–23.

Thomas, M. A., Quinodoz, S. and Schötz, E. M. (2012). Size matters!: Birth size and a size-independent stochastic term determine asexual reproduction dynamics in freshwater planarians. *Journal of Statistical Physics* 148, 663–675.

Van der Auwera, G. A. and O'Conner, B. D. (2020). *Genomics in the cloud: Using ocker, GATK, and WDL in Terra*. 1st Editio. O'Reilly Media.

Vergara, D., Jokela, J., Lively, C. M. and Bronstein, E. J. L. (2014). Infection dynamics in coexisting sexual and asexual host populations: Support for the red queen hypothesis. *The American Naturalist* 184, S22–S30.

Wang, Y., Zayas, R. M., Guo, T. and Newmark, P. A. (2007). Nanos function is essential for development and regeneration of planarian germ cells. *Proc Natl Acad Sci U S A* 104, 5901–5906.

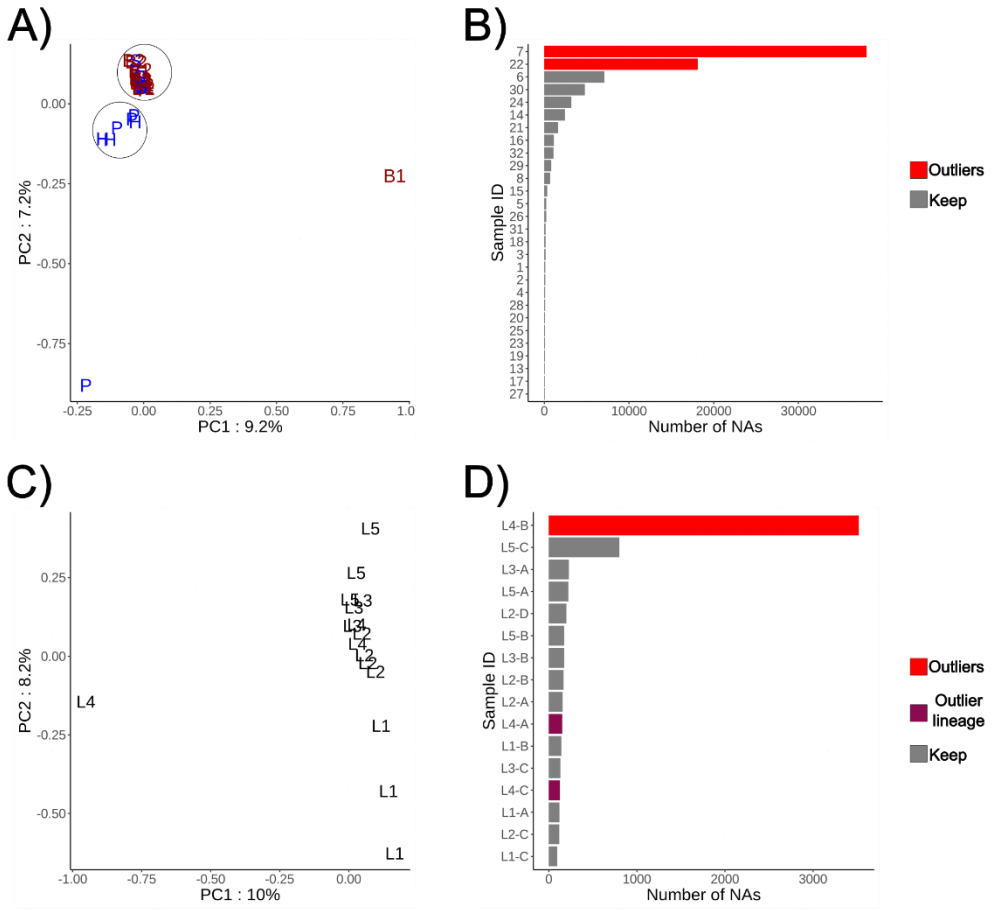
Wang, I. E., Wagner, D. E. and Reddien, P. W. (2018). Chapter 20: Irradiation and single-cell transplantation. 1774, 18–22.

West, S. A., Fisher, R. M., Gardner, A. and Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences* 112, 10112–10119.

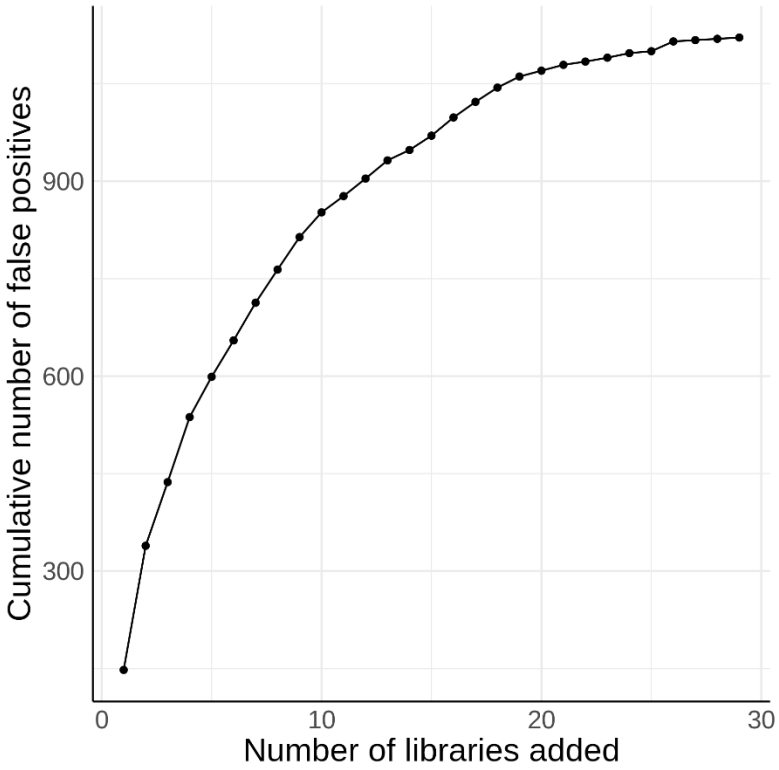
Wilson, C. G., Pieszko, T., Nowell, R. W. and Barraclough, T. G. (2024). Recombination in bdelloid rotifer genomes: asexuality, transfer and stress. *Trends in Genetics* 40, 422–436.

Yonemitsu, M. A., Giersch, R. M., Polo-Prieto, M., Hammel, M., Simon, A., Cremonte, F., Avilés, F. T., Merino-Véliz, N., Burioli, E. A., Muttray, A. F., et al. (2019). A single clonal lineage of transmissible cancer identified in two marine mussel species in South America and Europe. *eLife* 8, e47788.

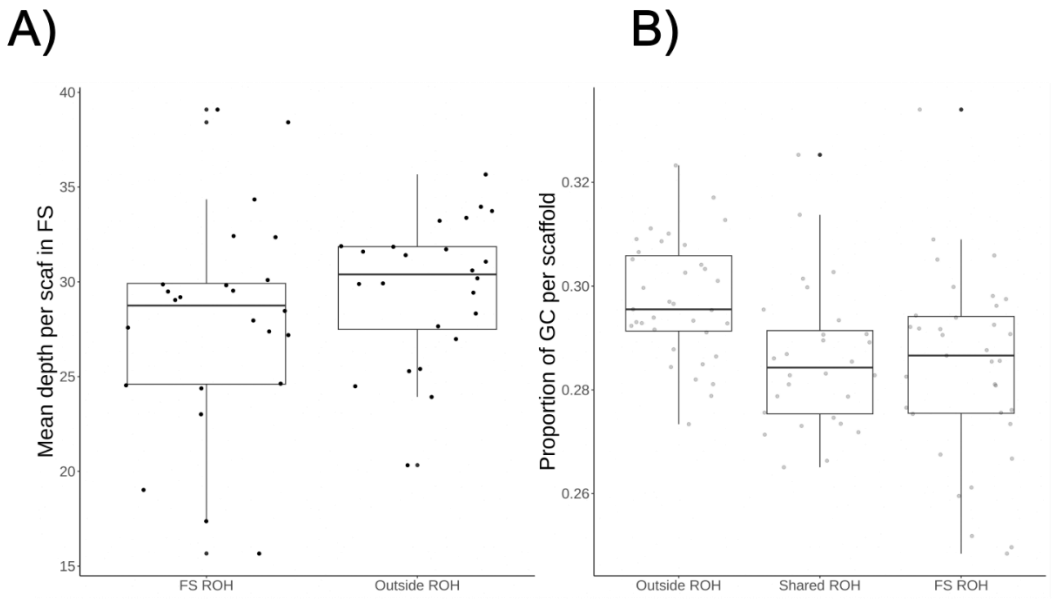
Supplementary figures



Supplementary figure 1 – Outlier removal from "family" and "head" datasets. A) PCA of all family data samples. The letters indicate the families (P,H,F,S,B1,B2) and red indicate the bottlenecked populations. One B1 and one P sample were outliers. B) The same samples have high levels of missing genotyping data. C) PCA of head data shows that one sample from lineage 4 (L4) is an outlier. D) The same sample has a high number of missing genotype data. We removed all of L4 since the detection of new mutations relies on other samples in the lineage.

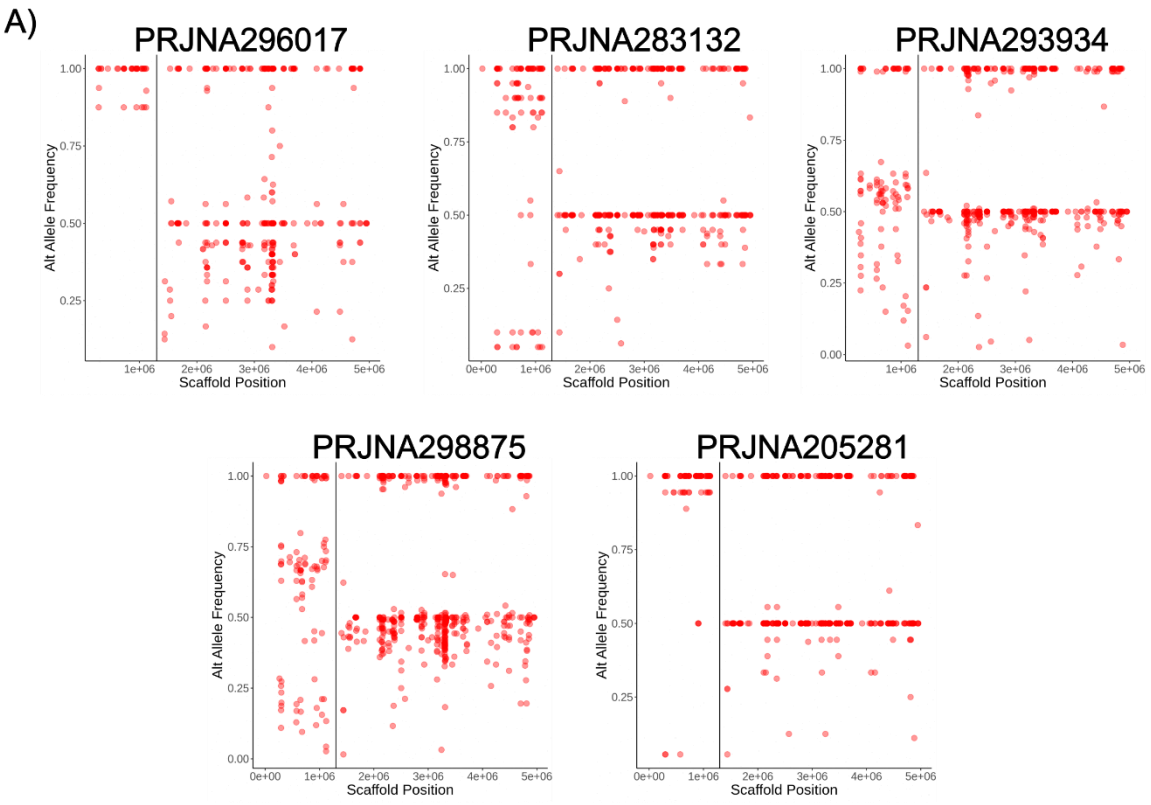


Supplementary figure 2 – The number of false positive *de novo* mutations decreases as I filter out those found in independent datasets. *De novo* mutations were identified from the "head dataset" (see methods) and then compared against unrelated datasets. This cross-referencing allowed us to eliminate most false positives—mutations that also appear in other datasets and are therefore unlikely to be true *de novo* mutations.

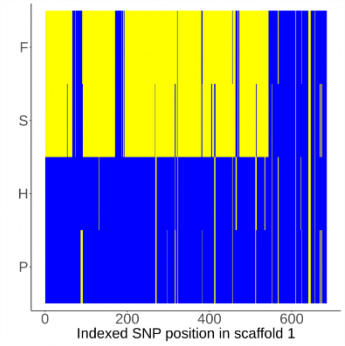


Supplementary figure 3 – Deletions and gene conversion cannot fully explain the FS specific runs of homozygosity. A) Mean depth in FS specific runs of homozygosity is not significantly lower than outside of runs of homozygosity. This was calculated per scaffold (Wilcoxon signed rank; V = 94, N = 24, $mean_{FS\text{-}roh}$ = 28, $mean_{outside\ roh}$ = 30, P = 0.114). B) There is no GC bias in FS specific runs of homozygosity. In fact, FS specific runs of homozygosity are AT rich compared to regions outside of runs of homozygosity. This was also calculated as a mean value per scaffold (Wilcoxon signed rank; V = 81, N = 35, $mean_{FS\text{-}roh}$ = 0.284, $mean_{outside\text{-}roh}$ = 0.297, P < 0.001).

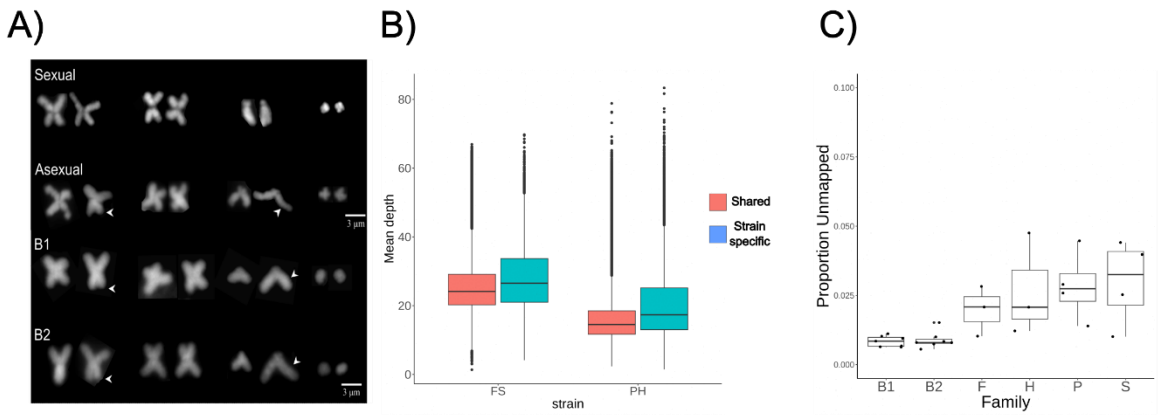
A)


PRJNA296017
PRJNA283132
PRJNA293934
PRJNA298875
PRJNA205281
PRJNA248468
PRJNA245936
PRJNA222859
BCN10
Alt Allele Frequency
Scaffold Position

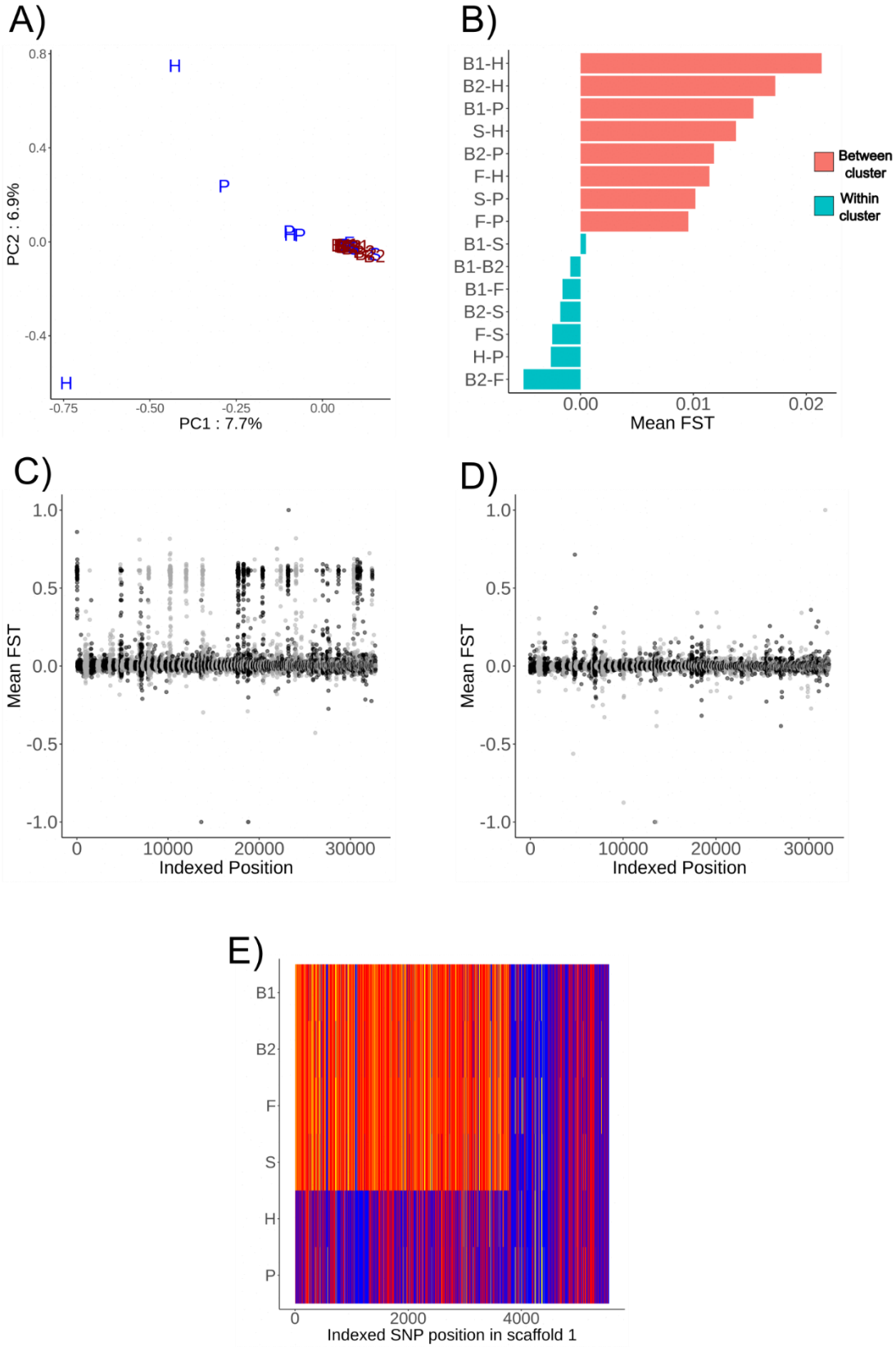
B)


F
S
H
P
0
200
400
600
Indexed SNP position in scaffold 1

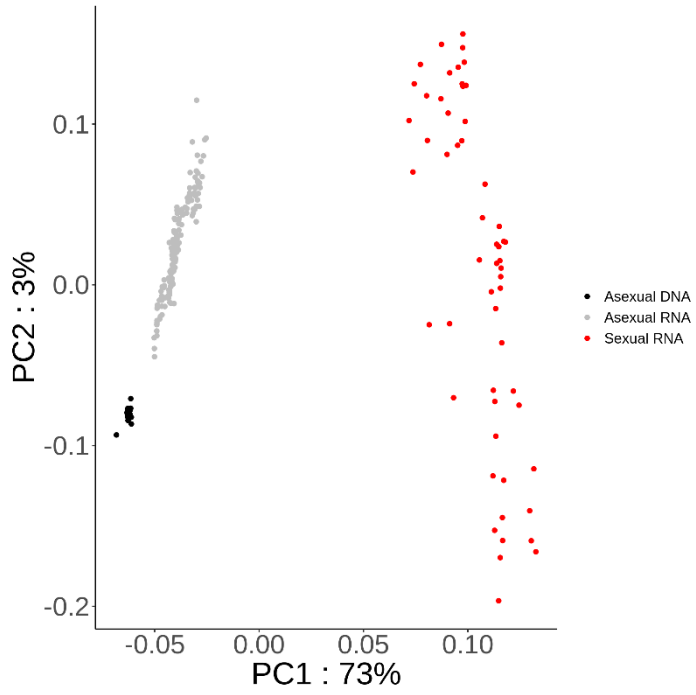
Supplementary figure 4 – CIW4 contains both FS and PH strains in laboratories across the world. A) A mixture of FS and PH strains are found in various, but not all, CIW4 populations. Allele frequency of variants in the dd_Smes_g4_1 scaffold from different BioProjects (stated above each plot). The plot with blue points is from a BCN10 strain, which seems to lack individuals with an FS like genotype. B) Using only variants that are not associated with repetitive regions or transposable elements doesn't allow us to assess whether F and S had independent losses of heterozygosity. Its worth noting that when including only non-repetitive regions, all homozygous variants in this run of homozygosity are of reference identity (Yellow).



Supplementary figure 5 – Allelic imbalance in FS specific variants is not due to polyploidy, local duplications, or contamination. A) Karyotyping on the bottlenecked population (which is are members of the FS strain) shows a normal diploid karyotype. B) FS specific variants have higher sequencing depth than shared variants in the FS strain, suggesting that the allelic imbalance is not due to local duplications (Wilcoxon rank sum; W = 662686528, N_{shared} = 242518, $N_{FS\text{-}specific}$ = 5026, $mean_{shared}$ = 25.7, $mean_{FS\text{-}specific}$ = 27.5, $P < 0.001$). C) Libraries have good mapping rates that are not indicative of a contamination problem.


A)
PC2 : 6.9%
PC1 : 7.7%
H
P
B)
B1-H
B2-H
B1-P
S-H
B2-P
F-H
S-P
F-P
B1-S
B1-B2
B1-F
B2-S
F-S
H-P
B2-F
Between cluster
Within cluster
Mean FST
C)
Mean FST
Indexed Position
D)
Mean FST
Indexed Position
E)
B1
B2
F
S
H
P
Indexed SNP position in scaffold 1

Supplementary figure 6 – The bottlenecked population comes from the FS strain. A) PCA showing that the bottlenecked population (in red), clusters with F and S. B) Comparisons of FST scores between B1/B2 and FS are lower than compared to PH. C) FST scores between the bottlenecked populations and PH show similar peaks to those seen between FS and PH. D) There are no peaks of high genetic divergence between the bottlenecked populations and FS. E) The bottlenecked populations contain the same run of homozygosity in scaffold dd_Smes_g4_1. Red = Homozygous alternative; Yellow = Homozygous reference; Blue = Heterozygous.



Supplementary figure 7 – Both PH and FS cluster with the asexual strains based on variants isolated from RNAseq data in chapter 1. Both PH and FS are shown in black (Asexual DNA), and are shown to cluster together next to the RNA based asexual cluster (none of them cluster next to the RNA based sexual cluster).

Chapter 3 - Comparative approach identifies two novel male germline regulators in planarians

Abstract

In most species, gametes are the only cells that contribute genetic material to the next generation, making them arguably the most important cells in the body. The ancestral metazoan is likely to have regulated germline development by using inductive signals form surrounding cells. However, little is known about the genes that control inductive germline development outside of vertebrates. One exception are planarians which have emerged as a key model for studying inductive germline development due to their ability to continuously generate a germline from adult pluripotent stem cells. Additionally, many planarian species contain asexual strains that lack a mature germline, making them useful for comparative studies to identify genes involved in germline development and more generally understand the changes that follow the transitions to an asexual lifestyle. In this study, we assessed the transcriptional differences between sexual and asexual strains of two planarian species *Schmidtea mediterranea* and *Girardia tigrina*. We found that asexual strains of both species upregulate the expression of cell-adhesion genes, and that genes with no detectable expression in asexual strains show signs of degradation. Additionally, we knocked down transcription factors whose expression differs between sexual and asexual strains, and found that Six1/2 and NK2.2 regulate male germline development in planarians. Overall, our findings support a conserved role for Six1/2 in male germline development in bilaterians and provide the first direct evidence for a role of NK2.2 in male germline development.

Introduction

The germline represents the genetic link between generations and is able to maintain totipotency over evolutionary time, making its development central to all sexually reproducing species (Ewen-Campen et al., 2010; Extavour and Akam, 2003). The mechanisms of germline development vary greatly, with some species relying on maternal determinants to drive germline cell fate, and others on inductive signals from the soma (Seervai and Wessel, 2013). Most research into germline development has focused on the former, despite evidence suggesting that the ancestral metazoan used inductive mechanisms to regulate germline development (Donoughe et al., 2014; Extavour and Akam, 2003; Johnson et al., 2003). To date, most research into the molecular mechanisms of inductive germline development has focused on mice, with little attention given to invertebrate species (Issigonis and Newmark, 2019; Seervai and Wessel, 2013). This limits our ability to identify genes with conserved roles in germline development and constrains our understanding of the diversity in the mechanisms animals use to develop their germline. Thus, more research into invertebrate species is needed to uncover conserved genes involved in inductive germline development and gain a fuller understanding of the mechanisms by which germlines develop across animals.

Although much of the research on the mechanisms of inductive germline development has been conducted in mice, emerging models in other species have improved our understanding of genes with conserved roles in this process. For example, BMP signalling was initially shown to be crucial for the induction of germline development in mice (Chang and Matzuk, 2001; Chu et al., 2004; Hayashi et al., 2002; Lawson et al., 1999; Lopes et al., 2004; Tremblay et al., 2001). However, it was unclear whether BMP signalling was a mammalian specific method for regulating germline development until researchers showed that the injection of BMP-4 RNA into the animal caps of axolotls drove the induction of germ cell fate (Johnson et al., 2003). Later work then found that crickets also rely on BMP signalling

for the induction of germ cell fate, suggesting that the ancestral bilaterian may have similarly relied on BMP signalling for germline development (Donoughe et al., 2014). Additionally, a small subset of genes have been broadly assessed for their roles in germline development across many species, identifying deeply conserved germline regulators. For example, nanos, an RNA binding protein, has been shown to be important in the development of germlines in chordates, ecdysozoans, spirallians, and cnidarians, suggesting that it played a role in germline development in the ancestral metazoan (Ewen-Campen et al., 2010; Extavour and Akam, 2003). This illustrates the value of using species from multiple phyla to infer genes with conserved roles in germline development.

Planarian flatworms are members of the phyla spiralia, which are generally understudied, and have emerged as a model of regeneration and inductive germline development in the last few decades (Grudniewska et al., 2016; Handberg-Thorsager and Saló, 2007; Hoshi et al., 2003; Sato et al., 2006; Shibata et al., 1999; Wang et al., 2007). The most studied planarian species is *S. mediterranea* which contains both sexual and asexual strains that depend on PIWI+ neoblasts to replenish all cell types in their body (Issigonis et al., 2022; Khan and Newmark, 2022; Newmark et al., 2008). The sexual strain is a cross fertilising hermaphrodite, with germ cells being continuously replenished during adulthood, forming an abundance of dorsolateral testes and two ventral ovaries positioned posterior to the brain (Fig. 1) (Issigonis and Newmark, 2019; Morgan, 1901; Wang et al., 2010). In contrast, the asexual strain fails to develop a mature germline and has only ever been observed to reproduce by fission (Issigonis et al., 2022). While asexual *S. mediterranea* do possess nanos+ cells (which in the sexual strain mark primordial germ cells) they do not differentiate into gonadal tissue (Handberg-Thorsager and Saló, 2007; Wang et al., 2007). Previous research has taken advantage of these differences between sexual and asexual strains to isolate genes important for planarian

germline induction, and together with other methods has helped elucidate the mechanisms of germline development in planarians. (Iii et al., 2010; Wang et al., 2010; Zayas et al., 2005).

There are many species of planarians capable of both sexual and asexual reproduction, allowing us to test which genes are associated with sexuality across multiple species (Goel et al., 2021; Issigonis and Newmark, 2019; Stocchino et al., 2019). For example, *Girardia tigrina* is a planarian native to North America that contains both sexual and asexual strains, with the asexual strain being highly invasive (Sluys et al., 2005; Stocchino et al., 2019). Similar to *S. mediterranea*, the sexual strain is a cross fertilising hermaphrodite capable of germline regeneration (Stocchino et al., 2019). However, in contrast to *S. mediterranea*, some asexual *G. tigrina* have been shown to be capable of undergoing rare sexualisation events which are thought to not be possible in *S. mediterranea* (Issigonis and Newmark, 2019; Stocchino et al., 2019). Given that each species will have idiosyncratic differences between the sexual and asexual strains, genes whose expression differs consistently between the strains of different planarian species are more likely to have a conserved role in germline development. Yet to date, no comparative studies have been done to isolate genes that are consistently associated with sexual or asexual strains across multiple planarians species.

In this study, we compare gene expression in the sexual and asexual strains of two planarian species *S. mediterranea* and *G. tigrina*,. After identifying transcription factors whose expression differs consistently between the sexual and asexual strains of both species, we knocked them down and assessed their impact on male germline development. We found two homeobox transcription factors, Six1/2 and NK2.2, with undescribed roles in planarian male germline development. While NK2.2 knockdown leads to the complete lack of testes development, Six1/2 knockdowns seem to inhibit testes maturation. Six1/2 is known to have roles in male germline development in other species such as drosophila and mice, but NK2.2 has no previously described roles in germline development. These findings support a

conserved role for Six1/2 in the bilaterian male germline and provide the first evidence for a functional role of NK2.2 in inductive germline development.

Methods

Worm husbandry

Both *S. mediterranea* and *G. tigrina* were cultured in 0.5% Instant Ocean water at 20°C. Animals were starved for at least 7 days before experiments.

Phylogenetic tree construction

We generated the phylogenetic tree of *S. mansoni* and both strains of *S. mediterranea*, and *G. tigrina*, using the shared BUSCO genes (Manni et al. 2021, Mulhair et al 2023). We ran BUSCO v5.3.2 with the metazoan gene list on the genomes of all groups using publicly available genomes for *S. mansoni* (Berriman et al., 2009), asexual *S. mediterranea* (Robb et al., 2015), and sexual *S. mediterranea* (Grohme et al., 2018b), and assembled draft genomes of sexual and asexual *G. tigrina*. We used MAFFT v7.4 to align the sequences (Katoh et al. 2005) and trimAl to trim the aligned fasta files (Capella-Gutiérrez et al. 2009). PhyKIT was then used to concatenate the aligned fasta files into a super matrix (Steenwyk et al. 2021) and IQ-tree was used to generate a tree from the alignments using 1000 bootstraps (Minh et al. 2020). MEGA was used to infer the phylogenetic trees of genes of interest with their potential human and drosophila orthologues (Tamura et al., 2021). First, the BLAST tool on Planmine was used to identity homologous genes in *S. mediterranea* (Rozanski et al., 2019). Then potential orthologues were similarly identified in the drosophila and human genomes. Since both genes of interest were homeobox genes, the homeobox domains were identified using Pfam, and aligned using MUSCLE (Edgar, 2004; Mistry et al., 2020). The best model (LG +

G) was identified in MEGA, and a maximum likelihood tree was constructed using 500 bootstraps.

Genome sequencing and assembly

To assemble a draft genome for both strains of *G. tigrina* DNA was extracted using phenol-chloroform extraction and sequenced with PacBio (Grohme et al., 2018a). To assemble the genome hifiasm was used at default settings (Cheng et al., 2021). The sexual genome had an N50 of 391,632 and a total length of 1,359,995,635. The asexual genome had an N50 of 114,967 and a total length of 1,039,849,991. The sexual genome contained 62.9% of BUSCO genes while the asexual genome contained 51.2% of BUSCO genes (Fig. S1). This is in contrast to a 67.8% and 55.4% BUSCO completeness in the publicly available sexual and asexual *S. mediterranea* genomes respectively.

Staining and imaging

S. mediterranea and *G. tigrina* were stained with for DAPI to visualise the differences in testes formation in sexuals and asexuals. To this end, worms were killed using 2% HCl and fixed with 4% PFA using standard published protocols (Forsthoefel et al., 2018). DAPI staining was performed at a concentration of 1ng/ml overnight at room temperature. For H3P staining of *S. mediterranea*, animals were killed and fixed using the nitric acid / formic acid protocol, followed by immunostaining for H3P (Guerrero-Hernández et al., 2021). Imaging was done on a Zeiss LSM 880 Airy Scan confocal microscope.

RNAseq and DESQ2

Bulk RNAseq was collected from large whole worms that were that were > 1cm in size using a TRIzol extraction (Liu and Rink, 2018). This large size was chosen to ensure sexual maturity of the sexual strain and to match as much as possible the size between sexually mature *S. mediterranea* and *G. tigrina*. Two individual worms were pooled per sample. Data

was trimmed using trimmomatic-0.39 (Bolger et al., 2014), and pseudo-aligned using Salmon (Patro et al., 2017). Pseudo-mapping was done using the Neiro et al annotation for *S. mediterranea* and the Wheeler annotation for *G. tigrina* (Neiro et al., 2022; Wheeler et al., 2015). The data was then analysed using DESeq2 in R 4.2.0 (Love et al., 2014). Genes were considered to be significantly differentially expressed when they had an adjusted P-value lower than 0.05 and a log2 fold-change higher than 1. We also identified genes that had no detected expression in the asexual strains, while being "highly expressed" in the sexual strains. We defined "highly expressed" as anything with > 5TPM in both sexual strains to exclude lowly expressed genes that may just be absent from the asexual strain by chance.

GO term enrichment analysis

We ran interproscan to on the transcriptomes of both species to get the Pfam annotation which was then used for the GO term analysis (Jones et al., 2014). TopGO was used for GO analysis (Rahnenfuhrer, 2024), and rrvgo and ggplot were used for visualisation of the analysis (Sayols, 2023; Wickham, 2009).

dN/dS analysis

dNdScv was used to assess the dN/dS ratios of genes associated with sexuality (Martincorena et al., 2017). We defined genes associated with sexuality in two ways; A) genes that were upregulated in the sexual strains of both species; B) genes that had no detectable expression in the asexual strains of both species. The variants used were called from DNAseq data of asexual *S. mediterranea* (see chapter 2 - "family dataset"). When comparing dN/dS values between sexual and asexual *S. mediterranea* strains in genes associated with sexuality, we used previously published dN/dS values per exon (Kershenbaum et al., 2024).

Identification of transcription factors

S. mediterranea transcription factors were taken from Neiro el al (Neiro et al., 2022). A similar method was then used to extract potential transcription factors from the *G. tigrina* transcriptome (SI1). Specifically, transcripts that contained an InterProScan description of "transcription factor", or a Pfam hit listed in the transcription factor database DNA binding domain v2.0, or a SUPERFMAILY hit for DNA binding domains were classified as potential transcription factors.

Orthologue identification

The orthologues output from OrthoFinder was used to identify potential orthologues between *G. tigrina* and *S. mediterranea* (SI2). These do not correspond to 1:1 orthologues, but rather a set of potential orthologues. Putative 1:1 orthologues were then extracted from this output in R, by filtering for groups that contained one entry from *S. mediterranea* and one entry from *G. tigrina*.

Gene knockdowns

Candidate genes were cloned into a pPR-T4P plasmid, and dsRNA was synthesised using previously described methods (Adler and Alvarado, 2018; Rouhana et al., 2013). Gene knockdown was achieved by feeding dsRNA in ground liver paste mixed with 1.5% low melting agarose (Rouhana et al., 2013). Specifically, 50µg of dsRNA were mixed in with 25µl of ground liver paste, and 25µl of 1.5% low melting agarose (agarose in 0.5% Instant Ocean planarian water). Worms were fed for 3 weeks (6 feeds), and then starved for a week prior to amputation. Following a week of regeneration, the feeding regime continued for a month (8 feeds), to allow the worms ample time to regenerate their germline. Worms were then starved for a week prior to RNA extraction and fixation.

Results

Independent transitions to asexuality in *Girardia tigrina* and *Schmidtea mediterranea*

Since asexual strains of *G. tigrina* have been previously shown to have the potential to undergo sexualisation, we wanted to test whether our population lacked a mature germline. To this end, we started by DAPI staining sexual and asexual strains of both *G. tigrina* and *S. mediterranea*, assessing the presence or absence of testes in each group. This showed that the asexual strains of both species lacked spermatogonia, and that the sexuals strains exhibited mature testes containing sperm (Fig. 1B-C). This shows that DAPI staining alone is sufficient to assess the presence or absence of testes in both *G. tigrina* and *S. mediterranea* (Wang et al., 2010), and that both asexual strains have their male germline development arrested prior to spermatogonia formation.

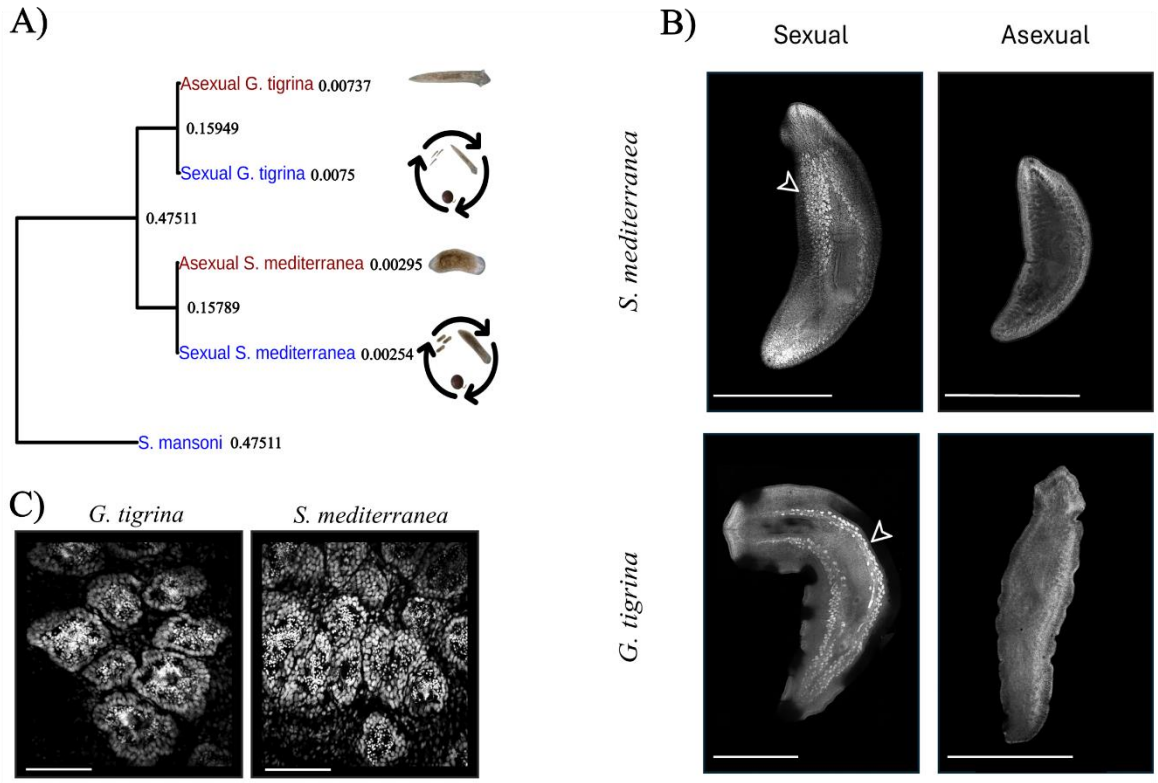


Figure 1 – DAPI staining shows the presence or absence of mature testes in both *G. tigrina* and *S. mediterranea*. A) Phylogenetic tree illustrating the independent transitions to asexuality in both species (red indicates obligate asexuality, blue indicates sexual reproduction). Numbers show sequence divergence from node. B) DAPI staining shows that sexual strains have mature testes while asexual strains do not. Scale bar = 3mm. Arrows point to dorsolateral testes in sexual strains. C) 25X magnification of DAPI stained testes in the sexual strains of *G. tigrina* and *S. mediterranea*. Scale bar = 100µm.

Whole body RNAseq between sexual and asexual strains

To assess the transcriptional differences between sexual and asexual planarians, we ran whole body bulk RNAseq on the sexual and asexual strains of *S. mediterranea* and *G. tigrina*. We found that *S. mediterranea* had more differentially expressed genes between their sexual and asexual strains, with 4,209 differentially expressed genes compared to 1,450 in *G. tigrina* (Fig. 2A, SI3-4). This is despite the strains of *G. tigrina* being more genetically distant than those in *S. mediterranea* (Fig. 1A). In both cases the majority of differentially expressed genes were upregulated in the sexual strain (Fig. 2A, Fig. S2, Binomial test; $N_{Gtig} = 1{,}450$, $N_{Upregulated} = 853$, $P < 0.001$; $N_{Smed} = 4{,}209$, $N_{Upregulated} = 3{,}072$, $P < 0.001$). When comparing putative 1:1 orthologues, we found that 24% were consistently upregulated in the sexual strains of both species, while only 1.8% were consistently downregulated. Given that both asexual strains lacked spermatogonia (Fig. 1B), we tested whether they also exhibited reduced expression of the primordial germ cell markers *nanos* and *klf4* (Issigonis et al., 2022; Wang et al., 2007). We found that while in *S. mediterranea* there was no difference between the strains in *nanos* and *klf4* expression, asexual *G. tigrina* had significantly reduced expression of these primordial germ cell markers (SI3-4). Taken together, we find that most differentially expressed genes are upregulated in the sexual strains, and that despite asexual

G. tigrina showing reduced expression of key primordial germ cell markers, its strains are transcriptionally more similar than those of *S. mediterranea*.

To better understand the broad patterns that consistently differ between sexual and asexual planarians, we performed GO term analysis on the differentially expressed genes between the strains. The genes upregulated in the sexual strains were enriched for GO terms associated with phosphorylation and carbohydrate metabolic processes in both strains (Fig. S2). GO terms associated with microtubule-based movement were also significantly enriched in sexual *S. mediterranea* but were only marginally enriched in the sexual strain of *G. tigrina* (Fig. S2). Genes upregulated in the asexual strain were enriched for cell-adhesion GO terms in both species (Fig. S2, SI5-8). Thus, the genes upregulate in sexual strains are enriched for genes with functions in phosphorylation, metabolic processes, and microtubule-based movement, while those upregulated in asexual strains are enriched for cell-adhesion related genes.

We then wanted to test whether genes associated with sexuality are degraded in the asexual strain of *S. mediterranea*. We found that genes downregulated in both asexual strains did not show higher dN/dS ratios compared to non-differentially expressed genes in asexual *S. mediterranea* (Wilcoxon rank sum; $N_{Background} = 16{,}674$, $N_{Upregulated\text{-}sexuals} = 77$, $P = 0.289$, P-adjustment = Bonferroni, Fig. 2C). However, genes that had no detectable expression in both asexual strains showed higher dN/dS ratios compared to non-differentially expressed genes (Wilcoxon rank sum; $N_{asexual\text{-}absent} = 23$, $N_{Background} = 16{,}674$, $P < 0.005$, P-adjustment = Bonferroni, Fig. 2C). Additionally, genes that had no detectable expression in asexual *S. mediterranea* had a higher proportion of nonsense mutations in their exons compared to genes that were not differentially expressed (Fig. S3). In total there were 23 genes that had no detectable expression in the asexual strains of both species, but only a single gene (sodium neurotransmitter symporter, slc6a-20) also showed high expression in the sexual strains of both species (Fig. 2D; see methods). Thus, while genes with no asexual expression degrade

over time, genes that are downregulated in asexual strains still maintain genetic integrity and thus may retain functions that are still necessary in asexual planarians.

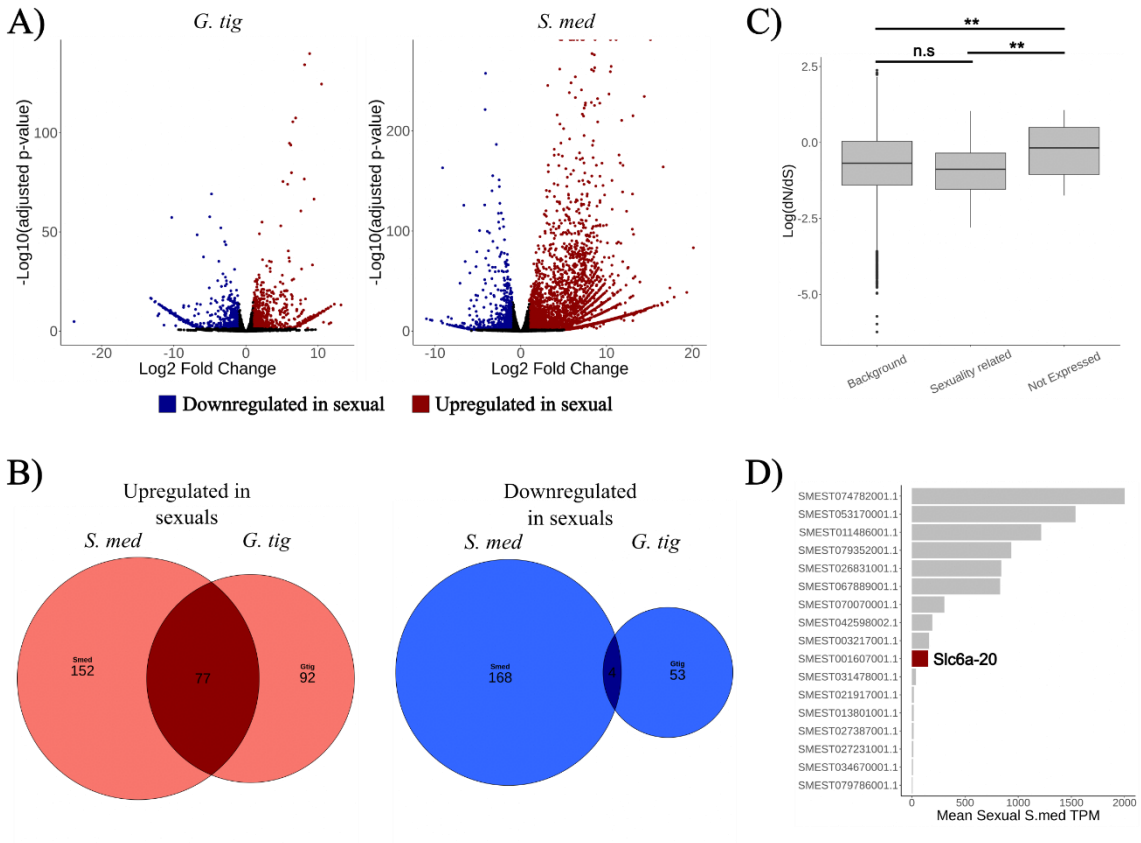


Figure 2 – Broad differences in gene expression between sexual and asexual strains. A) Volcano plots of differential gene expression in both species. *S. mediterranea* has more differentially expressed genes. B) Venn diagrams of differentially expressed genes between sexual and asexual strains in both species using only 1:1 orthologues. There is a 13X higher percentage of overlapping 1:1 orthologues in genes that are upregulated in the sexual strain compared to those downregulated. C) Genes that have no detectable expression in both asexual strains have higher dN/dS ratios indicative of degradation, while genes that are simply upregulated in both sexual strains show no sign of degradation, suggesting that they have pleotropic functions. D) The only gene that has no detectable expression in both asexual

strains, while being expressed in the sexual strains of both species is Slc61-20 (SMEST001607001.1 is *S. mediterranea*).

Knockdown of conserved transcription factors

Following the identification of genes that consistently differ between sexual and asexual strains, we focused on identifying transcription factors that function to regulate germline development. We identified ten transcription factors whose putative orthologues differ consistently in expression between sexual and asexual strains in both species. Of these transcription factors seven were consistently upregulated and three were consistently downregulated in the sexual strains (Fig. 3A). Two of our candidate transcription factors have previously been shown to play key roles in male germline formation in sexual *S. mediterranea*; dmd1 (Chong et al., 2013) and eIF3c (Wang et al., 2010). To test whether our candidate genes regulate germline development in sexual planarians, we knocked down the genes in sexual *S. mediterranea* using dsRNA. A principal component analysis of gene expression data showed that the knockdowns formed three broad clusters, with PC1 capturing 96% of the variance in the data (Fig. 3C, Fig. S4, SI9-18). The first cluster contains the GFP control group, untreated sexual *S. mediterranea*, and five transcription factor knockdowns that show very little impact on expression patterns and thus likely represents genes whose knockdown had little to no impact on germline development. The second cluster contained the knockdowns of genes known to control male germline development (eIF3c and dmd1, shown as #3 and #4 respectively), as well as untreated asexual *S. mediterranea*, and SMEST069539001.1 (gene #10). The third cluster was intermediate and consisted only of gene SMEST001687001.1 (gene #6). Finally, SMEST042630001.1 (gene #9) did not cluster to any specific group, and did not show a significant knockdown following dsRNA treatment,

thus it was excluded from further analysis (Fig. S4). Of all knockdowns, only SMEST001687001.1 (gene #6) led to an observable morphological phenotype, with all individuals showing a complete loss of eyes following 3 weeks of dsRNA feeding in homeostasis (Fig. 3D). Taken together, our knockdowns identified two novel genes with potential roles in germline development, SMEST001687001.1 and SMEST069539001.1.

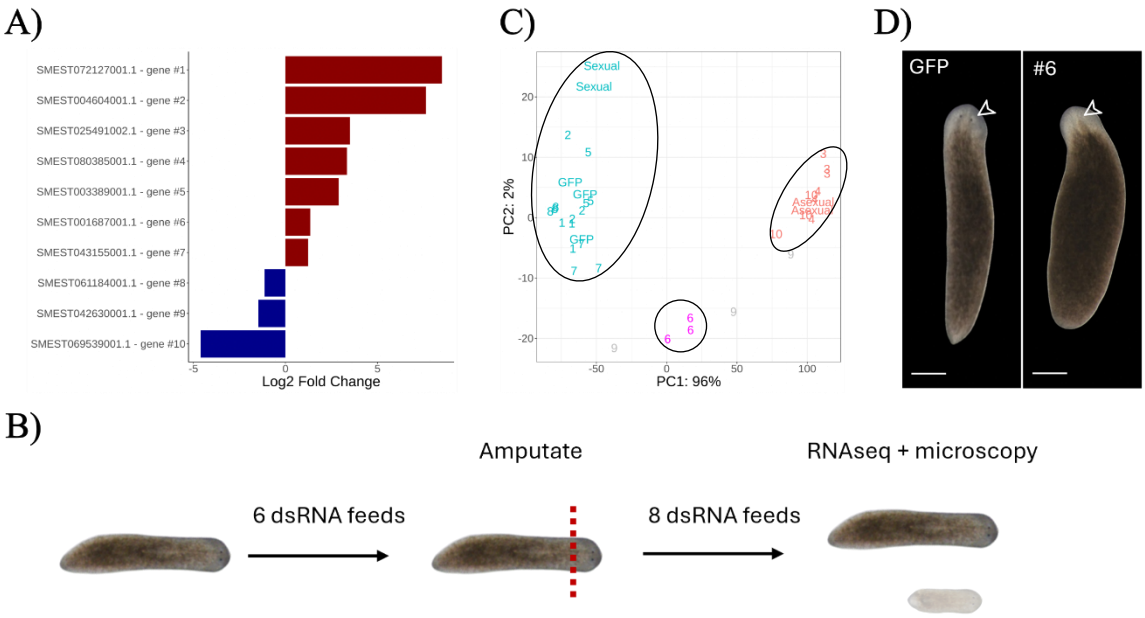


Figure 3 – Knockdowns of candidate transcription factors. A) Bar plot showing the 10 transcription factors that are consistently differentially expressed between sexuals and asexuals of both species. On the X axis is the log2 foldchange in *S. mediterranea*. B) Knockdown regime the worms were exposed to (see methods). Only trunk pieces were used for further analysis. C) PCA of gene expression data following gene knockdown (and wildtype sexual / asexual strains) shows that there are 3 broad clusters; i) Sexual cluster (blue); ii) Asexual cluster (red); iii) intermediate cluster (magenta). D) The knockdown of Gene #6 (SMEST001687001.1) led to a clear loss of eyes after 3 weeks of feeding in homeostasis (n = 15). Scale bar = 0.5mm.

Smed-NK2.2 and Smed-Six1/2 regulate male germline formation.

To test whether our candidate genes play a role in germline formation, we assessed whether their knockdowns led to a loss of testes development. We started by constructing gene trees to identify the orthologues of genes #6 and #10, which showed that they are orthologous to the human Six1/2 and NKx2.2 respectively (Fig. S5, SI19-20). We then analysed the number and identity of differentially expressed genes following the knockdown of NK2.2 and Six1/2, and compared these to the knockdown of dmd1, which was previously shown to regulate male germline formation in planarians (Chong et al., 2013). The knockdowns resulted in the downregulation of 3,025 genes in dmd1, 3,034 genes in NK2.2, and 2,535 genes in Six1/2 (Fig. 4A). In all cases there were fewer upregulated genes in response to the knockdown, with 69 in dmd1, 150 in NK2.2, and 50 in Six1/2 (Fig. 4A). While Six1/2 KDs had the fewest downregulated genes, they showed two genes with significantly lower P-values compared to the dmd1 and NK2.2 KDs. These genes corresponded to SMEST068770001.1 and SMEST003216001.1, with only the later having a significant BLAST result on planmine, revealing it to be a member of the endothelin-converting enzymes that are known to be expressed in human eyes (Rosenthal and Fromm, 2011). Most downregulated genes were shared by all three knockdowns (73.9%, corresponding to 2,401 genes), while only 6 genes were similarly upregulated by all three knockdowns (Fig. 4B). GO term analysis on the shared downregulated genes showed that they were enriched for cell division genes and microtubule-based movement, which could be the result of lost meiotic processes and a lack of sperm formation (Fig. 4C).

To explicitly test if NK2.2 and Six1/2 regulate germline development we focused on assessing whether their knockdowns led to a loss of testes development. While the control GFP knockdowns all showed testes development (n = 5), dmd1 and NK2.2 knockdown worms completely lacked testes (n = 3 and n = 4 respectively, Fig. 4D). Six1/2 knockdowns

had no mature testes, however, some worms showed signs of nuclei condensation at the dorsolateral flanks (2/3), suggesting that the initial stage of testes development can occur in the absence of Six1/2 expression and that its regulatory roles are downstream of dmd1 and NK2.2 (Fig. 4D). Taken together, our results suggest that NK2.2 is indispensable for the formation of testes in planarians, and that Six1/2 likely acts downstream of NK2.2 to drive testes maturation or maintenance.

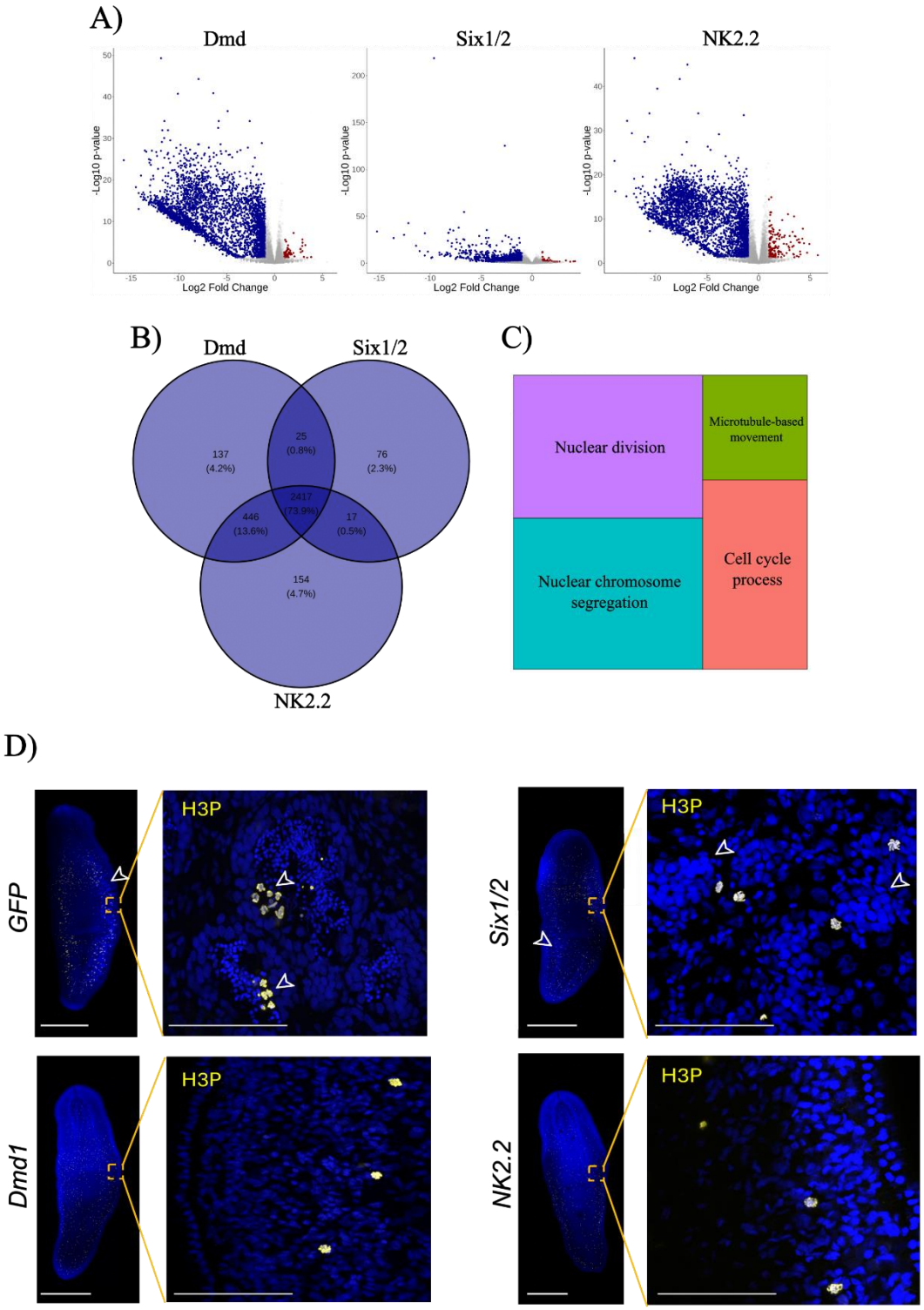


Figure 4 –NK2.2 and Six1/2 regulate testes development. A) Changes in gene expression following gene knockdown of dmd1, Six1/2, and NK2.2. dmd1 and NK2.2 have more differentially expressed genes and they have a higher log2 foldchange. Most significantly differentially expressed genes are downregulated. B) 73.9% of downregulated genes are

shared by dmd1, Six1/2, and NK2.2, suggesting that they regulate similar processes. C) GO term analysis on the genes downregulated in all three knockdowns. The enrichment of cell cycle related terms and microtubule related terms is in line with the loss of meiotic processes and sperm formation. D) Assessing the presence or absence of testes in the knockdowns using DAPI staining and H3P immunostaining to visualise the testes and proliferating spermatocytes. This shows that NK2.2 knockdown leads to the complete loss of testes formation (n = 4), and that Six1/2 knockdown inhibits testes maturation (n = 3), but may allow for the initial steps of testes development (n = 2/3). Scale bars; whole worm = 1mm; insets = 100µm.

Discussion

Using a comparative approach, we identify two previously unknown functions of NK2.2 and Six1/2 in planarian testes development. Six1 has previously been shown to regulate the SRY gene in mice, and its knockout (together with Six4) led to the feminisation of the gonads (Fujimoto et al., 2013). Additionally, Fujimoto *et al* (2013) found that the double knockout of Six1 and Six4 led to a reduction in gonad size in both sexes. This suggests, that Six1 has both male specific and general roles in the development of gonads in mice. In *Drosophila*, sine oculis (SO) is the orthologue of Six1/2 in mice, and was also shown to play a role in male germline development (Fabrizio et al., 2003). Specifically, it was shown that SO was required somatically to regulate spermatocyte differentiation, and that in its absence spermatocytes degenerated (Fabrizio et al., 2003). Our finding that Six1/2 is required for proper testes development in planarians is the first example of a role for Six1/2 in spiralian testes, and further supports an ancient role of this gene in male gonad development.

In contrast to Six1/2, NK2.2 has not been widely studied for its roles in germline development. NK2.2 has primarily been shown to act as a transcriptional repressor regulating neural and pancreatic cell identities (Cowden and Levine, 2003; Mio et al., 2022; Wei et al., 2005). That being said, it has been shown to be expressed in the human testes (Mio et al., 2022), suggesting that it may have a role in testes development in humans. A key question that our data raises is why NK2.2 expression is upregulated in the asexual strain if its knockdown leads to a complete lack of testes development. One possible explanation is that NK2.2 functions early in germline development and is later downregulated once the germline matures. This could lead to higher expression in asexuals since their germline never matures. However, in situs for NK2.2 at different stages of sexual planarian development will be needed to assess this hypothesis. Either way, we find that NK2.2 acts as a key regulator of testes development in planarians, a role that has not been previously shown for this gene in other taxa.

However, it is worth noting that despite comparing gene expression between strains in two species, our functional work focused on *S. mediterranea*. This is due to the ease of working with an established model. However, it raises the possibility that NK2.2 and Six1/2 have different roles in *G. tigrina*, despite having similar differences in expression between the strains. Additionally, we focused on testes development, meaning that it is possible that NK2.2 and Six1/2 have additional roles in the development of the female germline. This is particularly interesting in the context of Six1/2 which has been implicated in male specific roles in germline development, but to our knowledge has never been studied in a hermaphrodite.

In addition to identifying transcription factors involved in planarian germline development, we assessed the broad differences between sexual and asexual strains. We found that asexual strains upregulate genes associated with cell adhesion in both *S. mediterranea* and *G. tigrina*. Since asexual strains reproduce regularly by fission, they may require different cell adhesion properties, which might explain the changes in expression of cell-adhesion related genes (Arnold et al., 2021; Goel et al., 2021). However, another intriguing possibility is that the upregulation of cell-adhesion molecules reduces cell mixing in the soma, which may help limit somatic diversity in the absence of a single cell bottleneck (Howe et al., 2022). Thus, we found that asexual planarian strains of different species consistently upregulate genes with functions in cell-adhesion and propose that this may be due to their roles in fission and limiting cell mixing in the body.

Finally, we found that genes with no expression in the asexual strains show signs of degradation, while no such signature is seen in genes that are consistently downregulated in asexual strains (Fig. 2C, Fig. S3). Previous authors have found that genes with known roles in the germline are also highly expressed in neoblasts, suggesting that they have pleotropic functions, and potentially explaining why they are not degraded in asexual planarians (Guo et al., 2006; Reddien et al., 2005; Salvetti et al., 2005; Shibata et al., 1999; Solana et al., 2009). However, genes with no detectable expression in asexual strains are likely to have roles specific to sexual reproduction, and so might be expected to degrade over time in obligately asexual strains. Thus, our results suggesting the degradation of such genes in asexual *S. mediterranea*, but not in sexual *S. mediterranea*, further supports the hypothesis that asexual *S. mediterranea* are unable to reproduce sexually. Taken together, our results suggest that many genes upregulated in the sexual strains have pleotropic functions beyond the germline, while genes with no detectable expression in the asexual strains likely have functions related to sexual reproduction.

Overall, we assessed the differences in gene expression between sexual and asexual planarian strains of two species and identified transcription factors with crucial roles in planarian male germline development. We found that asexuality is associated with the upregulation of cell-adhesion genes in both *S. mediterranea* and *G. tigrina*. Additionally, we found that genes upregulated in sexual strains show no signs of degradation, suggesting that they may have pleotropic functions, while genes with no detectable expression in asexual strains are degraded. Finally, we identify two new transcription factors, Six1/2 and NK2.2, with roles in planarians male germline development. These discoveries may help us understand the changes that occur following the transition to asexuality in planarians, and how inductive germline development is controlled across metazoans more broadly.

Supplementary information

SI1- List of putative *G. tigrina* transcription factors

SI2- Orthofinder *S. mediterranea* and *G. tigrina* orthologues

SI3- Sexual vs Asexual *S. mediterranea* DESeq2 output

SI4- Sexual vs Asexual *G. tigrina* DESeq2 output

SI5- TopGO analysis, genes upregulated in sexual *G. tigrina*

SI6- TopGO analysis, genes upregulated in sexual *S. mediterranea*

SI7- TopGO analysis, genes upregulated in asexual *G. tigrina*

SI8- TopGO analysis, genes upregulated in asexual *S. mediterranea*

SI9-SI18 – DESeq2 outputs for transcription factor knockouts

SI19- Gene tree for NK2.2

SI20- Gene tree for Six1/2

Acknowledgements

I would like to thank Peter Mulhair for his helpful advice on building phylogenetic trees.

Data availability

Supplementary information and data can be found at:

https://github.com/SimonOX-Gene/ComparativeGermline

Additional information is available upon request.

Bibliography

Adler, C. E. and Alvarado, A. S. (2018). Systemic RNA interference in planarians by feeding of dsRNA containing bacteria. In *Planarian Regeneration: Methods and Protocols* (ed. Rink, J. C.), pp. 445–454. New York, NY: Springer.

Arnold, C. P., Lozano, A. M., Mann, F. G., Nowotarski, S. H., Haug, J. O., Lange, J. J., Seidel, C. W. and Alvarado, A. S. (2021). Hox genes regulate asexual reproductive behavior and tissue segmentation in adult animals. *Nat Commun* **12**, 6706.

Berriman, M., Haas, B. J., LoVerde, P. T., Wilson, R. A., Dillon, G. P., Cerqueira, G. C., Mashiyama, S. T., Al-Lazikani, B., Andrade, L. F., Ashton, P. D., et al. (2009). The genome of the blood fluke *Schistosoma mansoni*. *Nature* **460**, 352–358.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Chang, H. and Matzuk, M. M. (2001). Smad5 is required for mouse primordial germ cell development. *Mech Dev* **104**, 61–67.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170–175.

Chong, T., Collins, J. J., Brubacher, J. L., Zarkower, D. and Newmark, P. A. (2013). A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nat Commun* **4**, 1814.

Chu, G. C., Dunn, N. R., Anderson, D. C., Oxburgh, L. and Robertson, E. J. (2004). Differential requirements for Smad4 in TGFβ-dependent patterning of the early mouse embryo. *Development* **131**, 3501–3512.

Cowden, J. and Levine, M. (2003). Ventral dominance governs sequential patterns of gene expression across the dorsal–ventral axis of the neuroectoderm in the *Drosophila* embryo. *Developmental Biology* **262**, 335–349.

Donoughe, S., Nakamura, T., Ewen-Campen, B., Green, D. A., Henderson, L. and Extavour, C. G. (2014). BMP signaling is required for the generation of primordial germ cells in an insect. *Proceedings of the National Academy of Sciences* **111**, 4133–4138.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Ewen-Campen, B., Schwager, E. E. and Extavour, C. G. M. (2010). The molecular machinery of germ line specification. *Molecular Reproduction and Development* **77**, 3–18.

Extavour, C. G. and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**, 5869–5884.

Fabrizio, J. J., Boyle, M. and DiNardo, S. (2003). A somatic role for *eyes absent (eya)* and *sine oculis (so)* in *drosophila* spermatocyte development. *Developmental Biology* **258**, 117–128.

Forsthoefel, D. J., Ross, K. G., Newmark, P. A. and Zayas, R. M. (2018). Fixation, processing, and immunofluorescent labeling of whole mount planarians. *Methods Mol Biol* **1774**, 353–366.

Fujimoto, Y., Tanaka, S. S., Yamaguchi, Y. L., Kobayashi, H., Kuroki, S., Tachibana, M., Shinomura, M., Kanai, Y., Morohashi, K., Kawakami, K., et al. (2013). Homeoproteins Six1 and Six4 regulate male sex determination and mouse gonadal development. *Developmental Cell* **26**, 416–430.

Goel, T., Ireland, D., Shetty, V., Rabeler, C., Diamond, P. H. and Collins, E.-M. S. (2021). Let it rip: the mechanics of self-bisection in asexual planarians determines their population reproductive strategies. *Phys. Biol.* **19**, 016002.

Grohme, M. A., Vila-Farré, M. and Rink, J. C. (2018a). Small- and large-scale high molecular weight genomic DNA extraction from planarians. In *planarian regeneration: Methods and protocols* (ed. Rink, J. C.), pp. 267–275. New York, NY: Springer.

Grohme, M. A., Schloissnig, S., Rozanski, A., Pippel, M., Young, G. R., Winkler, S., Brandl, H., Henry, I., Dahl, A., Powell, S., et al. (2018b). The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature* **554**, 56–61.

Grudniewska, M., Mouton, S., Simanov, D., Beltman, F., Grelling, M., de Mulder, K., Arindrarto, W., Weissert, P. M., van der Elst, S. and Berezikov, E. (2016). Transcriptional signatures of somatic neoblasts and germline cells in *Macrostomum lignano*. *eLife* **5**, e20607.

Guerrero-Hernández, C., Doddihal, V., Mann, F. G. and Alvarado, A. S. (2021). A powerful and versatile new fixation protocol for immunohistology and in situ hybridization that preserves delicate tissues in planaria. 2021.11.01.466817.

Guo, T., Peters, A. H. F. M. and Newmark, P. A. (2006). A bruno-like Gene Is Required for Stem Cell Maintenance in Planarians. *Developmental Cell* **11**, 159–169.

Handberg-Thorsager, M. and Saló, E. (2007). The planarian nanos-like gene Smednos is expressed in germline and eye precursor cells during development and regeneration. *Dev Genes Evol* **217**, 403–411.

Hayashi, K., Kobayashi, T., Umino, T., Goitsuka, R., Matsui, Y. and Kitamura, D. (2002). SMAD1 signaling is critical for initial commitment of germ cell lineage from mouse epiblast. *Mechanisms of Development* **118**, 99–109.

Hoshi, M., Kobayashi, K., Arioka, S., Hase, S. and Matsumoto, M. (2003). Switch from Asexual to sexual reproduction in the planarian *Dugesia ryukyuensis*. *Integrative and Comparative Biology* **43**, 242–246.

Howe, J., Rink, J. C., Wang, B. and Griffin, A. S. (2022). Multicellularity in animals: The potential for within-organism conflict. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120457119.

Iii, J. J. C., Hou, X., Romanova, E. V., Lambrus, B. G., Miller, C. M., Saberi, A., Sweedler, J. V. and Newmark, P. A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLOS Biology* **8**, e1000509.

Issigonis, M. and Newmark, P. A. (2019). Chapter Four - From worm to germ: Germ cell development and regeneration in planarians. In *Current Topics in Developmental Biology* (ed. Lehmann, R.), pp. 127–153. Academic Press.

Issigonis, M., Redkar, A. B., Rozario, T., Khan, U. W., Mejia-Sanchez, R., Lapan, S. W., Reddien, P. W. and Newmark, P. A. (2022). A Krüppel-like factor is required for development and regeneration of germline and yolk cells from somatic stem cells in planarians. *PLoS Biol* **20**, e3001472.

Johnson, A. D., Crother, B., White, M. E., Patient, R., Bachvarova, R. F., Drum, M. and Masi, T. (2003). Regulative germ cell specification in axolotl embryos: a primitive trait conserved in the mammalian lineage. *Phil. Trans. R. Soc. Lond. B* **358**, 1371–1379.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Kershenbaum, S., Jara-Espejo, M., Griffin, A. S. and Aboobaker, A. (2024). Gene conversion limits the cost of asexuality in somatically immortal worms. 2023.03.20.533553.

Khan, U. W. and Newmark, P. A. (2022). Somatic regulation of female germ cell regeneration and development in planarians. *Cell Reports* **38**, 110525.

Lawson, K. A., Dunn, N. R., Roelen, B. A. J., Zeinstra, L. M., Davis, A. M., Wright, C. V. E., Korving, J. P. W. F. M. and Hogan, B. L. M. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424–436.

Liu, S.-Y. and Rink, J. C. (2018). Total RNA Isolation from Planarian Tissues. In *Planarian Regeneration: Methods and Protocols* (ed. Rink, J. C.), pp. 259–265. New York, NY: Springer.

Lopes, S. M. C. de S., Roelen, B. A. J., Monteiro, R. M., Emmens, R., Lin, H. Y., Li, E., Lawson, K. A. and Mummery, C. L. (2004). BMP signaling mediated by ALK2 in the visceral endoderm is necessary for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **18**, 1838–1849.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550.

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Loo, P. V., Davies, H., Stratton, M. R. and Campbell, P. J. (2017). Universal Patterns of Selection in Cancer and Somatic Tissues. *Cell* **171**, 1029-1041.e21.

Mio, C., Baldan, F. and Damante, G. (2022). NK2 homeobox gene cluster: Functions and roles in human diseases. *Genes Dis* **10**, 2038–2048.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., et al. (2020). Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412–D419.

Morgan, T. H. (1901). Growth and regeneration in *Planaria lugubris*. *Archiv für Entwicklungsmechanik der Organismen* **13**, 179–212.

Neiro, J., Sridhar, D., Dattani, A. and Aboobaker, A. (2022). Identification of enhancer-like elements defines regulatory networks active in planarian adult stem cells. 2022.02.03.479047.

Newmark, P. A., Wang, Y. and Chong, T. (2008). Germ cell specification and regeneration in planarians. *Cold Spring Harb Symp Quant Biol* **73**, 573–581.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C. (2017). Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat Methods* **14**, 417–419.

Rahnenfuhrer, A. A. (2024). TopGO: enrichment analysis for gene ontology.

Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. and Alvarado, A. S. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327–1330.

Robb, S. M. C., Gotting, K., Ross, E. and Sánchez Alvarado, A. (2015). SmedGD 2.0: The *Schmidtea mediterranea* genome database. *Genesis* **53**, 535–546.

Rosenthal, R. and Fromm, M. (2011). Endothelin antagonism as an active principle for glaucoma therapy. *Br J Pharmacol* **162**, 806–816.

Rouhana, L., Weiss, J. A., Forsthoefel, D. J., Lee, H., King, R. S., Inoue, T., Shibata, N., Agata, K. and Newmark, P. A. (2013). RNA interference by feeding in vitro synthesized double-stranded RNA to planarians: methodology and dynamics. *Dev Dyn* **242**, 718–730.

Rozanski, A., Moon, H., Brandl, H., Martín-Durán, J. M., Grohme, M. A., Hüttner, K., Bartscherer, K., Henry, I. and Rink, J. C. (2019). PlanMine 3.0—improvements to a mineable resource of flatworm biology and biodiversity. *Nucleic Acids Research* **47**, D812–D820.

Salvetti, A., Rossi, L., Lena, A., Batistoni, R., Deri, P., Rainaldi, G., Locci, M. T., Evangelista, M. and Gremigni, V. (2005). DjPum, a homologue of *Drosophila* Pumilio, is essential to planarian stem cell maintenance. *Development* **132**, 1863–1874.

Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S. and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed by

the expression of nanos-related gene in planarians. *Development, Growth & Differentiation* **48**, 615–628.

Sayols, S. (2023). rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. *microPublication Biology*.

Seervai, R. N. H. and Wessel, G. M. (2013). Lessons for inductive germline determination. *Mol Reprod Dev* **80**, 590–609.

Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K. and Agata, K. (1999). Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. *Developmental Biology* **206**, 73–87.

Sluys, R., Kawakatsu, M. and Ponce de León, R. (2005). Morphological stasis in an old and widespread group of species: Contribution to the taxonomy and biogeography of the genus *Girardia* (Platyhelminthes, *Tricladida*, *Paludicola*). *Studies on Neotropical Fauna and Environment* **40**, 155–180.

Solana, J., Lasko, P. and Romero, R. (2009). *Spoltud-1* is a chromatoid body component required for planarian long-term stem cell self-renewal. *Developmental Biology* **328**, 410–421.

Stocchino, G. A., Sluys, R., Harrath, A. H., Mansour, L. and Manconi, R. (2019). The invasive alien freshwater flatworm *Girardia tigrina* (Girard, 1850) (Platyhelminthes, *Tricladida*) in Western Europe: new insights into its morphology, karyology and reproductive biology.

Tamura, K., Stecher, G. and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol Biol Evol* **38**, 3022–3027.

Tremblay, K. D., Dunn, N. R. and Robertson, E. J. (2001). Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. *Development* **128**, 3609–3621.

Wang, Y., Zayas, R. M., Guo, T. and Newmark, P. A. (2007). Nanos function is essential for development and regeneration of planarian germ cells. *Proc Natl Acad Sci U S A* **104**, 5901–5906.

Wang, Y., Stary, J. M., Wilhelm, J. E. and Newmark, P. A. (2010). A functional genomic screen in planarians identifies novel regulators of germ cell development. *Genes Dev* **24**, 2081–2092.

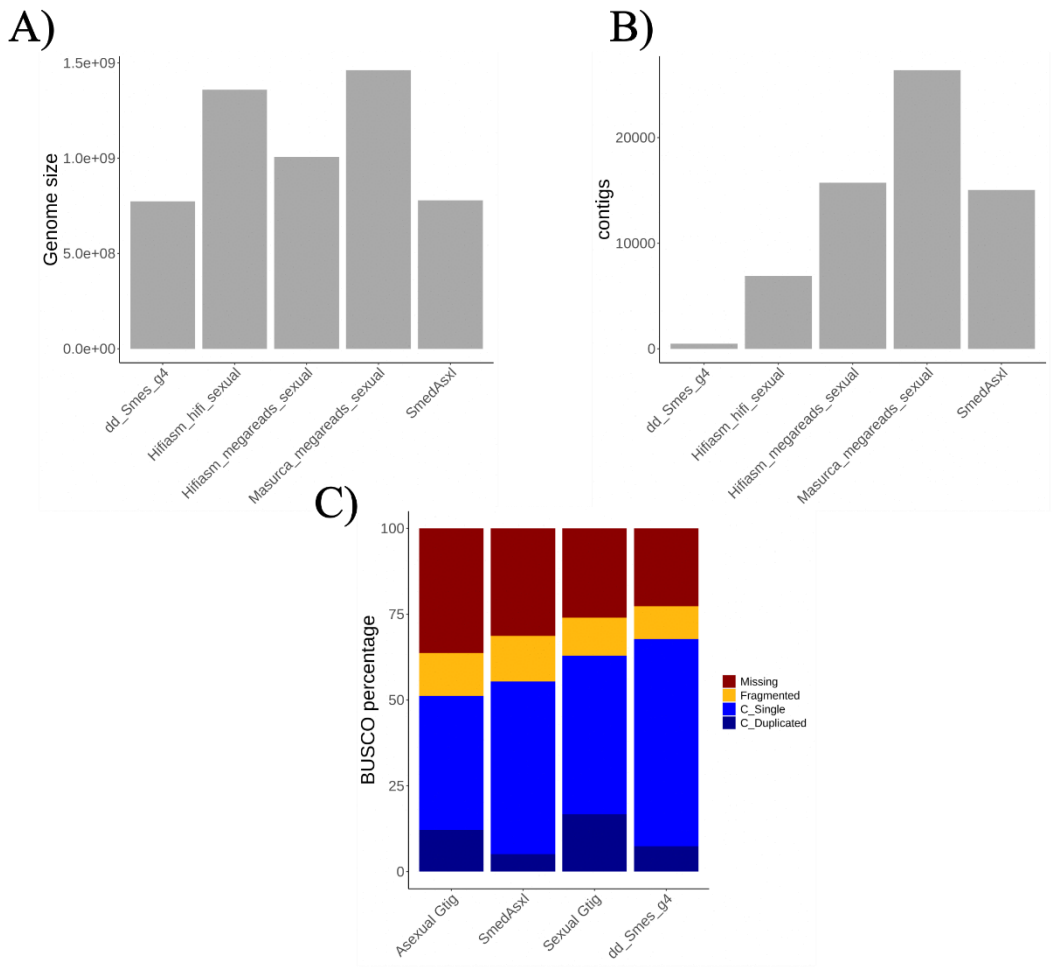
Wei, Q., Miskimins, W. K. and Miskimins, R. (2005). Stage-specific expression of myelin basic protein in oligodendrocytes involves Nkx2.2-mediated repression that is relieved by the Sp1 transcription factor. *Journal of Biological Chemistry* **280**, 16284–16294.

Wheeler, N. J., Agbedanu, P. N., Kimber, M. J., Ribeiro, P., Day, T. A. and Zamanian, M. (2015). Functional analysis of *Girardia tigrina* transcriptome seeds pipeline for anthelmintic target discovery. *Parasit Vectors* **8**, 34.

Wickham, H. (2009). *ggplot2: Elegant graphics for data analysis*. New York, NY: Springer.

Zayas, R. M., Hernández, A., Habermann, B., Wang, Y., Stary, J. M. and Newmark, P. A. (2005). The planarian *Schmidtea mediterranea* as a model for epigenetic germ cell

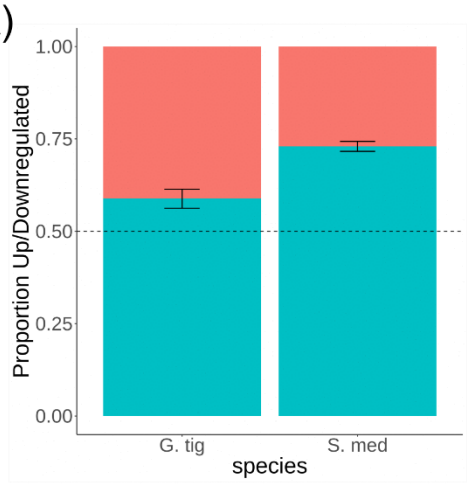
specification: Analysis of ESTs from the hermaphroditic strain. *Proc Natl Acad Sci U S A* **102**, 18491–18496.



Supplementary figure 1 – Draft assembly of sexual and asexual *G. tigrina* genomes. A) Genome size of sexual *G. tigrina* assembly using different assembly methods and comparing to published *S. mediterranea* genomes (sexual = dd_Smes_g4 (Grohme et al., 2018b); asexual SmedAsxl (Robb et al., 2015)). Using Hifiasm on PacBio reads assembles a 1,359,995,635 bp long sexual *G. tigrina* genome. B) Comparing contig number. Hifiasm

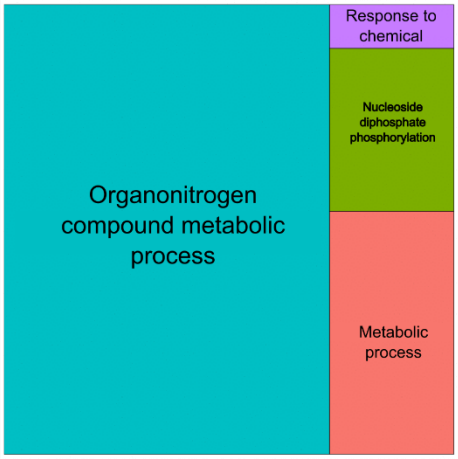
assembles 6,896 contigs for the sexual *G. tigrina* genome. Given the balance of contig number and genome length, we assessed BUSCO completeness on genomes assembled with Hifiasm. C) comparing BUSCO completeness in sexual and asexual strains of *G. tigrina* and *S. mediterranea*.

A)

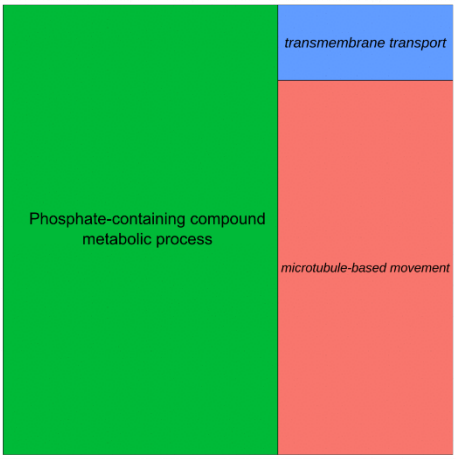


B)

G. tigrina - Up sexual

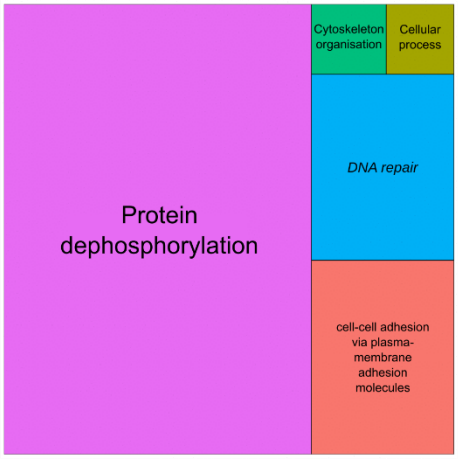


S. mediterranea - Up sexual

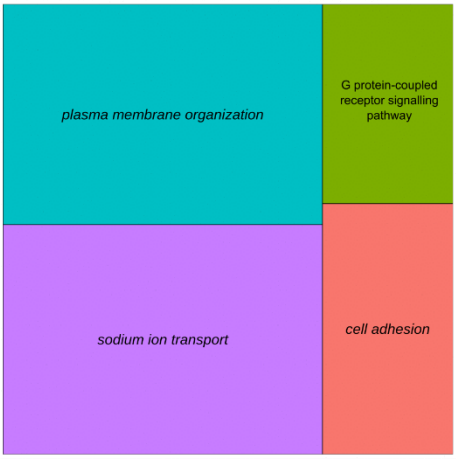


C)

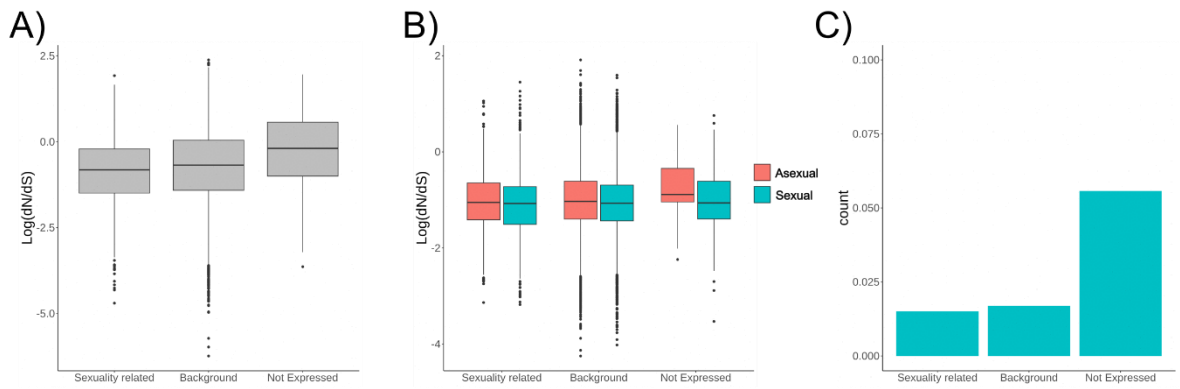
G. tigrina - Up asexual



S. mediterranea - Up asexual

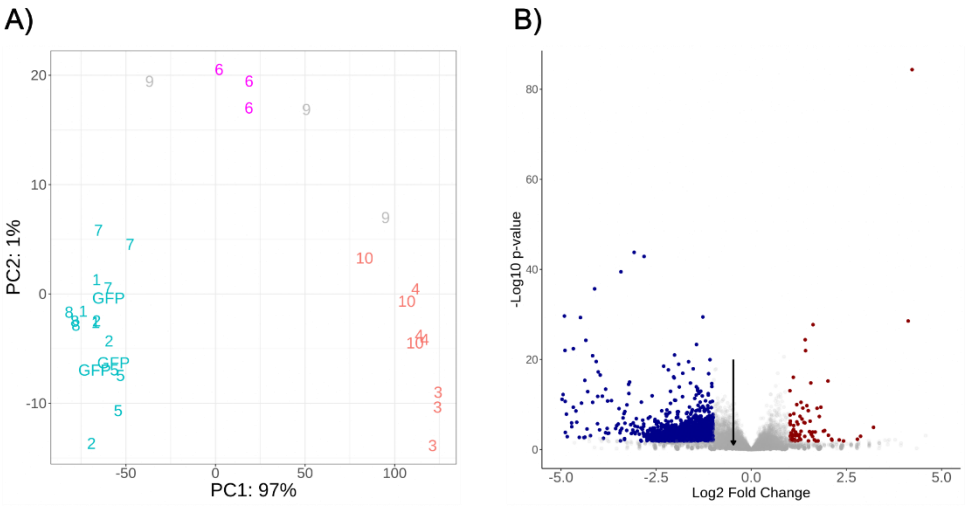


Supplementary figure 2 – Broad patterns of gene expression differences between sexual and sexual strains. A) In both *G. tigrina* and *S. mediterranea* there are more genes that are upregulated in the sexuals compared to the asexuals, and in *S. mediterranea* this is more significant (Binomial test; $N_{Gtig} = 1{,}450$, $N_{Upregulated} = 853$, $P < 0.001$; $N_{Smed} = 4{,}209$, $N_{Upregulated} = 3{,}072$, $P < 0.001$). B) GO terms associated with phosphorylation and carbohydrate metabolic processes are enriched in genes upregulated in sexual strains. C) Cell adhesion related GO terms are enriched in the asexual strains.

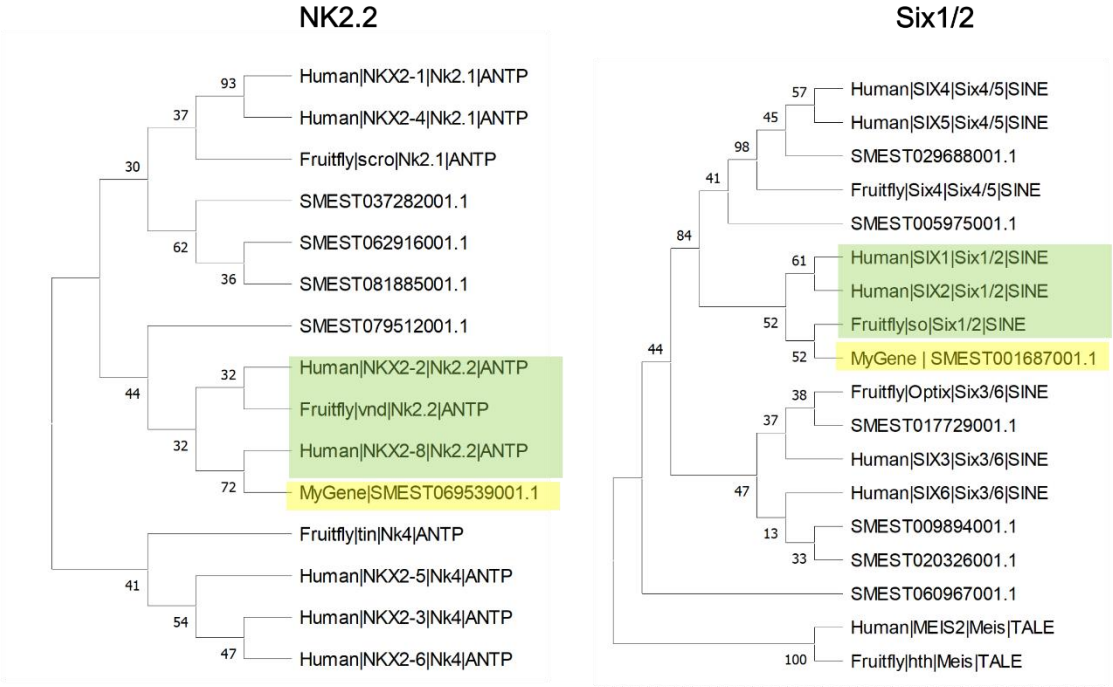


Supplementary figure 3 – Genes that are not expressed in asexual strains are degraded in *S. mediterranea*. A) Genes that are not expressed in asexual *S. mediterranea* have higher dN/dS ratios than genes that are not associated with sexuality (Wilcoxon rank sum; $N_{Not\text{-}Expressed} = 571$, $N_{Background} = 17196$, $Mean_{Not\text{-}Expressed} = 1.03$, $Mean_{Background} = 0.69$, $P < 0.001$, P.adjustment = Bonferroni). B) Sexual strains do not show a similar degradation in genes that are not expressed in asexuals (Wilcoxon rank sum; $N_{asex} = 43$, $N_{sex} = 127$, $Mean_{asex} = 0.52$, $Mean_{sex} = 0.47$, $P < 0.05$, P.adjustment = Bonferroni). For this comparison we used variants isolated from RNAseq data (see methods). C) Genes that are not expressed in asexual S. *mediterranea* contain a higher percentage of exons with nonsense mutations compared to genes that are not differentially expressed, with 5.6% and 1.7% respectively (Chi-sqr; $N_{Not\text{-}Expressed} = 1{,}184$, $N_{Background} = 101{,}697$, $P < 0.001$). However, this is not the case in genes that

are downregulated in asexual *S. mediterranea*, of which 1.5% of exons contain nonsense mutations.



Supplementary figure 4 – RNAseq of candidate transcription factors knockdown leads to three distinct clusters. A) PCA of all knockdowns (without wildtype strains). B) Gene #9 led did not cluster with any of the groups, and was not significantly knocked down in its own knockdown experiment. Animals treated with SMEST042630001.1 (gene #9) dsRNA, would stop eating as regularly, climb out of their petri dish and dry out. Thus, we believe gene #9 has an important role in behaviour, but more detailed experiments are needed to understand its specific roles.



Supplementary figure 5 – Gene trees for candidate transcription factors with newly described roles in male germline development. NK2.2 corresponds to gene #10, and Six1/2 corresponds to gene #6.

Discussion

Summary

In this thesis, I assess the evolutionary and developmental implications of asexuality by investigating asexual planarians that reproduce solely by fission.

In **chapter 1**, I used published RNAseq data from laboratories across the world to understand how asexual planarians persist despite the potential for mutation accumulation due to Muller's ratchet. I found that asexual planarians show relatively limited signatures of Muller's ratchet, with no evidence of excess mutation accumulation in highly conserved genes BUSCO genes, and that they show strong signatures of gene conversion, a process that has been hypothesised to help asexual lineages persist in the absence of sexual recombination.

In **chapter 2**, I assessed the levels of diversity present both between and within asexual planarians to better understand how clonal worms generate diversity in their population and how they avoid the accumulation of somatic variation in the absence of a zygotic bottleneck. I found previously unappreciated diversity in asexual planarians and described two genetically distinct strains that show different rates of reproduction and mortality. I then found that the strain with lower fitness also contained higher intra-organismal diversity, consistent with the hypothesis that increased somatic diversity could be linked to selfish cell adaptations. Finally, I assessed the mutation rate in the soma which ultimately drives variation both within and between worms and found that the per generation per bp mutation rate is orders of magnitude higher than in species that reproduce *via* a typical single cell bottleneck.

In **chapter 3**, I shift my focus to understanding the developmental implications of asexuality and what this can tell us about the genes important for sexual development. I compare gene expression in sexual and asexual planarians from two different species and find ten transcription factors that differ in their expression between the strains in both species. By knocking these genes down, I find two genes with previously undescribed roles in male germline development in planarians. Ultimately, this shows the utility of studying asexual species both for evolutionary and developmental questions.

Overall, I propose methods by which asexual planarians my overcome the challenges of asexuality, provide the first evidence for a cost to the lack of a zygotic bottleneck in asexual planarians, and find genes that consistently differ between sexual and asexual strains of different planarian species.

Implications, limitations, and suggested future work.

In **chapter 1**, I propose that gene conversion in the neoblasts of asexual planarians may help purge deleterious mutations from the population. Since other long-term asexuals have been shown to exhibit signatures of gene conversion (Flot et al., 2013; Schön et al., 2009), this may suggest that it is a general method that is used by asexual lineages to recombine deleterious mutations out of the genome.

However, we do not explicitly test whether gene conversion occurs at a high enough rate to limit mutation accumulation in asexual planarians. There is relatively little work done on the rates of gene conversion per site, but recent modelling suggests that a high, yet conceivable, rate of gene conversion ($\sim 10^{-4}$ per site per generation) can purge deleterious mutations from asexual genomes (Kopčak and Hartfield, 2024). While this is a high rate of gene conversion, some have suggested that rice may have a per site per generation gene conversion rate as high as 10^{-2} (Jia et al., 2021; Kopčak and Hartfield, 2024). Additionally, planarians can transmit

thousands of pluripotent stem cells to the next generation, increasing the probability of a gene conversion event removing a deleterious mutation in at least one stem cell in the next generation. Future work including both simulations and empirical measurements of gene conversion rates will be useful to confirm whether it is likely to be strong enough to limit mutation accumulation in asexual planarians. Overall, I propose that the high number of stem cells transmitted to the next generation may help increase the probability of beneficial gene conversion events, making gene conversion particularly useful to asexual species that reproduce by fission.

In **chapter 2**, I find that there are two distinct asexual planarian strains within the well-studied CIW4 population (Alvarado, 2006; Alvarado et al., 2002). Since the strain with lower fitness also shows higher intra-organismal diversity, I propose that high intra-organismal diversity may increase the probability of selfish cell adaptations evolving and driving reduced fitness. This correlation between the strain fitness and somatic diversity is an important finding since it is the first evidence for a cost of the lack of a single cell bottleneck in planarians.

However, cell competition assays are necessary to assess whether the reduced fitness in the FS strain is the result of selfish cell evolution or whether the FS strain is simply less fit overall. If the reduced fitness of FS individuals is truly due to selfish cell adaptations, neoblasts from the FS strain should outcompete neoblasts from the PH strain. Given that we provided an extensive list of strain specific variants, it should be possible to run competition experiments between neoblasts of these strains. I would suggest using multiple methods such as; A) Plug assays, where a tissue plug from one strain is transplanted into the body of the other strain (Peiris et al., 2016). qPCRs can then be used to track the changes in relative abundance of each strain over time; B) Rescue experiments with mixed populations of FS and PH cells. In this case a sexual planarian would be lethally irradiated and rescued by injecting

cells from both FS and PH at various ratios (Wagner et al., 2011). Abundance of progeny from each strain can then be assessed using qPCR on some of the distinguishing variants we provided. Overall, this will tell us whether the reduced fitness in FS is likely to be due to selfish cell adaptations that make them more competitive.

Additionally, the presence of two genetically distinct clonal strains requires further investigation and may have significant implications for our understanding of asexual planarian biology and genome evolution. We report long runs of homozygosity in the FS strain that are absent from the PH strain, which raises the possibility that some form of inbreeding took place in the FS strain (Ceballos et al., 2018). If true, this would suggest that asexual *S. mediterranea* are not obligately asexual. However, asexual processes could potentially lead to similar signatures (Engelstädter, 2017; LaFave and Sekelsky, 2009). Thus, more research is needed to establish the origins of these two strains. In particular, I would suggest that the date of divergence between the two strains be estimated to see whether they split after the CIW4 strain was established. Additionally, I would suggest that wild populations of asexual *S. mediterranea* be sequenced to assess whether there is any allele sharing in wild populations. It is worth noting that some wild populations of asexual *S. mediterranea* from the Balearic islands have been shown to have homozygous N13 (netrin receptor-like protein) genes while other populations from Barcelona have heterozygous N13 genes, suggesting that these two wild populations may represent the FS and PH strains respectively, but more sequencing will be necessary to assess whether this is the case (Lázaro et al., 2011a).

In **chapter 3**, I compare the sexual and asexual strains of *S. mediterranea* and *G. tigrina* and identify two genes with novel roles in male germline development in planarians, Six1/2 and NK2.2. Both these homeobox genes are expressed in the male germline of other species including mammals, with Six1/2 having known functional roles in the germline of both

Drosophila and mouse (Fabrizio et al., 2003; Fujimoto et al., 2013). Useful future experiments would include in situs for Six1/2, NK2.2, and nanos. In situs for Six1/2 and NK2.2 will help us understand where they are expressed and whether they function intrinsically or extrinsically to regulate the germline. In situs for nanos will help us understand whether these genes impact germline development as early as the primordial germline cell, and whether they also impact female germline development. Taken together, our work identifies genes with novel functions in inductive male germline development, providing the both the first evidence for a role of Six1/2 in spiralian germline development and the first evidence for a functional role of NK2.2 in male germline development.

Concluding remarks

In this thesis I assess some of the evolutionary implications of asexual reproduction by fission by investigating developmental and genomic changes in asexual planarians. While each chapter has its own conclusions and summaries, I want to conclude this thesis by giving a brief explanation of how I think asexual planarians persist despite their asexual nature, and whether the lack of a single cell bottleneck results in the accumulation of somatic diversity.

How do asexual planarians persist in the absence of sex?

Simply put, my research suggests that gene conversion may help purge deleterious mutations from the population, allowing them to limit the impact of Muller's ratchet. Additionally, these asexual planarians have high per generation mutation rates compared to closely related sexual planarians, which is likely due to the abundance of heritable pluripotent stem cells. This high mutation rate may help generate diversity in the absence of sex. Thus, I conclude that a combination of gene conversion and the abundance of heritable stem cells helps asexual planarians persist over time despite their asexual nature.

Does the lack of a single cell bottleneck lead to the accumulation of somatic diversity?

Some asexual planarians manage to limit the accumulation of somatic diversity despite the lack of a single cell bottleneck, suggesting that a single cell bottleneck is not strictly necessary for limiting the accumulation of somatic diversity over time. This begs the question of why zygotic single cell bottlenecks evolved convergently in all transitions to obligate multicellularity. I propose that single cell bottlenecks play a more critical role in the context of sexual reproduction, as the high diversity among cells contributing to the next generation increases the potential for genetic conflict and the evolution of selfish cell adaptations. However, limiting somatic diversity remains an important process also in clonal species, and the presence of a strain with both reduced fitness and high somatic diversity raises the possibility that selfish cell evolution can occur when somatic diversity is not limited. But more functional work is needed to assert that selfish cell evolution is driving the reduction of fitness in strains with high somatic diversity. Overall, the data suggests that the lack of a single cell bottleneck alone is not sufficient to drive the accumulation of somatic diversity in asexual planarians. However, there is likely variation in the ability to limit intra-organismal diversity between asexual planarian populations.

Bibliography

Adler, C. E. and Alvarado, A. S. (2018). Systemic RNA interference in planarians by feeding of dsRNA containing bacteria. In *Planarian Regeneration: Methods and Protocols* (ed. Rink, J. C.), pp. 445–454. New York, NY: Springer.

Afgan, E., Nekrutenko, A., Grüning, B. A., Blankenberg, D., Goecks, J., Schatz, M. C., Ostrovsky, A. E., Mahmoud, A., Lonie, A. J., Syme, A., et al. (2022). The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2022 update. *Nucleic Acids Research* **50**, W345–W351.

Aktipis, C. A., Boddy, A. M., Jansen, G., Hibner, U., Hochberg, M. E., Maley, C. C. and Wilkinson, G. S. (2015). Cancer across the tree of life: cooperation and cheating in

multicellularity. *Philosophical Transactions of the Royal Society B: Biological Sciences* **370**, 20140219.

Alvarado, A. S. (2006). Planarian regeneration: Its end is its beginning. *Cell* **124**, 241–245.

Alvarado, A. S., Newmark, P. A., Robb, S. M. C. and Juste, R. (2002). The *Schmidtea mediterranea* database as a molecular resource for studying platyhelminthes, stem cells and regeneration. *Development* **129**, 5659–5665.

Arnold, C. P., Lozano, A. M., Mann, F. G., Nowotarski, S. H., Haug, J. O., Lange, J. J., Seidel, C. W. and Alvarado, A. S. (2021). Hox genes regulate asexual reproductive behavior and tissue segmentation in adult animals. *Nat Commun* **12**, 6706.

Barraclough, T. G., Fontaneto, D., Ricci, C. and Herniou, E. A. (2007). Evidence for inefficient selection against deleterious mutations in cytochrome oxidase I of asexual bdelloid rotifers. *Molecular Biology and Evolution* **24**, 1952–1962.

Bast, J., Parker, D. J., Dumas, Z., Jalvingh, K. M., Tran Van, P., Jaron, K. S., Figuet, E., Brandt, A., Galtier, N. and Schwander, T. (2018). Consequences of asexuality in natural populations: Insights from stick insects. *Molecular Biology and Evolution* **35**, 1668–1677.

Berriman, M., Haas, B. J., LoVerde, P. T., Wilson, R. A., Dillon, G. P., Cerqueira, G. C., Mashiyama, S. T., Al-Lazikani, B., Andrade, L. F., Ashton, P. D., et al. (2009). The genome of the blood fluke *Schistosoma mansoni*. *Nature* **460**, 352–358.

Birdsell, J. A. (2002). Integrating genomics, bioinformatics, and classical genetics to study the effects of recombination on genome evolution. *Molecular Biology and Evolution* **19**, 1181–1197.

Birky, C. W., Wolf, C., Maughan, H., Herbertson, L. and Henry, E. (2005). Speciation and selection without sex. *Hydrobiologia* **546**, 29–45.

Bolger, A. M., Lohse, M. and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120.

Bosch, T. C. G., Anton-Erxleben, F., Hemmrich, G. and Khalturin, K. (2010). The Hydra polyp: Nothing but an active stem cell community. *Development, Growth & Differentiation* **52**, 15–25.

Boyer, L., Jabbour-Zahab, R., Mosna, M., Haag, C. R. and Lenormand, T. (2021). Not so clonal asexuals: Unraveling the secret sex life of *Artemia parthenogenetica*. *Evol Lett* **5**, 164–174.

Brandt, A., Schaefer, I., Glanz, J., Schwander, T., Maraun, M., Scheu, S. and Bast, J. (2017). Effective purifying selection in ancient asexual oribatid mites. *Nat Commun* **8**, 873.

Brandt, A., Tran Van, P., Bluhm, C., Anselmetti, Y., Dumas, Z., Figuet, E., François, C. M., Galtier, N., Heimburger, B., Jaron, K. S., et al. (2021). Haplotype divergence supports long-term asexuality in the oribatid mite *Oppiella nova*. *Proceedings of the National Academy of Sciences* **118**, e2101485118.

Buss, L. W. (1982). Somatic cell parasitism and the evolution of somatic tissue compatibility. *Proceedings of the National Academy of Sciences* **79**, 5337–5341.

Butlin, R., Schön, I. and Martens, K. (1998). Asexual reproduction in nonmarine ostracods. *Heredity* **81**, 473–480.

Carter, J. A., Lind, C. H., Truong, M. P. and Collins, E.-M. S. (2015). To each his own. *Journal of Statistical Physics* **161**, 250–272.

Ceballos, F. C., Joshi, P. K., Clark, D. W., Ramsay, M. and Wilson, J. F. (2018). Runs of homozygosity: windows into population history and trait architecture. *Nat Rev Genet* **19**, 220–234.

Chang, H. and Matzuk, M. M. (2001). Smad5 is required for mouse primordial germ cell development. *Mech Dev* **104**, 61–67.

Charlesworth, B. and Barton, N. H. (1996). Recombination load associated with selection for increased recombination. *Genetics Research* **67**, 27–41.

Chen, J.-M., Cooper, D. N., Chuzhanova, N., Férec, C. and Patrinos, G. P. (2007). Gene conversion: mechanisms, evolution and human disease. *genetics* 14.

Chen, S., Zhou, Y., Chen, Y. and Gu, J. (2018). fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* **34**, i884–i890.

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H. and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat Methods* **18**, 170–175.

Chong, T., Collins, J. J., Brubacher, J. L., Zarkower, D. and Newmark, P. A. (2013). A sex-specific transcription factor controls male identity in a simultaneous hermaphrodite. *Nat Commun* **4**, 1814.

Chu, G. C., Dunn, N. R., Anderson, D. C., Oxburgh, L. and Robertson, E. J. (2004). Differential requirements for Smad4 in TGFβ-dependent patterning of the early mouse embryo. *Development* **131**, 3501–3512.

Colegrave, N. (2012). The evolutionary success of sex. *EMBO Rep* **13**, 774–778.

Comeron, J. M., Williford, A. and Kliman, R. M. (2008). The Hill–Robertson effect: evolutionary consequences of weak selection and linkage in finite populations. *Heredity* **100**, 19–31.

Cowden, J. and Levine, M. (2003). Ventral dominance governs sequential patterns of gene expression across the dorsal–ventral axis of the neuroectoderm in the *Drosophila* embryo. *Developmental Biology* **262**, 335–349.

Currie, K. W. and Pearson, B. J. (2013). Transcription factors lhx1/5-1 and pitx are required for the maintenance and regeneration of serotonergic neurons in planarians. *Development* **140**, 3577–3588.

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, M. A., Handsaker, R. E., Lunter, G., Marth, G. T., Sherry, S. T., et al. (2011). The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158.

Davies, E. L., Lei, K., Seidel, C. W., Kroesen, A. E., McKinney, S. A., Guo, L., Robb, S. M., Ross, E. J., Gotting, K. and Alvarado, A. S. (2017). Embryonic origin of adult stem cells required for tissue homeostasis and regeneration. *eLife* **6**, e21052.

de Visser, J. A. G. M. and Elena, S. F. (2007). The evolution of sex: empirical insights into the roles of epistasis and drift. *Nat Rev Genet* **8**, 139–149.

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M. and Gingeras, T. R. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21.

Domazet-Lošo, T., Klimovich, A., Anokhin, B., Anton-Erxleben, F., Hamm, M. J., Lange, C. and Bosch, T. C. G. (2014). Naturally occurring tumours in the basal metazoan Hydra. *Nat Commun* **5**, 4222.

Donoughe, S., Nakamura, T., Ewen-Campen, B., Green, D. A., Henderson, L. and Extavour, C. G. (2014). BMP signaling is required for the generation of primordial germ cells in an insect. *Proceedings of the National Academy of Sciences* **111**, 4133–4138.

Duncan, E. M., Chitsazan, A. D., Seidel, C. W. and Alvarado, A. S. (2015). Set1 and MLL1/2 target distinct sets of functionally different genomic loci in vivo. *Cell Rep* **13**, 2741–2755.

Dunkel, J., Talbot, J. and Schötz, E.-M. (2011). Memory and obesity affect the population dynamics of asexual freshwater planarians. *Physical biology* **8**, 026003.

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* **32**, 1792–1797.

Edwards, P. B., Wanjura, W. J., Brown, W. V. and Dearn, J. M. (1990). Mosaic resistance in plants. *Nature* **347**, 434–434.

Emms, D. M. and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. *Genome Biology* **20**, 238.

Engelstädter, J. (2017). Asexual but Not Clonal: Evolutionary Processes in Automictic Populations. *Genetics* **206**, 993–1009.

Ewen-Campen, B., Schwager, E. E. and Extavour, C. G. M. (2010). The molecular machinery of germ line specification. *Molecular Reproduction and Development* **77**, 3–18.

Extavour, C. G. and Akam, M. (2003). Mechanisms of germ cell specification across the metazoans: epigenesis and preformation. *Development* **130**, 5869–5884.

Fabrizio, J. J., Boyle, M. and DiNardo, S. (2003). A somatic role for *eyes absent (eya)* and *sine oculis (so)* in *drosophila* spermatocyte development. *Developmental Biology* **258**, 117–128.

Felsenstein, J. (1974). The evolutionary advantage of recombination. *Genetics* **78**, 737–756.

Fields, C. and Levin, M. (2018). Are planaria individuals? What regenerative biology is telling us about the nature of multicellularity. *Evol Biol* **45**, 237–247.

Fisher, R. M., Cornwallis, C. K. and West, S. A. (2013). Group formation, relatedness, and the evolution of multicellularity. *Current Biology* **23**, 1120–1125.

Flot, J.-F., Hespeels, B., Li, X., Noel, B., Arkhipova, I., Danchin, E. G. J., Hejnol, A., Henrissat, B., Koszul, R., Aury, J.-M., et al. (2013). Genomic evidence for ameiotic evolution in the bdelloid rotifer *Adineta vaga*. *Nature* **500**, 453–457.

Forsthoefel, D. J., Ross, K. G., Newmark, P. A. and Zayas, R. M. (2018). Fixation, processing, and immunofluorescent labeling of whole mount planarians. *Methods Mol Biol* **1774**, 353–366.

Freitas, S., Parker, D. J., Labédan, M., Dumas, Z. and Schwander, T. (2023). Evidence for cryptic sex in parthenogenetic stick insects of the genus *Timema*. *Proceedings of the Royal Society B: Biological Sciences* **290**, 20230404.

Fujimoto, Y., Tanaka, S. S., Yamaguchi, Y. L., Kobayashi, H., Kuroki, S., Tachibana, M., Shinomura, M., Kanai, Y., Morohashi, K., Kawakami, K., et al. (2013). Homeoproteins Six1 and Six4 Regulate Male Sex Determination and Mouse Gonadal Development. *Developmental Cell* **26**, 416–430.

Gibson, A. K., Delph, L. F. and Lively, C. M. (2017). The two-fold cost of sex: Experimental evidence from a natural system. *Evol Lett* **1**, 6–15.

Goel, T., Ireland, D., Shetty, V., Rabeler, C., Diamond, P. H. and Collins, E.-M. S. (2021). Let it rip: the mechanics of self-bisection in asexual planarians determines their population reproductive strategies. *Phys. Biol.* **19**, 016002.

Grohme, M. A., Vila-Farré, M. and Rink, J. C. (2018a). Small- and large-scale high molecular weight genomic DNA extraction from planarians. In *Planarian regeneration: Methods and protocols* (ed. Rink, J. C.), pp. 267–275. New York, NY: Springer.

Grohme, M. A., Schloissnig, S., Rozanski, A., Pippel, M., Young, G. R., Winkler, S., Brandl, H., Henry, I., Dahl, A., Powell, S., et al. (2018b). The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature* **554**, 56–61.

Grohme, M. A., Schloissnig, S., Rozanski, A., Pippel, M., Young, G. R., Winkler, S., Brandl, H., Henry, I., Dahl, A., Powell, S., et al. (2018c). The genome of *Schmidtea mediterranea* and the evolution of core cellular mechanisms. *Nature* **554**, 56–61.

Grudniewska, M., Mouton, S., Simanov, D., Beltman, F., Grelling, M., de Mulder, K., Arindrarto, W., Weissert, P. M., van der Elst, S. and Berezikov, E. (2016). Transcriptional signatures of somatic neoblasts and germline cells in *Macrostomum lignano*. *eLife* **5**, e20607.

Grum-Grzhimaylo, A. A., Bastiaans, E., van den Heuvel, J., Berenguer Millanes, C., Debets, A. J. M. and Aanen, D. K. (2021). Somatic deficiency causes reproductive parasitism in a fungus. *Nat Commun* **12**, 783.

Guerrero-Hernández, C., Doddihal, V., Mann, F. G. and Alvarado, A. S. (2021). A powerful and versatile new fixation protocol for immunohistology and in situ hybridization that preserves delicate tissues in planaria. 2021.11.01.466817.

Guo, T., Peters, A. H. F. M. and Newmark, P. A. (2006). A bruno-like gene is required for stem cell maintenance in planarians. *Developmental Cell* **11**, 159–169.

Guo, L., Accorsi, A., He, S., Guerrero-Hernández, C., Sivagnanam, S., McKinney, S., Gibson, M. and Sánchez Alvarado, A. (2018). An adaptable chromosome preparation methodology for use in invertebrate research organisms. *BMC Biology* **16**, 1–14.

Hamilton, W. D., Axelrod, R. and Tanese, R. (1990). Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences* **87**, 3566–3573.

Handberg-Thorsager, M. and Saló, E. (2007). The planarian nanos-like gene Smednos is expressed in germline and eye precursor cells during development and regeneration. *Dev Genes Evol* **217**, 403–411.

Hartfield, M., Wright, S. I. and Agrawal, A. F. (2016). Coalescent times and patterns of genetic diversity in species with facultative sex: Effects of gene conversion, population structure, and heterogeneity. *Genetics* **202**, 297–312.

Hayashi, K., Kobayashi, T., Umino, T., Goitsuka, R., Matsui, Y. and Kitamura, D. (2002). SMAD1 signaling is critical for initial commitment of germ cell lineage from mouse epiblast. *Mechanisms of Development* **118**, 99–109.

Henry, L., Schwander, T. and Crespi, B. J. (2012). Deleterious mutation accumulation in asexual *Timema* stick insects. *Molecular Biology and Evolution* **29**, 401–408.

Hill, W. G. and Robertson, A. (1966). The effect of linkage on limits to artificial selection. *Genetics Research* **8**, 269–294.

Hoshi, M., Kobayashi, K., Arioka, S., Hase, S. and Matsumoto, M. (2003). Switch from asexual to sexual reproduction in the planarian *Dugesia ryukyuensis*. *Integrative and Comparative Biology* **43**, 242–246.

Howe, J., Rink, J. C., Wang, B. and Griffin, A. S. (2022). Multicellularity in animals: The potential for within-organism conflict. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2120457119.

Iii, J. J. C., Hou, X., Romanova, E. V., Lambrus, B. G., Miller, C. M., Saberi, A., Sweedler, J. V. and Newmark, P. A. (2010). Genome-wide analyses reveal a role for peptide hormones in planarian germline development. *PLOS Biology* **8**, e1000509.

Issigonis, M. and Newmark, P. A. (2019). Chapter Four - From worm to germ: Germ cell development and regeneration in planarians. In *Current Topics in Developmental Biology* (ed. Lehmann, R.), pp. 127–153. Academic Press.

Issigonis, M., Redkar, A. B., Rozario, T., Khan, U. W., Mejia-Sanchez, R., Lapan, S. W., Reddien, P. W. and Newmark, P. A. (2022). A Krüppel-like factor is required for development and regeneration of germline and yolk cells from somatic stem cells in planarians. *PLoS Biol* **20**, e3001472.

Jia, X., Zhang, Q., Jiang, M., Huang, J., Yu, L., Traw, M. B., Tian, D., Hurst, L. D. and Yang, S. (2021). Mitotic gene conversion can be as important as meiotic conversion in driving genetic variability in plants and other species without early germline segregation. *PLOS Biology* **19**, e3001164.

Johnson, S. G. and Howard, R. S. (2007). Contrasting patterns of synonymous and nonsynonymous sequence evolution in asexual and sexual freshwater snail lineages. *Evolution* **61**, 2728–2735.

Johnson, A. D., Crother, B., White, M. E., Patient, R., Bachvarova, R. F., Drum, M. and Masi, T. (2003). Regulative germ cell specification in axolotl embryos: a primitive trait conserved in the mammalian lineage. *Phil. Trans. R. Soc. Lond. B* **358**, 1371–1379.

Jones, P., Binns, D., Chang, H.-Y., Fraser, M., Li, W., McAnulla, C., McWilliam, H., Maslen, J., Mitchell, A., Nuka, G., et al. (2014). InterProScan 5: genome-scale protein function classification. *Bioinformatics* **30**, 1236–1240.

Judson, O. P. and Normark, B. B. (1996). Ancient asexual scandals. *Trends in Ecology & Evolution* **11**, 41–46.

Kershenbaum, S., Jara-Espejo, M., Griffin, A. S. and Aboobaker, A. (2024). Gene conversion limits the cost of asexuality in somatically immortal worms. 2023.03.20.533553.

Khan, U. W. and Newmark, P. A. (2022). Somatic regulation of female germ cell regeneration and development in planarians. *Cell Reports* **38**, 110525.

Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annu. Rev. Earth Planet. Sci.* **39**, 217–239.

Kopčak, D. and Hartfield, M. (2024). How do self-fertilising and facultative sexual populations differ in mutation accumulation? 2024.08.30.610445.

LaFave, M. C. and Sekelsky, J. (2009). Mitotic recombination: Why? When? How? Where? *PLoS Genet* **5**, e1000411.

Laine, V. N., Sackton, T. B. and Meselson, M. (2022). Genomic signature of sexual reproduction in the bdelloid rotifer *Macrotrachella quadricornifera*. *Genetics* **220**, iyab221.

Laird, D. J., De Tomaso, A. W. and Weissman, I. L. (2005). Stem Cells Are Units of Natural Selection in a Colonial Ascidian. *Cell* **123**, 1351–1360.

Lawson, K. A., Dunn, N. R., Roelen, B. A. J., Zeinstra, L. M., Davis, A. M., Wright, C. V. E., Korving, J. P. W. F. M. and Hogan, B. L. M. (1999). Bmp4 is required for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **13**, 424–436.

Lázaro, E. M., Harrath, A. H., Stocchino, G. A., Pala, M., Baguñà, J. and Riutort, M. (2011a). *Schmidtea mediterranea* phylogeography: an old species surviving on a few Mediterranean islands? *BMC Evolutionary Biology* **11**, 274.

Lehtonen, J., Jennions, M. D. and Kokko, H. (2012). The many costs of sex. *Trends in Ecology & Evolution* **27**, 172–178.

Leria, L., Vila-Farré, M., Solà, E. and Riutort, M. (2019). Outstanding intraindividual genetic diversity in fissiparous planarians (*Dugesia*, Platyhelminthes) with facultative sex. *BMC Evolutionary Biology* **19**, 130.

Li, H. (2013). Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. *arXiv*.

Liu, S.-Y. and Rink, J. C. (2018). Total RNA isolation from planarian tissues. In *planarian regeneration: Methods and protocols* (ed. Rink, J. C.), pp. 259–265. New York, NY: Springer.

Lively, C. M. and Morran, L. T. (2014). The ecology of sexual reproduction. *Journal of Evolutionary Biology* **27**, 1292–1303.

Loewe, L. and Lamatsch, D. K. (2008). Quantifying the threat of extinction from Muller's ratchet in the diploid Amazon molly (*Poecilia formosa*). *BMC Evolutionary Biology* **8**, 88.

Lopes, S. M. C. de S., Roelen, B. A. J., Monteiro, R. M., Emmens, R., Lin, H. Y., Li, E., Lawson, K. A. and Mummery, C. L. (2004). BMP signaling mediated by ALK2 in the visceral endoderm is necessary for the generation of primordial germ cells in the mouse embryo. *Genes Dev.* **18**, 1838–1849.

Love, M. I., Huber, W. and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biology* **15**, 550.

Lynch, M. and Gabriel, W. (1990). Mutation load and the survival of small populations. *Evolution* **44**, 1725–1737.

Lynch, M., Bürger, R., Butcher, D. and Gabriel, W. (1993). The mutational meltdown in asexual populations. *Journal of Heredity* **84**, 339–344.

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A. and Zdobnov, E. M. (2021). BUSCO update: Novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral Genomes. *Molecular Biology and Evolution* **38**, 4647–4654.

Mark Welch, D. B. and Meselson, M. (2000). Evidence for the evolution of bdelloid rotifers without sexual reproduction or genetic exchange. *Science* **288**, 1211–1215.

Martincorena, I., Raine, K. M., Gerstung, M., Dawson, K. J., Haase, K., Loo, P. V., Davies, H., Stratton, M. R. and Campbell, P. J. (2017). Universal patterns of selection in cancer and somatic tissues. *Cell* **171**, 1029-1041.e21.

Martínez, D. E. (1998). Mortality patterns suggest lack of senescence in hydra. *Exp Gerontol* **33**, 217–225.

Mio, C., Baldan, F. and Damante, G. (2022). NK2 homeobox gene cluster: Functions and roles in human diseases. *Genes Dis* **10**, 2038–2048.

Miralles, R., Gerrish, P. J., Moya, A. and Elena, S. F. (1999). Clonal interference and the evolution of RNA viruses. *Science* **285**, 1745–1747.

Mistry, J., Chuguransky, S., Williams, L., Qureshi, M., Salazar, G. A., Sonnhammer, E. L. L., Tosatto, S. C. E., Paladin, L., Raj, S., Richardson, L. J., et al. (2020). Pfam: The protein families database in 2021. *Nucleic Acids Res* **49**, D412–D419.

Morgan, T. H. (1901). Growth and regeneration in *Planaria lugubris*. *Archiv für Entwicklungsmechanik der Organismen* **13**, 179–212.

Muller, H. J. (1964). The relation of recombination to mutational advance. 8.

Nakagawa, H., Sekii, K., Maezawa, T., Kitamura, M., Miyashita, S., Abukawa, M., Matsumoto, M. and Kobayashi, K. (2018). A comprehensive comparison of sex-inducing activity in asexual worms of the planarian *Dugesia ryukyuensis*: the crucial sex-inducing substance appears to be present in yolk glands in Tricladida. *Zoological Lett* **4**, 14.

Neiman, M., Hehman, G., Miller, J. T., Logsdon, J. M. and Taylor, D. R. (2010). Accelerated mutation accumulation in asexual lineages of a freshwater snail. *Mol Biol Evol* **27**, 954–963.

Neiro, J., Sridhar, D., Dattani, A. and Aboobaker, A. (2022a). Identification of enhancer-like elements defines regulatory networks active in planarian adult stem cells. 2022.02.03.479047.

Neiro, J., Sridhar, D., Dattani, A. and Aboobaker, A. (2022b). Identification of putative enhancer-like elements predicts regulatory networks active in planarian adult stem cells. *eLife* **11**, e79675.

Newmark, P. A. and Sánchez Alvarado, A. (2000). Bromodeoxyuridine specifically labels the regenerative stem cells of planarians. *Developmental Biology* **220**, 142–153.

Newmark, P. A., Wang, Y. and Chong, T. (2008). Germ cell specification and regeneration in planarians. *Cold Spring Harb Symp Quant Biol* **73**, 573–581.

Ní Leathlobhair, M. and Lenski, R. E. (2022). Population genetics of clonally transmissible cancers. *Nat Ecol Evol* **6**, 1077–1089.

Nishant, K. T., Singh, N. D. and Alani, E. (2009). Genomic mutation rates: What high-throughput methods can tell us. *Bioessays* **31**, 912–920.

Nishimura, O., Hosoda, K., Kawaguchi, E., Yazawa, S., Hayashi, T., Inoue, T., Umesono, Y. and Agata, K. (2015). Unusually large number of mutations in asexually reproducing clonal planarian *Dugesia japonica*. *PLOS ONE* **10**, e0143525.

Omilian, A. R., Cristescu, M. E. A., Dudycha, J. L. and Lynch, M. (2006). Ameiotic recombination in asexual lineages of *Daphnia*. *Proceedings of the National Academy of Sciences* **103**, 18638–18643.

Paland, S. and Lynch, M. (2006). Transitions to asexuality result in excess amino acid substitutions. *Science* **311**, 990–992.

Patil, A. B. and Vijay, N. (2021). Repetitive genomic regions and the inference of demographic history. *Heredity (Edinb)* **127**, 151–166.

Patro, R., Duggal, G., Love, M. I., Irizarry, R. A. and Kingsford, C. (2017). Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. *Nat Methods* **14**, 417–419.

Peiris, T. H., Ramirez, D., Barghouth, P. G., Ofoha, U., Davidian, D., Weckerle, F. and Oviedo, N. J. (2016). Regional signals in the planarian body guide stem cell fate in the presence of genomic instability. *Development* **143**, 1697–1709.

Póti, Á., Szüts, D. and Vermezovic, J. (2024). Mutational profile of the regenerative process and de novo genome assembly of the planarian *Schmidtea polychroa*. *Nucleic Acids Research* **52**, 1779–1792.

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A. R., Bender, D., Maller, J., Sklar, P., de Bakker, P. I. W., Daly, M. J., et al. (2007). PLINK: A tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* **81**, 559–575.

Quinlan, A. R. and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842.

Quinodoz, S., Thomas, M. A., Dunkel, J. and Schötz, E. M. (2011). The more the merrier?: Entropy and statistics of asexual reproduction in freshwater planarians. *Journal of Statistical Physics* **142**, 1324–1336.

Rahnenfuhrer, A. A. (2024). TopGO: enrichment analysis for gene ontology.

Reddien, P. W., Oviedo, N. J., Jennings, J. R., Jenkin, J. C. and Alvarado, A. S. (2005). SMEDWI-2 is a PIWI-like protein that regulates planarian stem cells. *Science* **310**, 1327–1330.

Reusch, T. B. H., Baums, I. B. and Werner, B. (2021). Evolution via somatic genetic variation in modular species. *Trends in Ecology & Evolution* **36**, 1083–1092.

Rink, J. C. (2013). Stem cell systems and regeneration in planaria. *Dev Genes Evol* **223**, 67–84.

Robb, S. M. C., Gotting, K., Ross, E. and Sánchez Alvarado, A. (2015). SmedGD 2.0: The *Schmidtea mediterranea* genome database. *Genesis* **53**, 535–546.

Rouhana, L., Weiss, J. A., Forsthoefel, D. J., Lee, H., King, R. S., Inoue, T., Shibata, N., Agata, K. and Newmark, P. A. (2013). RNA interference by feeding in vitro synthesized double-stranded RNA to planarians: methodology and dynamics. *Dev Dyn* **242**, 718–730.

Rozanski, A., Moon, H., Brandl, H., Martín-Durán, J. M., Grohme, M. A., Hüttner, K., Bartscherer, K., Henry, I. and Rink, J. C. (2019). PlanMine 3.0—improvements to a mineable resource of flatworm biology and biodiversity. *Nucleic Acids Research* **47**, D812–D820.

Saberi, A., Jamal, A., Beets, I., Schoofs, L. and Newmark, P. A. (2016). GPCRs direct germline development and somatic gonad function in planarians. *PLoS Biol* **14**, e1002457.

Sahu, S., Dattani, A. and Aboobaker, A. A. (2017). Secrets from immortal worms: What can we learn about biological ageing from the planarian model system? *Seminars in Cell & Developmental Biology* **70**, 108–121.

Salvetti, A., Rossi, L., Lena, A., Batistoni, R., Deri, P., Rainaldi, G., Locci, M. T., Evangelista, M. and Gremigni, V. (2005). DjPum, a homologue of *Drosophila* Pumilio, is essential to planarian stem cell maintenance. *Development* **132**, 1863–1874.

Sato, K., Shibata, N., Orii, H., Amikura, R., Sakurai, T., Agata, K., Kobayashi, S. and Watanabe, K. (2006). Identification and origin of the germline stem cells as revealed by the expression of nanos-related gene in planarians. *Development, Growth & Differentiation* **48**, 615–628.

Sayols, S. (2023). rrvgo: a Bioconductor package for interpreting lists of Gene Ontology terms. *microPublication Biology*.

Schön, I., Martens, K., Van Doninck, K. and Butlin, R. K. (2003). Evolution in the slow lane: molecular rates of evolution in sexual and asexual ostracods (Crustacea: *Ostracoda*): Molecular evolution in sexual and asexual ostracods. *Biological Journal of the Linnean Society* **79**, 93–100.

Schön, I., Rossetti, G. and Martens, K. (2009). Darwinulid ostracods: Ancient asexual scandals or scandalous gossip? In *lost sex: The evolutionary biology of parthenogenesis* (ed. Schön, I.), Martens, K.), and Dijk, P.), pp. 217–240. Dordrecht: Springer Netherlands.

Scimone, M. L., Kravarik, K. M., Lapan, S. W. and Reddien, P. W. (2014). Neoblast specialization in regeneration of the planarian *Schmidtea mediterranea*. *Stem Cell Reports* **3**, 339–352.

Seervai, R. N. H. and Wessel, G. M. (2013). Lessons for inductive germline determination. *Mol Reprod Dev* **80**, 590–609.

Segers, H. (2007). Annotated checklist of the rotifers (Phylum *Rotifera*), with notes on nomenclature, taxonomy and distribution. *Zootaxa* **1564**, 1–104.

Shibata, N., Umesono, Y., Orii, H., Sakurai, T., Watanabe, K. and Agata, K. (1999). Expression of vasa(vas)-related genes in germline cells and totipotent somatic stem cells of planarians. *Developmental Biology* **206**, 73–87.

Simberloff, D. and Leppanen, C. (2019). Plant somatic mutations in nature conferring insect and herbicide resistance. *Pest Management Science* **75**, 14–17.

Sluys, R., Kawakatsu, M. and Ponce de León, R. (2005). Morphological stasis in an old and widespread group of species: Contribution to the taxonomy and biogeography of the genus *Girardia* (Platyhelminthes, *Tricladida*, *Paludicola*). *Studies on Neotropical Fauna and Environment* **40**, 155–180.

Solana, J., Lasko, P. and Romero, R. (2009). *Spoltud-1* is a chromatoid body component required for planarian long-term stem cell self-renewal. *Developmental Biology* **328**, 410–421.

Stammnitz, M. R., Gori, K., Kwon, Y. M., Harry, E., Martin, F. J., Billis, K., Cheng, Y., Baez-Ortega, A., Chow, W., Comte, S., et al. (2023). The evolution of two transmissible cancers in Tasmanian devils. *Science* **380**, 283–293.

Stocchino, G. A., Sluys, R., Harrath, A. H., Mansour, L. and Manconi, R. (2019). The invasive alien freshwater flatworm *Girardia tigrina* (Girard, 1850) (Platyhelminthes, *Tricladida*) in Western Europe: new insights into its morphology, karyology and reproductive biology.

Stoner, D. S., Rinkevich, B. and Weissman, I. L. (1999). Heritable germ and somatic cell lineage competitions in chimeric colonial protochordates. *Proceedings of the National Academy of Sciences* **96**, 9148–9153.

Swanstrom, J., Chen, K., Castillo, K., Barraclough, T. G. and Fontaneto, D. (2011). Testing for evidence of inefficient selection in bdelloid rotifers: do sample size and habitat differences matter? *Hydrobiologia* **662**, 19–25.

Tamura, K., Stecher, G. and Kumar, S. (2021). MEGA11: Molecular evolutionary genetics analysis version 11. *Mol Biol Evol* **38**, 3022–3027.

Tan, T. C. J., Rahman, R., Jaber-Hijazi, F., Felix, D. A., Chen, C., Louis, E. J. and Aboobaker, A. (2012). Telomere maintenance and telomerase activity are differentially regulated in asexual and sexual worms. *Proceedings of the National Academy of Sciences* **109**, 4209–4214.

Tang, C. Q., Obertegger, U., Fontaneto, D. and Barraclough, T. G. (2014). Sexual species are separated by larger genetic gaps than asexual species in rotifers. *Evolution* **68**, 2901–2916.

Thomas, M. A. and Schötz, E.-M. (2011). SAPling: a Scan-Add-Print barcoding database system to label and track asexual organisms. *The Journal of experimental biology* **214**, 3518–23.

Thomas, M. A., Quinodoz, S. and Schötz, E. M. (2012). Size matters!: Birth size and a size-independent stochastic term determine asexual reproduction dynamics in freshwater planarians. *Journal of Statistical Physics* **148**, 663–675.

Tremblay, K. D., Dunn, N. R. and Robertson, E. J. (2001). Mouse embryos lacking Smad1 signals display defects in extra-embryonic tissues and germ cell formation. *Development* **128**, 3609–3621.

Tu, K. C., Cheng, L.-C., TK Vu, H., Lange, J. J., McKinney, S. A., Seidel, C. W. and Sánchez Alvarado, A. (2015). Egr-5 is a post-mitotic regulator of planarian epidermal differentiation. *eLife* **4**, e10501.

Tucker, A. E., Ackerman, M. S., Eads, B. D., Xu, S. and Lynch, M. (2013). Population-genomic insights into the evolutionary origin and fate of obligately asexual *Daphnia pulex*. *Proc Natl Acad Sci U S A* **110**, 15740–15745.

Van der Auwera, G. A. and O'Conner, B. D. (2020). *Genomics in the cloud: Using Docker, GATK, and WDL in Terra*. 1st Editio. O'Reilly Media.

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., del Angel, G., Levy-Moonshine, A., Jordan, T., Shakir, K., Roazen, D., Thibault, J., et al. (2013). From FastQ data to high-confidence variant calls: The genome analysis toolkit vest practices pipeline. *Current Protocols in Bioinformatics* **43**, 11.10.1-11.10.33.

van Wolfswinkel, J. C., Wagner, D. E. and Reddien, P. W. (2014). Single-cell analysis reveals functionally distinct classes within the planarian stem cell compartment. *Cell Stem Cell* **15**, 326–339.

Vergara, D., Jokela, J., Lively, C. M. and Bronstein, E. J. L. (2014). Infection dynamics in coexisting sexual and asexual host populations: Support for the red queen hypothesis. *The American Naturalist* **184**, S22–S30.

Vogg, M. C., Owlarn, S., Pérez Rico, Y. A., Xie, J., Suzuki, Y., Gentile, L., Wu, W. and Bartscherer, K. (2014). Stem cell-dependent formation of a functional anterior regeneration pole in planarians requires Zic and Forkhead transcription factors. *Developmental Biology* **390**, 136–148.

Wagner, D. E., Wang, I. E. and Reddien, P. W. (2011). Clonogenic neoblasts are pluripotent adult stem cells that underlie planarian regeneration. *Science* **332**, 811–816.

Wang, Y., Zayas, R. M., Guo, T. and Newmark, P. A. (2007). Nanos function is essential for development and regeneration of planarian germ cells. *Proc Natl Acad Sci U S A* **104**, 5901–5906.

Wang, Y., Stary, J. M., Wilhelm, J. E. and Newmark, P. A. (2010). A functional genomic screen in planarians identifies novel regulators of germ cell development. *Genes Dev* **24**, 2081–2092.

Wang, I. E., Wagner, D. E. and Reddien, P. W. (2018). Chapter 20 - Irradiation and single-cell transplantation. **1774**, 18–22.

Wei, Q., Miskimins, W. K. and Miskimins, R. (2005). Stage-specific expression of myelin basic protein in oligodendrocytes involves Nkx2.2-mediated repression that is relieved by the Sp1 transcription factor. *Journal of Biological Chemistry* **280**, 16284–16294.

Welch, D. B. M. and Meselson, M. S. (2001). Rates of nucleotide substitution in sexual and anciently asexual rotifers. *Proceedings of the National Academy of Sciences* **98**, 6720–6724.

West, S. A., Fisher, R. M., Gardner, A. and Kiers, E. T. (2015). Major evolutionary transitions in individuality. *Proceedings of the National Academy of Sciences* **112**, 10112–10119.

Wheeler, N. J., Agbedanu, P. N., Kimber, M. J., Ribeiro, P., Day, T. A. and Zamanian, M. (2015). Functional analysis of *Girardia tigrina* transcriptome seeds pipeline for anthelmintic target discovery. *Parasit Vectors* **8**, 34.

Wilson, C. G., Nowell, R. W. and Barraclough, T. G. (2018). Cross-contamination explains “Inter and intraspecific horizontal genetic transfers” between asexual bdelloid rotifers. *Current Biology* **28**, 2436-2444.e14.

Wilson, C. G., Pieszko, T., Nowell, R. W. and Barraclough, T. G. (2024a). Recombination in bdelloid rotifer genomes: asexuality, transfer and stress. *Trends in Genetics* **40**, 422–436.

Yonemitsu, M. A., Giersch, R. M., Polo-Prieto, M., Hammel, M., Simon, A., Cremonte, F., Avilés, F. T., Merino-Véliz, N., Burioli, E. A., Muttray, A. F., et al. (2019). A single clonal

lineage of transmissible cancer identified in two marine mussel species in South America and Europe. *eLife* **8**, e47788.

Zayas, R. M., Hernández, A., Habermann, B., Wang, Y., Stary, J. M. and Newmark, P. A. (2005). The planarian *Schmidtea mediterranea* as a model for epigenetic germ cell specification: Analysis of ESTs from the hermaphroditic strain. *Proc Natl Acad Sci U S A* **102**, 18491–18496.

Zhu, S. J., Hallows, S. E., Currie, K. W., Xu, C. and Pearson, B. J. (2015). A mex3 homolog is required for differentiation during planarian stem cell lineage development. *eLife* **4**, e07025.