The retrotransposon R2 maintains Drosophila ribosomal DNA repeats

Jonathan O. Nelson[1], Alyssa Slicko[1,3], and Yukiko M. Yamashita[1,2,3*]

[1] Whitehead Institute for Biomedical Research,

[2]Department of Biology, MIT,

[3]Howard Hughes Medical Institute,

455 Main Street, Cambridge, MA 02142

*Corresponding Author

Ribosomal RNAs (rRNAs) account for 80-90% of all transcripts in eukaryotic cells[1]. To meet this demand, the ribosomal DNA (rDNA) gene that codes for rRNA is tandemly repeated hundreds of times, comprising rDNA loci on eukaryotic chromosomes[2–6]. This repetitiveness imposes a challenge to maintaining sufficient copy number due to spontaneous intra-chromatid recombination between repetitive units causing copy number loss[7]. The progressive shrinking of rDNA loci from generation to generation could lead to extinction of the lineage[8], yet the mechanism(s) to counteract spontaneous copy number loss remained unclear. Here, we show that the rDNA-specific retrotransposon R2 is essential for rDNA copy number (CN) maintenance in the Drosophila male germline, despite the perceived disruptive nature of transposable elements. Depletion of R2 led to defective rDNA CN maintenance in multiple contexts, causing a decline in fecundity over generations and eventual extinction of the lineage. Our data suggests that DNA double strand breaks generated by R2 is the initiating event of rDNA CN expansion, stimulating the repair processes proposed to underlie rDNA CN expansion[9]. This study reveals that retrotransposons can provide a benefit to their hosts, contrary to their reputation as genomic parasitic, which may contribute to their widespread success throughout taxa.

rDNA loci contain the hundreds of tandemly repeated copies of rRNA genes needed to support cellular viability. Ironically, this repetitive structure is susceptible to intra-chromatid recombination that causes rDNA CN loss **(Fig. 1a)**, which is a major cause of replicative senescence in budding yeast[8]. Evidence of similar rDNA CN instability has been noted in some tissues from aged dogs and humans[10,11]. Critically, age-associated rDNA CN loss also occurs in the Drosophila male germline and is inherited by the next generation[12]. The essential yet unstable nature of rDNA raises the question as to how the degeneration of rDNA loci over successive generations is prevented to avoid the extinction of the lineage. Intensive studies have revealed that sister chromatid recombination mediates rDNA CN expansion in yeast, thereby maintaining

36 rDNA repeat abundance over generations[8,13]. rDNA CN is variable between individuals of most
37 species but maintained within a consistent range throughout the population[2–6], implying that trans-
38 generational dynamic rDNA CN changes (loss and restoration) are a common feature of rDNA
39 maintenance. However, the mechanism of trans-generational rDNA CN maintenance in
40 multicellular organisms has remained a mystery. Over 50 years ago, the phenomenon of 'rDNA
41 magnification' in Drosophila was first described as the process wherein aberrant rDNA loci
42 bearing minimal rDNA repeats recover a normal rDNA CN[14]. Similar to yeast, rDNA magnification
43 was shown to require genes involved in homologous recombination-mediated repair, which can
44 create duplications of tandemly repeated elements[15]. Despite the robust rDNA CN expansion
45 activity, it remained unclear if the mechanisms of rDNA magnification served a physiological
46 function in natural populations. We recently demonstrated that offspring who inherited reduced
47 rDNA CN from old fathers could also recover normal rDNA CN in their germline[12]. This recovery
48 depends on the same set of the genes as rDNA magnification, leading us to propose that rDNA
49 magnification is a manifestation of the physiological mechanisms to maintain rDNA CN across
50 generations[16]. However, the underlying factors responsible for rDNA magnification and their
51 normal physiological functions remain poorly understood.
52
53 **R2 knockdown compromises GSC maintenance**
54
55 Metazoan rDNA genes are frequently inserted by rDNA-specific transposable elements
56 (TEs), such as the retrotransposon *R2* in Drosophila. R2 is found throughout arthropods[17] and
57 R2-like elements are widely present across taxa, including Cnidaria[18], Planaria[19], nematodes[20],
58 fish[21,22], birds, and reptiles[22]. These TEs use their sequence-specific nuclease to mobilize
59 specifically within rDNA loci, inserting in rDNA genes and disrupting their function[23] **(Extended**
60 **Data Fig. 1a)**. Surprisingly, we found that RNAi-mediated knockdown of R2 in the Drosophila
61 male germline (nos-gal4 driving R2 RNAi expression) **(Extended Data Fig. 1a-b)** resulted in

62 premature loss of germline stem cells (GSCs) during aging **(Fig. 1b-f)**. GSCs continuously
63 produce differentiating germ cells to sustain sperm production throughout adulthood and thus are
64 the source of the genome passed to the next generation[24]. Whereas newly eclosed R2 RNAi
65 males contained similar numbers of GSCs to controls, GSC number more rapidly declined during
66 aging in R2 knockdown males compared to controls **(Fig. 1b-f)**. Given that R2 is specifically
67 inserted at rDNA, we also examined the effect of R2 knockdown on rDNA stability. Using highly
68 quantitative droplet digital PCR (ddPCR), we found that RNAi-mediated knockdown of R2
69 enhanced rDNA CN loss in the testis during aging compared to controls **(Fig. 1g)**. rDNA CN
70 insufficiency is likely the primary cause of GSC loss in R2 RNAi animals, because increasing total
71 rDNA CN via introduction of a mini-chromosome harboring an rDNA locus[25] suppressed the
72 premature GSC loss caused by R2 knockdown **(Fig. 1h)**. These results revealed that R2
73 contributes to sustaining GSC population during aging by supporting rDNA CN maintenance,
74 uncovering an unanticipated benefit of the R2 retrotransposon to the host, despite the widely-held
75 view of TEs being genetic parasites.
76
77 **R2 facilitates rDNA magnification**
78
79 Why is R2 necessary for rDNA copy number maintenance? We found that R2 is required
80 for rDNA magnification, a phenomenon of rDNA CN expansion classically assayed by the
81 restoration of normal cuticular development in animals inheriting rDNA loci that had insufficient
82 rDNA CN in the previous generation **(Fig. 2a)**. The *Drosophila melanogaster* genome has two
83 rDNA loci, one on each X and Y chromosome[26]. rDNA magnification almost exclusively occurs to
84 X chromosome rDNA loci harboring the minimal viable amount of rDNA (bb^{Z9}, **Extended Data**
85 **Fig. 2a)** when combined with a Y chromosome completely lacking rDNA (bb^{Z9}/Ybb^{0}, 'magnifying
86 males' hereafter)[14] **(Extended Data Fig. 2b)**. Importantly, rDNA magnification never occurs in
87 males with a normal Y chromosome containing intact rDNA (bb^{Z9}/Y^{+}, 'non-magnifying males'

88 hereafter)[27], indicating the presence of mechanisms to monitor rDNA CN to activate expansion.
89 We found that RNAi-mediated knockdown of R2 eliminated rDNA magnification observed in
90 magnifying males from 13.73% to 0% (n = 414, p = 5.6×10^{-7}) **(Fig. 2b)**. Moreover, quantification
91 of rDNA CN by ddPCR revealed that 87.5% of bb^{Z9} chromosomes increased rDNA CN in
92 magnifying males compared to the non-magnifying males (n = 96, p = 1.8×10^{-4}), with an average
93 increase of 18.29 rDNA copies across all samples (n = 96, p = 3.1×10^{-12}) **(Fig. 2c)**. This
94 widespread rDNA CN increase indicates that rDNA magnification broadly increases rDNA CN
95 throughout the germline, despite only 13.73% of bb^{Z9} chromosomes having enough CN increase
96 to support normal cuticle development. We found this rDNA CN increase is also eliminated upon
97 R2 knockdown in magnifying males (n = 96, p = 4.7×10^{-10}) **(Fig. 2c)**. These results reveal that R2
98 is required for rDNA CN expansion during rDNA magnification. Interestingly, we found that R2
99 knockdown only disrupted rDNA magnification when the R2 RNAi construct was expressed by
100 the *nanos*-gal4 driver in early germ cells (including GSCs)[28], but not when expressed in later germ
101 cells (excluding GSCs) by the bam-gal4 driver[29] **(Fig. 2b)**. These results indicate that R2 primarily
102 functions in the earliest stages of germ cells (including GSCs) to support rDNA magnification.
103
104 We further found that R2 is sufficient for rDNA CN expansion. Ectopic expression of R2 in
105 the germline induced rDNA magnification of the bb^{Z9} locus in non-magnifying males (bb^{Z9}/Y^{+}) **(Fig.**
106 **2d-e)**. While assessment of dorsal cuticle phenotype never detects rDNA magnification from
107 control males, we found 3.3% of bb^{Z9} rDNA loci from non-magnifying males expressing a
108 transgenic copy of R2 were phenotypically magnified **(Fig. 2d,** n = 877, p = 3.2×10^{-5}).
109 Quantification of rDNA CN by ddPCR revealed that ectopic overexpression of R2 in non-
110 magnifying males (bb^{Z9}/Y^{+}) also increases the average rDNA CN at bb^{Z9} rDNA loci among all
111 offspring compared to controls **(Fig. 2e**, n = 94, p = 0.0256), revealing R2 is sufficient to induce
112 rDNA CN expansion. Critically, expression of a nuclease dead R2 transgene (NucDeadR2) in
113 non-magnifying males **(Extended Data Fig. 3a)**[30] failed to induce phenotypic detection of rDNA

114 magnification **(Fig. 2d)**, suggesting that the nuclease activity of R2 is essential for its ability to
115 induce rDNA CN expansion.
116
117 **R2 forms DSBs during rDNA magnification**
118
119 How does R2's nuclease activity contribute to rDNA magnification? In yeast, rDNA CN
120 expansion is initiated by double-stranded breaks (DSBs) at the rDNA intergenic sequence, which
121 induces sister chromatid recombination that results in rDNA gene duplication[31]. While the
122 mechanism of rDNA magnification remains unclear, all proposed models of Drosophila rDNA CN
123 expansion require an initiating DSB at the rDNA locus, the most prominent model being sister
124 chromatid recombination resulting in unequal sister chromatid recombination[32] **(Fig. 3a;**
125 **Extended Data Fig. 4a-b** for additional models**)**. Indeed, artificial introduction of DSBs at rDNA
126 loci by I-CreI endonuclease expression has been reported to induce rDNA magnification[33], but
127 the endogenous factor that induces rDNA magnification remained unclear. R2 is capable of
128 creating DSBs through sequential nicking of opposite DNA strands during retrotransposition[34]. It
129 has been speculated that DSBs created during R2 retrotranspostion may be an initiating event of
130 rDNA magnification[9], although this possibility has yet to be empirically tested. We first confirmed
131 that ectopic overexpression of R2, but not NucDeadR2, indeed induces chromosomal breaks at
132 rDNA loci identified by chromosome spreads **(Extended Data Fig. 3b-d)**. R2 overexpression (but
133 not NucDeadR2) in the germline also increased the frequency of GSCs with DSBs, identified by
134 γH2Av expression **(Extended Data Fig. 3e)**. Next, we found rDNA magnification is associated
135 with an elevation in DSBs in GSCs: the frequency of γH2Av-positive GSCs is increased in
136 magnifying males (bb^{Z9}/Ybb^{0}) compared to non-magnifying males (bb^{Z9}/Y^{+}) (**Fig. 3b-c, e**; n = 519,
137 $p = 8.8x10^{-4}$). Strikingly, we observed that knockdown of R2 in magnifying males reduced the
138 frequency of γH2Av-positive GSCs to levels comparable to non-magnifying males **(Fig. 3d-e;** n =

537, $p = 7.1 \times 10^{-4}$**)**, indicating that R2 is responsible for the DSBs formed in GSCs during rDNA magnification. Taken together, these results suggest that rDNA-specific endonuclease activity of R2, which functions during R2 retrotransposition, creates DSBs at the rDNA loci that may in turn induce rDNA CN expansion.

rDNA CN regulates R2 expression

Given the threat R2 mobilization poses to the host genome, both by disruption of rRNA function and causing excessive DSB formation[34], how is the potential benefit of R2 to rDNA CN maintenance balanced with the detriment of R2 retrotransposition? We found R2 expression in the germline is specifically de-repressed under conditions of reduced rDNA CN, potentially explaining how the conflicting consequences of R2 expression are resolved. Using RNA fluorescence *in situ* hybridization (RNA FISH) to examine R2 expression at a single cell resolution, we found that the frequency of GSCs expressing R2 was significantly increased in magnifying males (bb^{Z9}/Ybb^{0}), whereas non-magnifying males (bb^{Z9}/Y^{+}) rarely expressed R2 **(Fig. 3f-h**; n = 231, $p = 1.7 \times 10^{-10}$). Moreover, we found that GSCs from aged animals and the sons of old fathers, which inherit reduced rDNA CN[12]**,** also exhibited a higher frequency of R2 expression compared to GSCs from young, wild type flies **(Extended Data Fig. 5a-b, d**; n = 1,247, $p = 8.3 \times 10^{-4}$ for old animals; n = 1,107, $p = 1.5 \times 10^{-4}$ for offspring). Importantly, the frequency of R2 expression among GSCs in the sons of old fathers returned to the basal level after 20 days of age, when rDNA CN was shown to have recovered[12] **(Extended Data Fig. 5c-d**; n = 617, p = 0.036**)**. These results indicate that R2 expression is dynamically regulated in response to changing rDNA CN. Taken together, we propose that R2 expression is finely tuned to function when most beneficial to the host while minimizing unnecessary exposure to the harmful effects of transposition.

R2 is required for germline immortality

165

166 Based on the finding that R2 plays a critical role in maintaining germline rDNA CN, we
167 postulated that R2 is essential to prevent continuous multi-generational rDNA loss capable of
168 causing the extinction of the lineage. In *C. elegans*, the loss of genome integrity is known to cause
169 gradual loss of fertility, a phenotype known as mortal germline (morg)[35]. To test whether R2-
170 mediated rDNA maintenance is required to maintain fertility through generations, we established
171 multiple independent lines expressing R2 RNAi in their germline and tracked their fecundity at
172 each generation through the ability of each line to produce sufficient offspring to establish a new
173 generation **(Extended Data Fig. 6)**. We found many R2 RNAi lines failed to produce sufficient
174 progeny, with over half failing by the fourth generation, while nearly all control lines survived
175 throughout the duration of the experiments **(Fig. 4a)**, indicating that R2 is essential for continuity
176 of the germline lineage. Males of extinguishing R2 RNAi lineages had ~20% reduction in rDNA
177 CN compared to control lines ($n = 22$, $p = 0.031$) **(Fig 4b)**. These results suggest that R2-mediated
178 maintenance of rDNA contributes to germline immortality.

179

180 Together, our findings reveal an unanticipated 'function' of retrotransposons to benefit the
181 host genome through a role in rDNA CN maintenance. Repetitive DNA sequences such as rDNA
182 are among the most vulnerable elements of the eukaryotic genome[36,37]. We propose that DSBs
183 generated by R2 in GSCs with diminished rDNA loci stimulate sister chromatid recombination that
184 results in rDNA CN expansion **(Fig. 4c)**. Our proposed 'function' for R2 in rDNA maintenance may
185 represent a novel mutualistic host-TE relationship, which are rarely described in eukaryotes[38].
186 R2's de-repression when it can be beneficial (i.e. decreased rDNA CN) may be the key to this
187 mutualistic host-TE relationship. It awaits future investigation to understand how the mechanisms
188 that repress R2[39–41] may tune this activity for host's benefit. The widespread presence of R2 and
189 other rDNA-specific TEs in both vertebrates and invertebrates[22] suggests that similar host-TE
190 mutualism may support rDNA CN maintenance throughout Metazoa. Interestingly, many of these

191 other R2-like elements have little sequence similarity to R2, instead appearing to be derived from
192 other non-specific TEs[22], suggesting this host-TE mutualism may have evolved multiple times
193 over the course of evolution. In summary, our study provides an example of beneficial TEs in the
194 maintenance of eukaryotic genomes, and we propose that more beneficial TEs are yet to be
195 discovered.

196 **References**

197

198 1. Palazzo, A. F. & Lee, E. S. Non-coding RNA: what is functional and what is junk? *Frontiers Genetics*
199 **6**, 2 (2015).

200 2. Schmidtke, J., Zenzes, M. T., Weiler, C., Bross, K. & Engel, W. Gene action in fish of tetraploid
201 origin. IV. Ribosomal DNA amount in clupeoid and salmonoid fish. *Biochem Genet* **14**, 293–297 (1976).

202 3. Miller, L. & Knowland, J. The number and activity of ribosomal RNA genes in Xenopus laevis
203 embryos carrying partial deletions in both nucleolar organizers. *Biochem Genet* **6**, 65–73 (1972).

204 4. Henderson, A. S., Eicher, E. M., Yu, M. T. & Atwood, K. C. The chromosomal location of ribosomal
205 DNA in the mouse. *Chromosoma* **49**, 155–160 (1974).

206 5. Lofgren, L. A. *et al.* Genome-based estimates of fungal rDNA copy number variation across
207 phylogenetic scales and ecological lifestyles. *Mol Ecol* **28**, 721–730 (2019).

208 6. Rabanal, F. A. *et al.* Unstable Inheritance of 45S rRNA Genes in Arabidopsis thaliana. *G3 Genes*
209 *Genomes Genetics* **7**, 1201–1209 (2017).

210 7. Park, P. U., Defossez, P.-A. & Guarente, L. Effects of Mutations in DNA Repair Genes on Formation
211 of Ribosomal DNA Circles and Life Span inSaccharomyces cerevisiae. *Mol Cell Biol* **19**, 3848–3856
212 (1999).

213 8. Ganley, A. R. D. & Kobayashi, T. Ribosomal DNA and cellular senescence: new evidence supporting
214 the connection between rDNA and aging. *Fems Yeast Res* **14**, 49–59 (2014).

215 9. Hawley, R. S. & Marcus, C. H. Recombinational Controls of rDNA Redundancy in Drosophila. *Annu*
216 *Rev Genet* **23**, 87–120 (1989).

217 10. Strehler, B. L., Chang, M.-P. & Johnson, L. K. Loss of hybridizable ribosomal DNA from human
218 post-mitotic tissues during aging: I. Age-dependent loss in human myocardium. *Mech Ageing Dev* **11**,
219 371–378 (1979).

220 11. Johnson, R. & Strehler, B. L. Loss of Genes coding for Ribosomal RNA in Ageing Brain Cells.
221 *Nature* **240**, 412–414 (1972).

222 12. Lu, K. L., Nelson, J. O., Watase, G. J., Warsinger-Pepe, N. & Yamashita, Y. M. Transgenerational
223 dynamics of rDNA copy number in Drosophila male germline stem cells. *Elife* **7**, e32421 (2018).

224 13. Kobayashi, T. Ribosomal RNA gene repeats, their stability and cellular senescence. *Proc Jpn Acad*
225 *Ser B* **90**, 119–129 (2014).

226 14. Ritossa, F. M. Unstable redundancy of genes for ribosomal RNA. *Proc National Acad Sci* **60**, 509–
227 516 (1968).

228 15. Reams, A. B. & Roth, J. R. Mechanisms of Gene Duplication and Amplification. *Csh Perspect Biol* **7**,
229 a016592 (2015).

230 16. Nelson, J. O., Watase, G. J., Warsinger-Pepe, N. & Yamashita, Y. M. Mechanisms of rDNA Copy
231 Number Maintenance. *Trends Genet* **35**, 734–742 (2019).

232 17. Burke, W. D., Malik, H. S., Jones, J. P. & Eickbush, T. H. The domain structure and
233 retrotransposition mechanism of R2 elements are conserved throughout arthropods. *Mol Biol Evol* **16**,
234 502–511 (1999).

235 18. Kojima, K. K., Kuma, K., Toh, H. & Fujiwara, H. Identification of rDNA-Specific Non-LTR
236 Retrotransposons in Cnidaria. *Mol Biol Evol* **23**, 1984–1993 (2006).

237 19. Burke, W. D., Singh, D. & Eickbush, T. H. R5 Retrotransposons Insert into a Family of Infrequently
238 Transcribed 28S rRNA Genes of Planaria. *Mol Biol Evol* **20**, 1260–1270 (2003).

239 20. Burke, W. D., Müller, F. & Eickbush, T. H. R4, a non-LTR retrotransponson specific to the large
240 subunit rRNA genes of nematodes. *Nucleic Acids Res* **23**, 4628–4634 (1995).

241 21. Kojima, K. K. & Fujiwara, H. Cross-Genome Screening of Novel Sequence-Specific Non-LTR
242 Retrotransposons: Various Multicopy RNA Genes and Microsatellites Are Selected as Targets. *Mol Biol*
243 *Evol* **21**, 207–217 (2004).

244 22. Kojima, K. K. & Fujiwara, H. Long-Term Inheritance of the 28S rDNA-Specific Retrotransposon R2.
245 *Mol Biol Evol* **22**, 2157–2165 (2005).

246 23. Yang, J., Malik, H. S. & Eickbush, T. H. Identification of the endonuclease domain encoded by R2
247 and other site-specific, non-long terminal repeat retrotransposable elements. *Proc National Acad Sci* **96**,
248 7847–7852 (1999).

249 24. Fuller, M. T. & Spradling, A. C. Male and Female Drosophila Germline Stem Cells: Two Versions of
250 Immortality. *Science* **316**, 402–404 (2007).

251 25. Lindsley, D. L. & Sandler, L. The Meiotic Behavior of Grossly Deleted X Chromosomes in
252 Drosophila Melanogaster. *Genetics* **43**, 547–63 (1958).

253 26. Ritossa, F. M., Atwood, K. C. & Spiegelman, S. A molecular explanation of the bobbed mutants of
254 Drosophila as partial deficiencies of "ribosomal" DNA. *Genetics* **54**, 819–34 (1966).

255 27. Tartof, K. D. Regulation of ribosomal RNA gene multiplicity in Drosophila melanogaster. *Genetics*
256 **73**, 57–71 (1973).

257 28. Doren, M. V., Williamson, A. L. & Lehmann, R. Regulation of zygotic gene expression in Drosophila
258 primordial germ cells. *Curr Biol* **8**, 243–246 (1998).

259 29. Chen, D. & McKearin, D. M. A discrete transcriptional silencer in the bam gene determines
260 asymmetric division of the Drosophila germline stem cell. *Development* **130**, 1159–1170 (2003).

261 30. Govindaraju, A., Cortez, J. D., Reveal, B. & Christensen, S. M. Endonuclease domain of non-LTR
262 retrotransposons: loss-of-function mutants and modeling of the R2Bm endonuclease. *Nucleic Acids Res*
263 **44**, 3276–3287 (2016).

264 31. Gangloff, S., Zou, H. & Rothstein, R. Gene conversion plays the major role in controlling the stability
265 of large tandem repeats in yeast. *Embo J* **15**, 1715–1725 (1996).

266 32. Tartof, K. D. Unequal Mitotic Sister Chromatid Exchange as the Mechanism of Ribosomal RNA
267 Gene Magnification. *Proc National Acad Sci* **71**, 1272–1276 (1974).

268 33. Paredes, S. & Maggert, K. A. Expression of I-CreI Endonuclease Generates Deletions Within the
269 rDNA of Drosophila. *Genetics* **181**, 1661–1671 (2009).

270 34. Eickbush, T. H. & Eickbush, D. G. Integration, regulation, and long-term stability of R2
271 retrotransposons. *Microbiol Spectr* **3**, 1127–1146 (2015).

272 35. Smelick, C. & Ahmed, S. Achieving immortality in the C. elegans germline. *Ageing Res Rev* **4**, 67–82
273 (2005).

274 36. Brown, R. E. & Freudenreich, C. H. Structure-forming repeats and their impact on genome stability.
275 *Curr Opin Genet Dev* **67**, 41–51 (2021).

276 37. Pâques, F., Leung, W.-Y. & Haber, J. E. Expansions and Contractions in a Tandem Repeat Induced
277 by Double-Strand Break Repair. *Mol Cell Biol* **18**, 2045–2054 (1998).

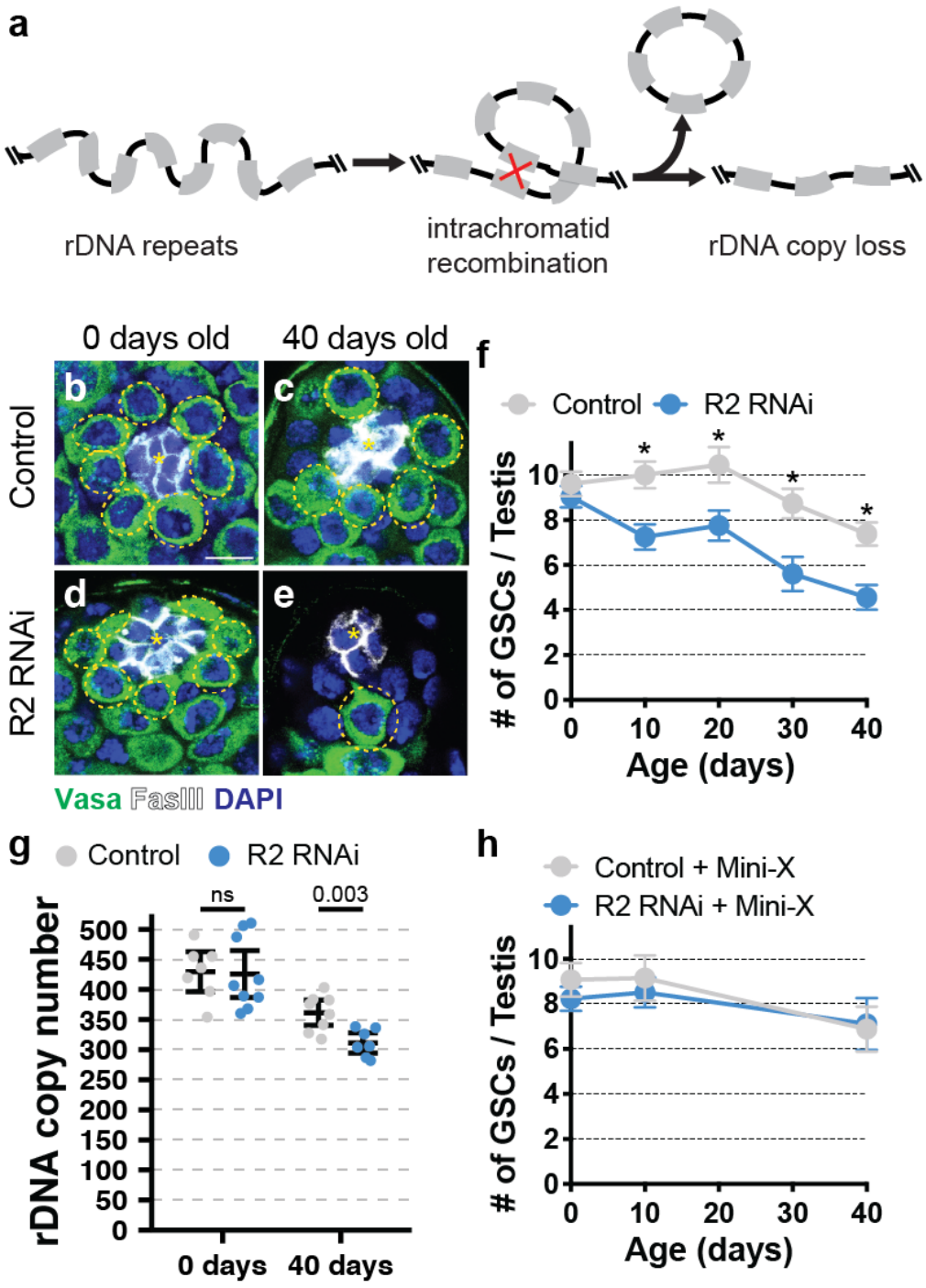
278 38. Cosby, R. L., Chang, N.-C. & Feschotte, C. Host–transposon interactions: conflict, cooperation, and
279 cooption. *Gene Dev* **33**, 1098–1116 (2019).

280 39. Quénerch'du, E., Anand, A. & Kai, T. The piRNA pathway is developmentally regulated during
281 spermatogenesis in Drosophila. *Rna* **22**, 1044–1054 (2016).

282 40. He, F., James, A., Raje, H., Ghaffari, H. & DiMario, P. Deletion of Drosophila Nopp140 induces
283 subcellular ribosomopathies. *Chromosoma* **124**, 191–208 (2015).

284 41. Guerrero, P. A. & Maggert, K. A. The CCCTC-Binding Factor (CTCF) of Drosophila Contributes to
285 the Regulation of the Ribosomal DNA and Nucleolar Stability. *Plos One* **6**, e16401 (2011).

286



287

288 **Figure 1: R2 is required for GSC maintenance via rDNA CN maintenance during**

289 **Drosophila male germline aging. (a)** Tandem rDNA copy repeats are lost through

290 circularization and extrusion caused by intra-chromatid recombination between rDNA copies. **(b-**

291 **e)** GSCs in 0- and 40-day old control (*nos-gal4* driver alone) (b-c) and R2 RNAi (*nos-gal4*

292 driving *UAS-R2 RNAi*) testes (d-e). Yellow dotted circle = GSC. The major component of the
293 GSC signaling niche, the hub, is indicated by *. Green = Vasa, White = Fascillin III, Blue =
294 DAPI. Scale bar = 7.5 μM **(f)** Average number of GSCs per testis in control and R2 RNAi
295 animals during aging. * indicates $p < 10^{-5}$ determined by Student's t-test. Error bars = 95%
296 confidence interval (CI). **(g)** rDNA CN determined by ddPCR in testes of 0- and 40-day old
297 control and R2 RNAi animals. P-value determined by Student's t-test. Error bars = 95% CI. **(h)**
298 Average number of GSCs per testis during aging in Control and R2 RNAi animals with mini-X
299 chromosome. Error bars = 95% CI.
300

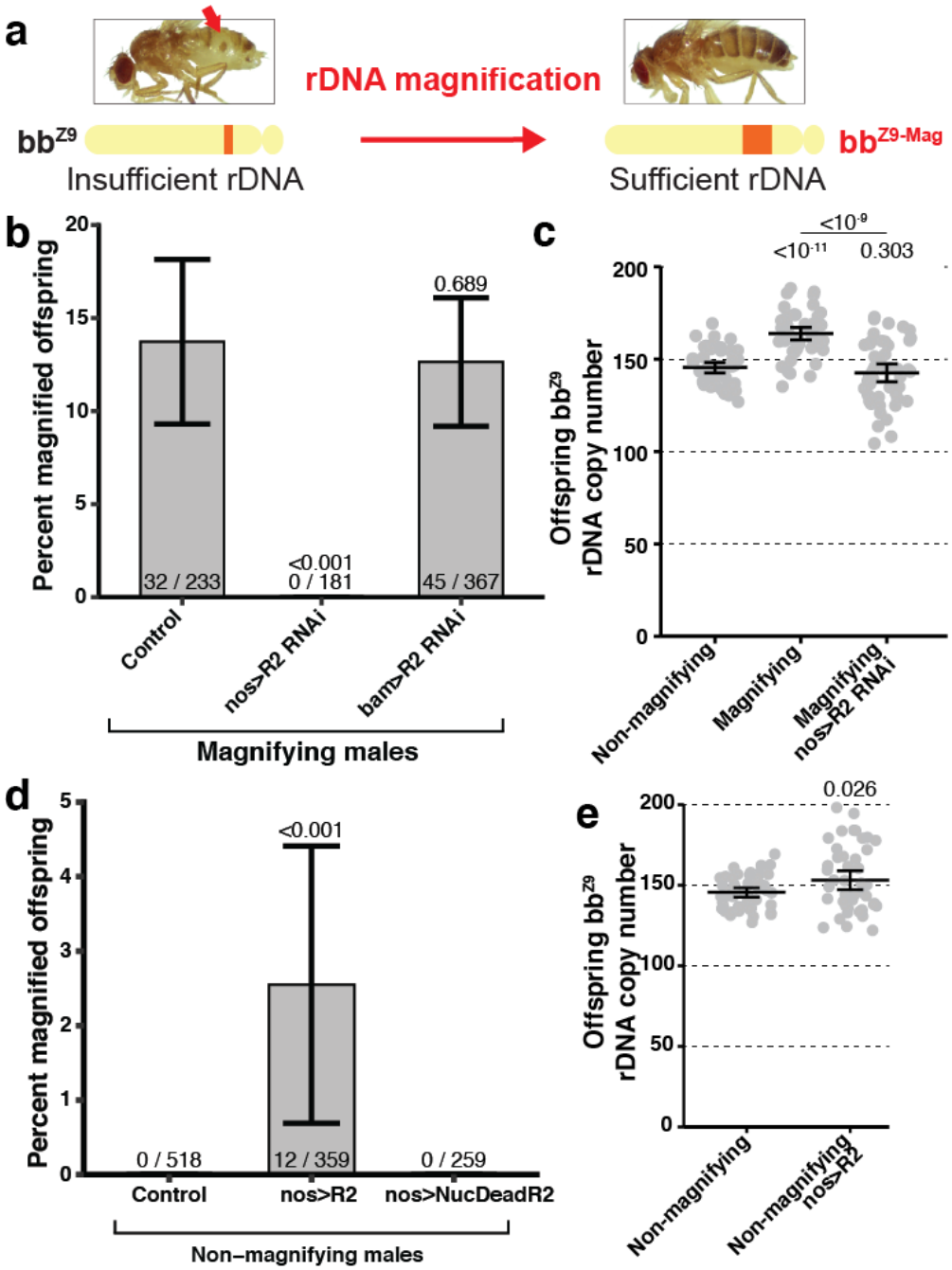


Figure 2: R2 is necessary and sufficient for rDNA magnification. (a) Diagram of rDNA magnification. Animals with insufficient rDNA CN due to having the bb^{Z9} locus as their only source of rDNA have distinct developmental defects resulting in gaps in the dorsal cuticle (red arrow). rDNA CN expansion at the bb^{Z9} locus due to rDNA magnification is detected by reversion of this phenotype to normal cuticles. **(b)** Percent magnified offspring determined by

307 cuticular phenotype in offspring from magnifying males of the indicated genotypes. P-value
308 determined by chi-squared test. Error bars = 95% CI. **(c)** bb^{Z9} locus rDNA CN determined by
309 ddPCR in daughters from males of indicated genotypes. P-value determined by Student's t-test.
310 Error bars = 95% CI. Non-magnifiying condition is the same data as panel c. **(d)** Percent
311 magnified offspring from non-magnifying males of the indicated genotypes. P-value determined
312 by chi-squared test. Error bars = 95% CI. **(e)** bb^{Z9} locus rDNA CN determined by ddPCR in
313 daughters from non-magnifying males of the indicated genotype. P-value determined by
314 Student's t-test. Error bars = 95% CI.

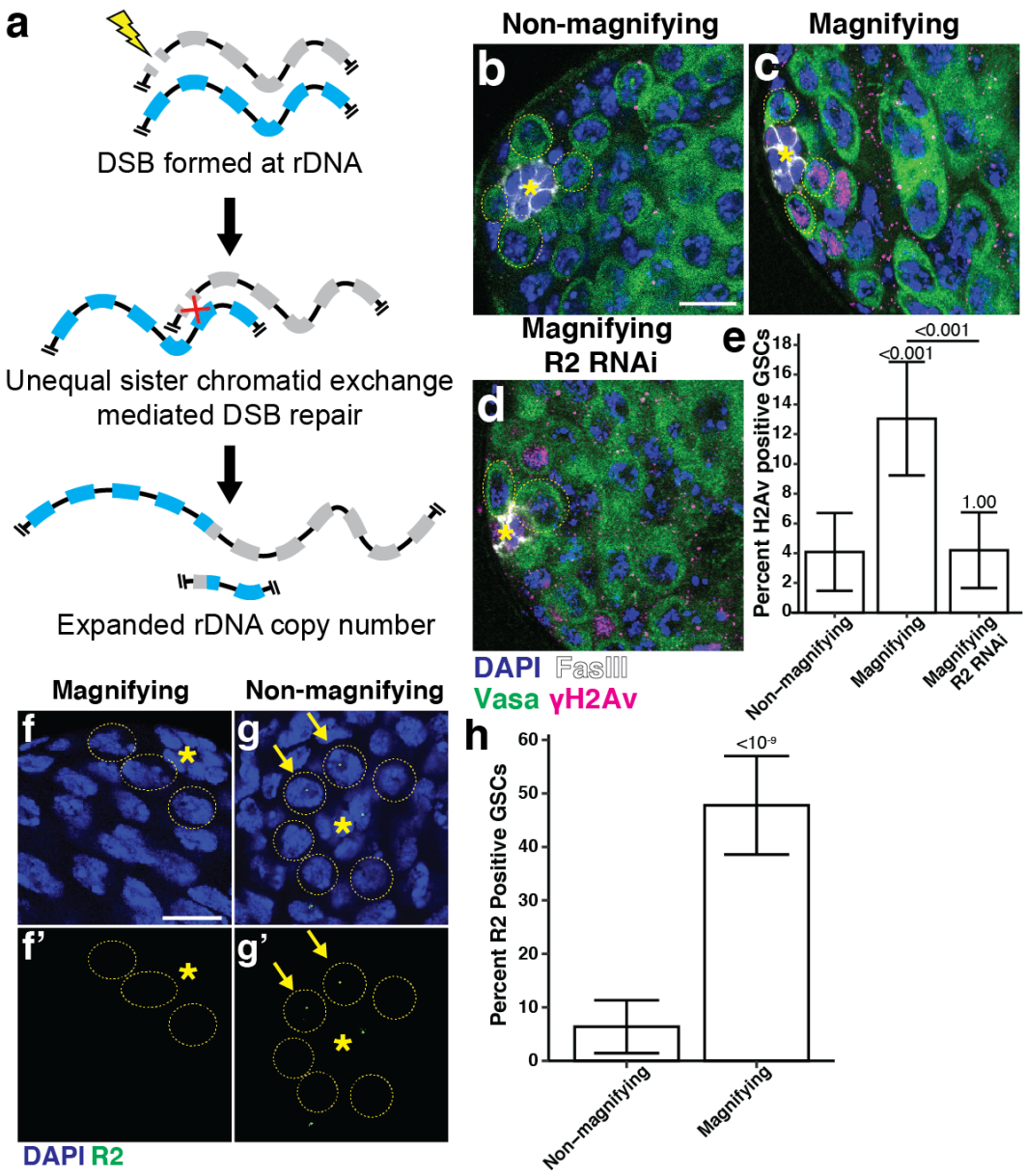


Fig 3: Derepressed R2 creates DSBs in GSCs during rDNA magnification. (a) Diagram of rDNA CN expansion by unequal sister chromatid exchange during DSB repair at rDNA loci. Recombination between misaligned rDNA copies during DSB repair result in crossovers that create unequal sister chromatid exchange that increases rDNA CN on one chromatid. **(b-d)** Detection of DSBs in the early adult male germline by anti-γH2Av staining. R2 RNAi expressed under the *nos-gal4* driver. Non-RNAi conditions contain the *nos-gal4* driver alone. GSCs indicated by yellow dotted circle. Blue = DAPI, Green = vasa, Magenta = γH2Av, white = FasIII.

323 The hub is indicated by *. Scale bar = 10 μM. **(e)** Percentage of γH2Av positive GSCs in the
324 indicated conditions. P-value determined by chi-squared test. Error bars = 95% CI. **(f-g)** R2
325 expression in GSCs (yellow dotted circle). Blue = DAPI, Green = R2 mRNA. Isolated R2
326 channel in f'-g'. The hub is indicated by *. R2 positive cells GSCs are marked by yellow
327 arrowhead. **(h)** Percentage of R2 positive GSCs. P-value determined by chi-squared test. Error
328 bars = 95% CI.

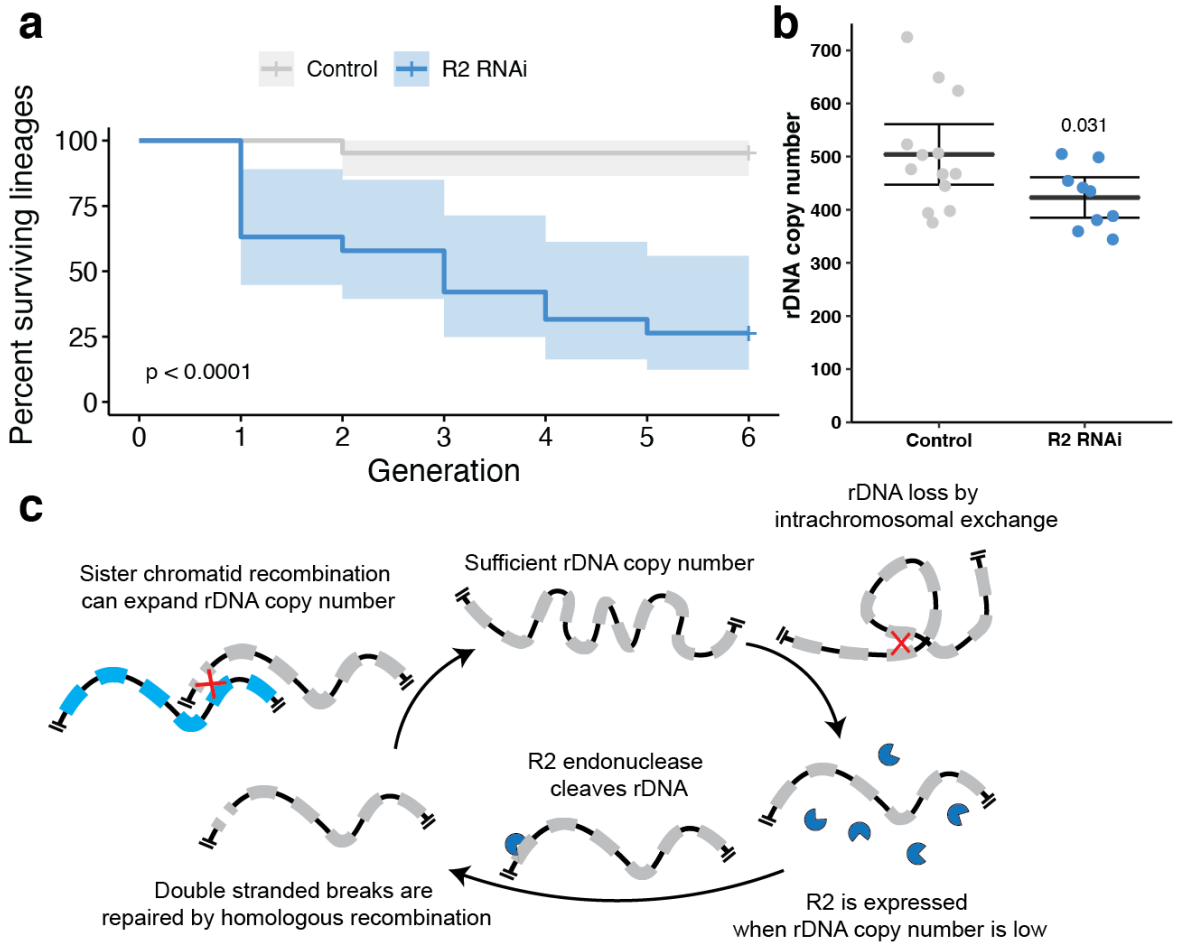


Fig 4: R2 is required to maintain rDNA CN and fertility over successive generations. (a) Kaplan-Meier curve of lineage survival in R2 RNAi and control lineages. Each lineage constitutes an individual data point. P-value determined by log rank test. Shade represents 95% CI. **(b)** rDNA CN determined by ddPCR in males of control (*nos-gal4* driver alone) animals at the 6th generation or R2 RNAi (RNAi expressed under the *nos-gal4* driver) lineages at their terminating generation. P-value determined by Student's t-test. Error = 95% CI. **(c)** Model of the role of R2 in germline rDNA CN maintenance. R2 derepression in cells with reduced rDNA CN leads to DSBs formed by rDNA-specific R2 endonuclease that can induce rDNA CN expansion through sister chromatid recombination during their repair.

340 **Methods**

341

342 *Immunofluorescence*

343 Immunofluorescence staining of testes was performed as previously described[42]. Briefly, testes
344 were dissected in PBS, fixed in 4% formaldehyde in PBS for 30 minutes, then briefly washed two
345 times in PBS containing 0.1% Triton-X (PBS-T), followed by washing in PBS-T for 30 minutes.
346 After washes, samples were incubated at 4°C overnight with primary antibody in 3% bovine serum
347 albumin (BSA) in PBS-T. Samples were washed three consecutive times for 20 minutes in PBS-
348 T, then incubated at 4°C overnight with secondary antibody in 3% BSA in PBS-T, washed three
349 times again in PBS-T for 20 minutes, and mounted in VECTASHIELD with DAPI (Vector Labs).
350 The following primary antibodies were used: rat anti-vasa (1:20; DSHB; developed by A.
351 Spradling), mouse anti-Fascillin III (1:200; DSHB; developed by C. Goodman), and rabbit anti-γ-
352 H2AvD pS137 (1:200; Rockland). Images were taken with a Leica Stellaris 8 confocal microscope
353 with 63x oil-immersion objectives and processed using Fiji (ImageJ) software.

354

355 *RNA FISH*

356 RNA FISH samples were prepared as previously described[12]. In short, dissected testes were fixed
357 in 4% formaldehyde in PBS for 30 minutes, briefly washed in PBS, and permeabilized in 70%
358 ethanol overnight at 4º. Samples were then briefly rinsed in 2xSSC with 10% formamide prior to
359 hybridization with 50nM Stellaris probes overnight at 37º. Samples were washed twice in 2xSSC
360 with 10% formamide for 30 minutes and mounted in VECTASHIELD with DAPI (Vector Labs).
361 Samples were imaged using a Leica Stellaris 8 confocal microscope with 63x oil-immersion
362 objectives and processed using Fiji (ImageJ) software.

363

364 *DNA isolation*

365 Testis DNA was isolated from 50 pooled dissected testes frozen in liquid N2. DNA isolation was
366 performed according to previously described methods for isolation from Drosophila tissues[43]. DNA
367 was isolated from individual Drosophila animals using a modified protocol of the DNeasy Blood
368 and Tissue DNA extraction kit (Qiagen). In short, individual animals were homogenized in 200 µL
369 Buffer ATL containing proteinase K using a pipette tip in Eppendorf tubes, vortexed for 15
370 seconds, and incubated for 1.5 hours at 56º. Samples were then prepared following the
371 manufacturer's protocol after incubation. All DNA samples were quantified and checked for purity
372 by Nanodrop One spectrophotometer (ThermoFisher).
373
374 *rDNA copy number measurement by Droplet Digital PCR (ddPCR)*
375 30 ng of DNA sample were used per 20 µL ddPCR reaction for control gene reactions (RpL and
376 Upf1), and 0.3 ng of DNA per 20 µL ddPCR reaction for 28S rDNA reactions. Primers and probes
377 for reactions are listed at **Extended Data Table 1**. ddPCR reactions were carried out according
378 to the manufacturer's (Bio Rad) protocol. In short, master mixes containing ddPCR Supermix for
379 Probes (No dUTP) (Bio Rad), DNA samples, primer / probe mixes, and HindIII-HF restriction
380 enzyme (New England Biolabs) for 28S rDNA reactions (no restriction enzyme for control gene
381 reactions) were prepared in 0.2 mL Eppendorf tubes, and incubated at room temperature for 15
382 minutes to allow for restriction enzyme digestion. ddPCR droplets were generated from samples
383 using QX200 Droplet Generator (Bio Rad) and underwent complete PCR cycling on a C100 deep-
384 well thermocycler (Bio Rad). Droplet fluorescence was read using the QX200 Droplet Reader (Bio
385 Rad). Sample copy number was determined using Quantasoft software (Bio Rad). rDNA copy
386 number per genome was determined by 28S sample copy number multiplied by 100 (due to the
387 100x dilution of sample in the 28S reaction compared to control reaction) divided by control gene
388 copy number multiplied by the expected number of control gene copies per genome (2 for RpL in
389 all samples; 2 for Upf1 in female samples; 1 for Upf1 in male samples). The 28S copy number

390 values determined by each control gene was averaged to determine 28S copy number for each
391 sample.
392
393 *Mitotic chromosome spread and DNA FISH*
394 Larval brain dissections, mitotic chromosome spreads, DNA FISH, and imaging were all done as
395 previously[44]. In short, brains were dissected from male third instar larvae in PBS and fixed in 25µ
396 µL of acetid acid and 4% formaldehyde in PBS. Samples were applied to Superfrost plus slides
397 and manually squashed under a coverslip, then immediately frozen in liquid N2. After freezing
398 slides were removed from N2, coverslip removed, and slides were dehydrated in 100% ethanol
399 and dried at room temperature. DNA FISH hybridization was performed in 20 µL of 50%
400 formamide, 10% dextran sulfate, 2X SSC buffer and 0.5 µM each probe applied directly to the
401 sample on the slide and covered with a cover slip. Samples were incubated at 95° for 5 minutes,
402 cooled and wrapped in parafilm, then incubated overnight at room temperature in a dark humid
403 chamber. Coverslips were removed and slides were washed three times for 15 minutes in 0.1x
404 SSC, dried, and then mounted in VECTASHIELD with DAPI (Vector Labs). Samples were imaged
405 using a Leica Stellaris 8 confocal microscope with 63x oil-immersion objectives and processed
406 using Fiji (ImageJ) software. Probes used for this study are as follows: 359, 5'-
407 AGGATTTAGGGAAATTAATTTTTGGATCAATTTTCGCATTTTTTGTAAG-3'-Cy5; $(TAGA)_6$-
408 Cy5; IGS, 5'-
409 AGTGAAAAATGTTGAAATATTCCCATATTCTCTAAGTATTATAGAGAAAAGCCATTTTAGTGA
410 ATGGA-3'-Alexa488; $(AATAC)_6$-Cy3; and $(AATAAAC)_6$-Cy3.
411
412 *Generation of rDNA deletion animals*
413 rDNA copy number loss was induced during larval development in *yw* / Y ;; *HS-I-CreI*, *Sb* / TM6B
414 males with a *y, w* X chromosome by I-CreI expression as previously described[33]. In brief, parental
415 animals mated and laid eggs for three days, then removed from food. After one additional day of

416 larval development, animals were exposed to 37° C heat shock for 45 minutes on two consecutive
417 days. identify X chromosomes with significant rDNA copy number reduction (bb), adult males that
418 experienced I-CreI expression were mated to *bb*^{158} / *FM6* females, and virgin non-*FM6* daughters
419 (bb / bb^{158}) were screened for the *bobbed* phenotype. 28 out of 946 non-*FM6* daughters screened
420 were *bobbed*. To isolate potentially reduced rDNA loci and remove HS-I-CreI from the
421 background, TM6B containing *bobbed* females were individually mated to wild-type males. Male
422 offspring from each individual female candidates were subsequently individually mated to *bb*^{158} /
423 *FM6* females. Any mating that failed to produce non-FM6 daughters were eliminated (due to
424 having the bb^{158} and not the candidate bb chromosome). All viable non-FM6 daughters were
425 double checked for the *bobbed* phenotype, and stocks with all *bobbed* non-FM6 daughters had
426 FM6 containing siblings collected and used to establish a bb / FM6 stock. This method isolated
427 the novel rDNA deletion allele, bb^{Z9}, used in this study.
428
429 *Drosophila genetics*
430 All Drosophila lines used in this study are found in **Extended Data Table 2**. All animals we reared
431 on standard Bloomington medium at 25º. All aging was done in roughly 1:1 mixed presence of
432 males and females, provided fresh food every four to six days. UAS:R2 RNAi strains were
433 designed using SNAP-DRAGON shRNA target software, and oligos containing target hairpin
434 sequence were cloned into the WALLIUM20 vector for phiC31 site-directed integration into the
435 Drosophila genome. All UAS:R2 transgenes were synthesized into pUAST:attB by VetcorBuilder
436 (Chicago, IL) for site-directed integration. All injections and selection of animals containing
437 integrated transgenes were performed by BestGene, Inc (Chino Hills, CA).
438
439 *Lineage survival assay*
440 Independent lineages of nos:Gal4 / CyO; UAS:R2 RNAi / Tm6B or nos:Gal4 / CyO; TM2 / TM6B
441 animals were established by collecting siblings of the indicated genotypes from nos:Gal4 / CyO;

442 TM2 / TM6B males mated to Sp / CyO; UAS:R2 RNAi / TM6B females. At each generation in
443 each lineage, 3 males were mated with 5 females for 5 days, and offspring were collected 10 days
444 after mated ended. Any lineages that did not have at least 3 males and 5 females at collection
445 time were terminated due to insufficient animals.
446
447 *Statistics*
448 For all comparisons of percentage of samples with categorical values (percent γH2Av or R2
449 positive cells), significance was determined by chi-squared test. For all comparisons of samples
450 with independent values (number of GSCs; rDNA copy number), significance was determined by
451 Student's t-test between experimental and control conditions, unless otherwise indicated.
452

453 **Methods References**

454 42. Cheng, J. *et al.* Centrosome misorientation reduces stem cell division during ageing. *Nature* **456**, 599–
455 604 (2008).

456 43. Huang, A. M., Rehm, E. J. & Rubin, G. M. Quick Preparation of Genomic DNA from Drosophila.
457 *Cold Spring Harb Protoc* **2009**, pdb.prot5198-pdb.prot5198 (2009).

458 44. Jagannathan, M., Warsinger-Pepe, N., Watase, G. J. & Yamashita, Y. M. Comparative Analysis of
459 Satellite DNA in the Drosophila melanogaster Species Complex. *G3 Genes Genomes Genetics* **7**, 693–
460 704 (2017).

461 **Acknowledgements**

462

463 We thank the Bloomington Drosophila Stock Center, Kyoto Drosophila Stock Center and
464 Developmental Studies Hybridoma Bank for reagents. We thank the Yamashita lab members
465 and Dr. Andy Clark for discussion and comments on the manuscript. We thank ATCC for design
466 of the 28S and RpL ddPCR assays. This research was supported by the Howard Hughes
467 Medical Institute. Jonathan Nelson is supported by a Postdoctoral Fellowship (133949-PF-19-
468 133-01-DMC) from the American Cancer Society.

469

470 **Author Contributions**

471 J.O.N. contributed to the conception, design, and execution of experiments, the collection,
472 management, and analysis of data, and the writing and editing of the manuscript. A.S.
473 contributed to the execution of experiments, the collection and management of data, and the
474 editing of the manuscript. Y.M.Y contributed to the conception and design of experiments, the
475 management and analysis of data, and the writing and editing of the manuscript.

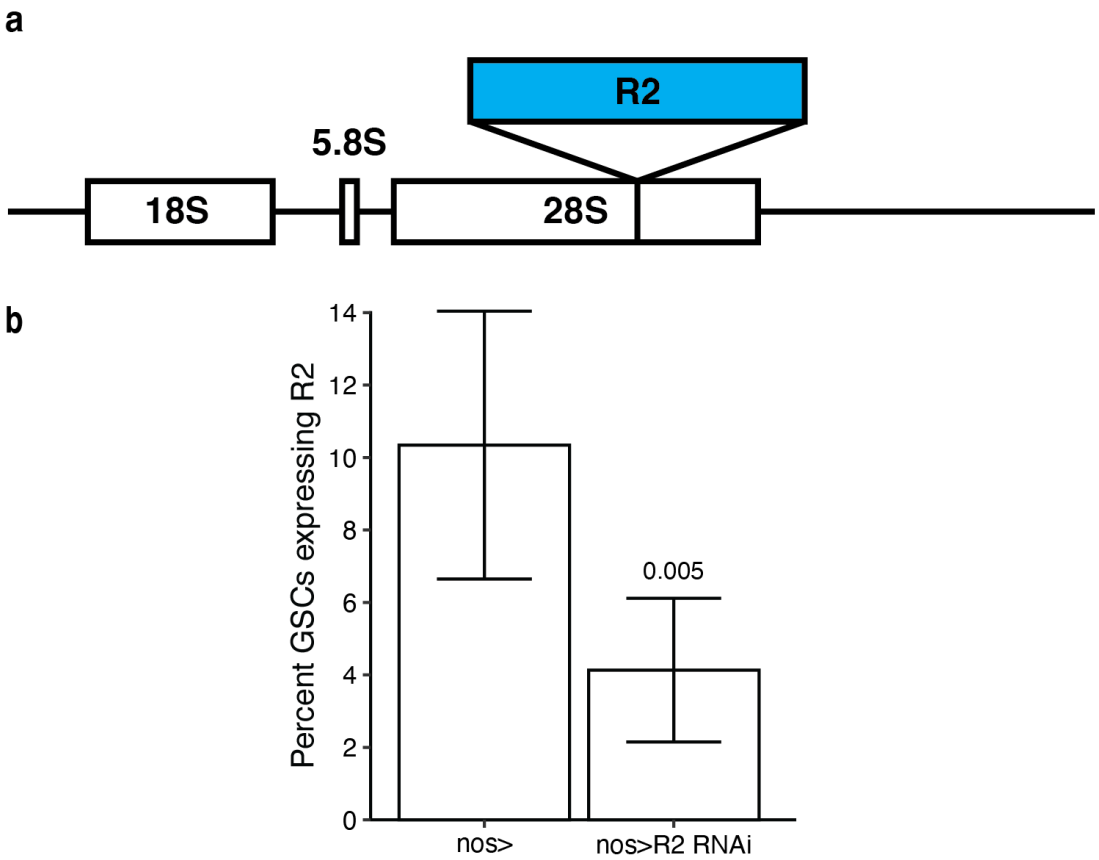
476

Competing Interests

The authors declare no competing interests.

479 **Extended Data**

480



481

482 **Extended Data Figure 1: Design and efficiency of R2 RNAi. (a)** Diagram the 45S rRNA cistron

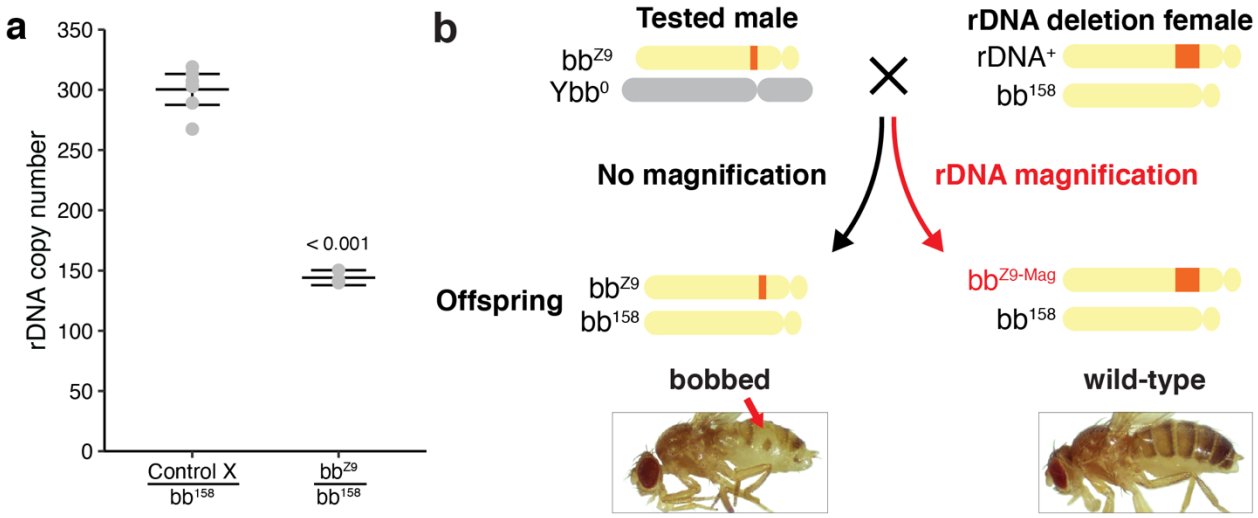
483 constituting an 'rDNA copy' and the sequence-specific R2 insertion site in the 28S rRNA gene. **(b)**

484 Confirmation of R2 expression knockdown by R2 RNAi. Precent of GSCs expressing R2 detected by

485 RNA FISH in *nos-gal4* driver only condition (nos>) and R2 RNAi expressed by the *nos-gal4* driver

486 (nos>R2 RNAi). P-value determined by chi-squared test. Error bars = 95% CI.

487



488

489 **Extended Data Figure 2: Diagram of phenotypic assessment to detect rDNA**

490 **magnification at the bb^{Z9} allele. (a)** rDNA CN quantification of the bb^{Z9} allele determined by

491 ddPCR. All samples from individual females with either the bb^{Z9} allele or X chromosome from

492 which the allele was isolated heterozygous with an X chromosome completely lacking rDNA

493 (bb^{158}). P-value determined by Student's t-test. Error bars = 95% CI. **(b)** Scheme to detect rDNA

494 magnification of bb^{Z9} allele. Tested magnifying males (bb^{Z9} / Ybb^{0}) are crossed to females

495 heterozygous for the bb^{158} rDNA deletion allele. Resultant bb^{Z9} / bb^{158} daughters completely rely

496 on the bb^{Z9} locus for their rDNA. (Left) bb^{Z9} / bb^{158} offspring with the 'bobbed' cuticular defects

497 associated with rDNA insufficiency (red arrow) are determined to have inherited a bb^{Z9} allele

498 that did not undergo rDNA magnification. (Right) bb^{Z9} / bb^{158} offspring with wild-type cuticles are

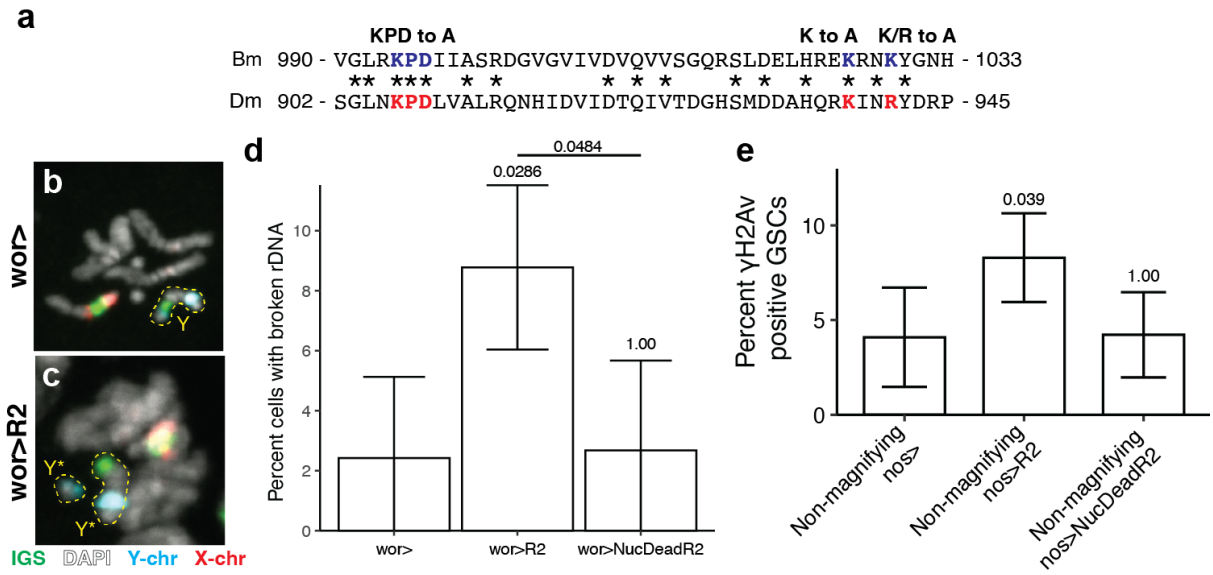
499 determined to have inherited a bb^{Z9} allele that restored normal rDNA content through rDNA

500 magnification ($bb^{Z9\text{-}Mag}$). Thus, the portion of bb^{Z9} / bb^{158} offspring with wild-type cuticles

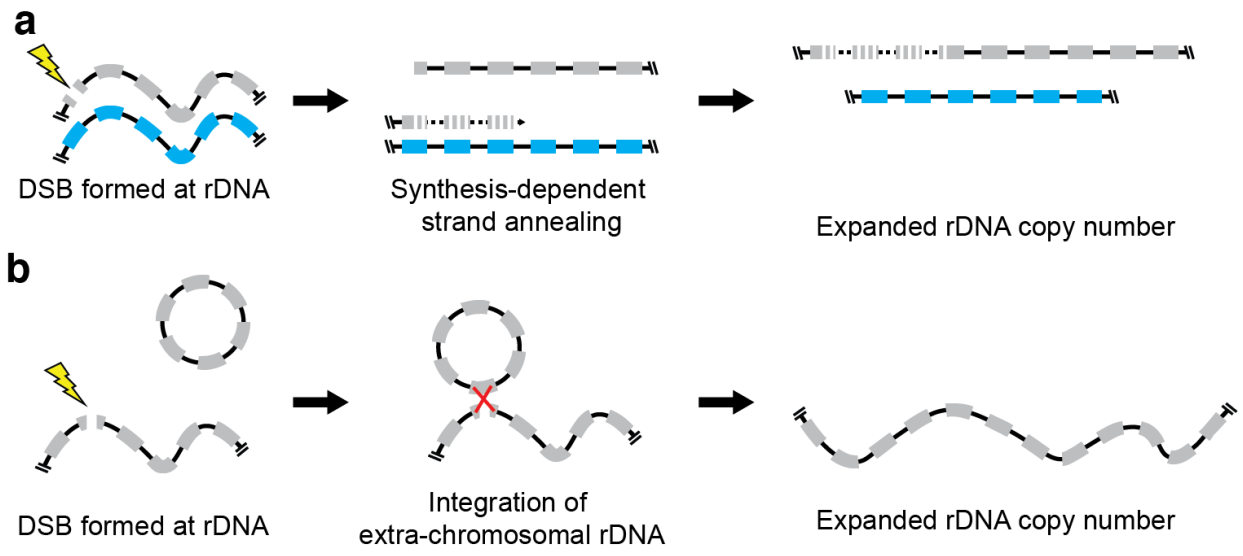
501 represents the portion of bb^{Z9} alleles that underwent rDNA magnification. bb^{Z9} / bb^{158} animals

502 are also collected for rDNA CN quantification, and total rDNA CN in these animals represents

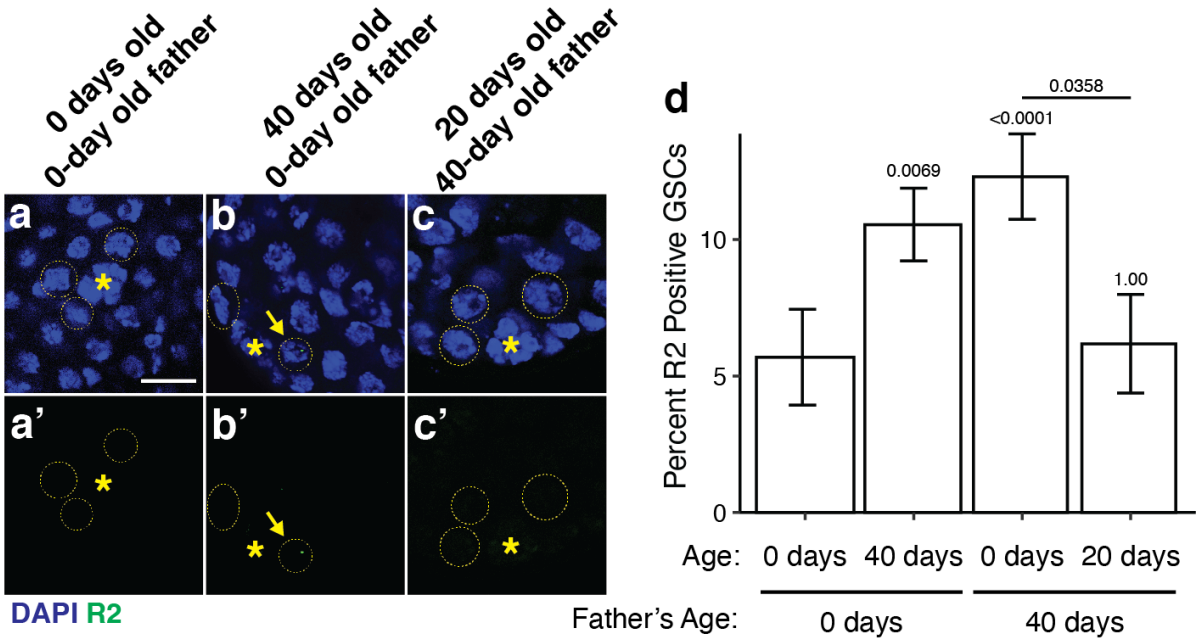
503 the rDNA from the bb^{Z9} locus alone.



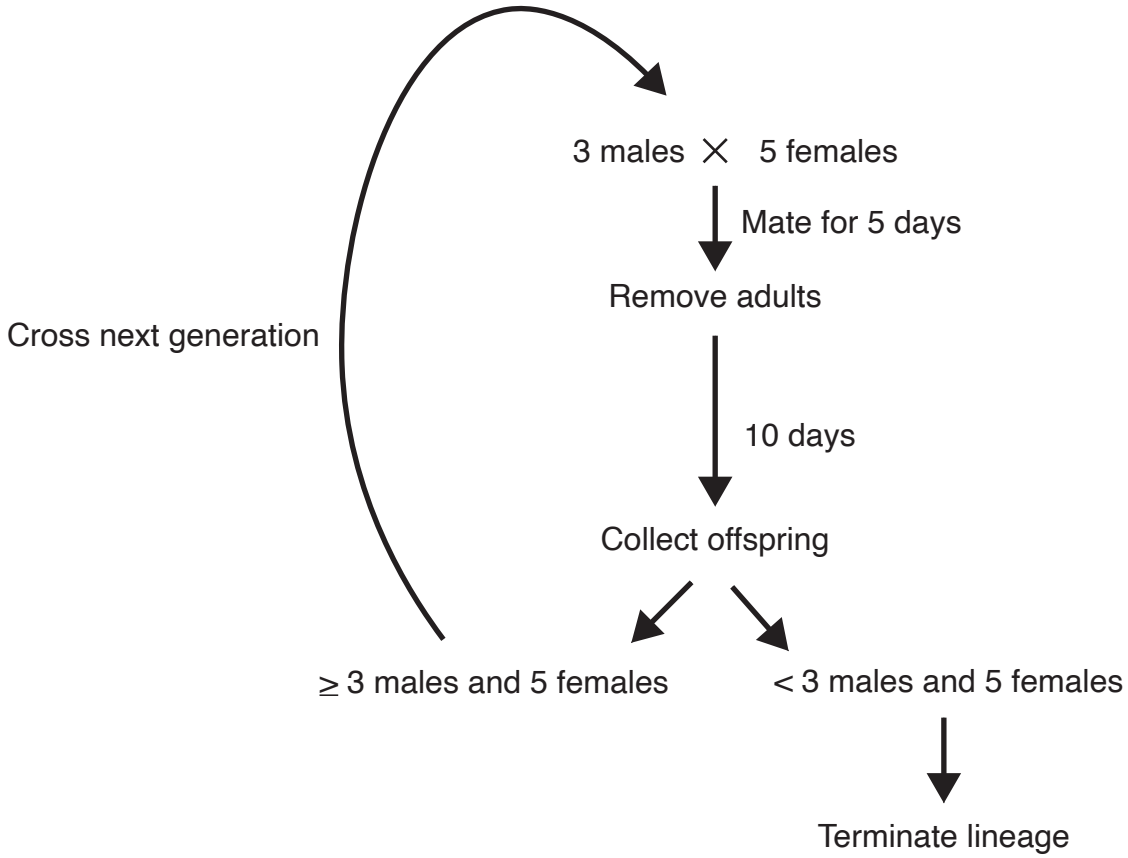
Extended Data Figure 3: Transgenic R2 expression induces double stranded breaks at rDNA loci. (a) Alignment of 44 amino acids in the endonuclease domain of *Bombyx mori* R2 (Bm) and *Drosophila melanogaster* R2 (Dm). Residues shown to be essential for Bm R2 endonuclease activity[30] labeled in blue and respective Dm R2 residues in red. Mutations found to render Bm R2 endonuclease dead[30] and their respective amino acid changes in Dm R2 are indicated above highlighted residues. **(b-c)** DNA FISH in mitotic chromosome spreads of larval neuroblasts in animals with the neuroblast driver *wor-gal4* (wor>, B) or transgenic R2 expressed by the *wor-gal4* driver (wor>R2, B') animals. rDNA is labelled by probes targeting the rDNA intergenic sequence (IGS) (Green), Y- chromosome is labelled by probes targeting AATAC and AATAAAC satellite sequences (Cyan) and X-chromosome is labelled by probes against 359 and TAGA satellite sequences (Red). Y* indicates Y-chromosome fragments broken at the rDNA locus. **(d)** Percentage of neuroblasts with DSBs at rDNA loci determined by DNA FISH in mitotic spreads in indicated genotypes. P-values determined by chi-squared test. Error bars = 95% CI. **(e)** Percentage of γH2Av positive GSCs in the indicated genotypes. Transgenic R2 constructs expressed in the germline by the *nos-gal4* driver (nos>). P-value determined by chi-squared test. Error bars = 95% CI.



Extended Data Figure 4: Models of rDNA CN expansion initiated by DSBs at rDNA locus. **(a)** rDNA CN expansion by strand-dependent strand annealing (SDSA). DSBs at rDNA loci can be repaired via recombination of a single strand with sister chromatid and synthesis. Synthesis can run on to replicate multiple rDNA copies before re-annealing with broken strand. Upon repair after synthesis, rDNA CN is expanded on one sister chromatid while CN of the other sister remains intact. **(b)** DSBs formed at rDNA loci induce homology-dependent repair that causes recombination with existing extrachromosomal rDNA circles (ERCs), allowing for reintegration of the ERC rDNA copies into the genomic rDNA locus.



Extended Data Figure 5: R2 expression in GSCs is regulated in response to changing rDNA copy number during aging. (a-c) R2 RNA FISH in the male germline of young (0-days old), (a) old (40-days old) (b), and recovered offspring of old animals (c). Isolated R2 channel in a'-c'. GSCs indicated by yellow dotted circle. * indicates GSC niche signal hub. R2 positive cells indicated by yellow arrowhead. **(c)** Percentage of R2 positive GSCs at indicated ages in the offspring of 0- or 40-day old fathers. P-value determined by chi-squared test. Error bars = 95% CI.



Extended Data Figure 6: Scheme for testing lineage survival in multi-generational R2 RNAi experiment. Each lineage is established using 3 males and 5 females of the given genotype. Males and females are mated for 5 days and removed from the vial. 10 days later, all offspring are collected. If total offspring from the mating contains at least 3 males and 5 females, then 3 males and 5 females are randomly selected to establish the next generation of the lineage. If the total offspring contained fewer than 3 males or 5 females, then the lineage was terminated and future generations were not established. This procedure was repeated at each generation for each lineage for 6 generations or until termination.

549 **Extended Data Table 1: List of oligomers used in this study for ddPCR assays.**

Oligo Name	Type	Target	Sequence	Modifications
dd-RpL32 F	Primer	RpL32	GCTTCAAGGGACAGTATCTG	
dd-RpL32 R	Primer	RpL32	AACGCGGTTCTGCATGAG	
dd-RpL32 Probe	Probe	RpL32	ATGCCCAACATCGGTTAC	5' HEX AND 3' Iowa Black FQ
dd-28S F	Primer	28S	GAGCTGCCATTGGTACAG	
dd-28S R	Primer	28S	GCTTTCGCCTTGAACTTAG	
dd-28S Probe	Probe	28S	TGGTGGATAGTAGCAAATAATCG	5' 6-FAM AND 3' Iowa Black FQ
dd-Upf1 F	Primer	Upf1	CACACTTTATGTCCACCATTATTG	
dd-Upf1 R	Primer	Upf1	GAGTTTCCGTAGGGACCAC	
dd-Upf1 Probe	Probe	Upf1	CCG TAA CCG CCA CTG CGG T	5' 6-FAM AND 3' Iowa Black FQ

550

551

552 **Extended Data Table 2: List of Drosophila stocks used in this study.**

Genotype	Source	Use
Dp(1;f)122; C(1)RM, y^1 v^1 f^1 / C(1;Y)6, Df(1)259, w^1	Kyoto Drosophila Genomics and Genetic Resource Center (DGRC) #107274	Mini-x Chromosome (*Dp(1;f)*)
yw	Bloomington Stock Center (BSC) #1495	Source for isolating novel rDNA deletion strain
w[1118]; P{v[+t1.8]=hs-I-CreI.R}1A Sb[1]/TM6	BSC #6937	I-CreI endonucelase
nos-gal4	PMID: 9501989	Early germ cell driver
wor-gal4	BSC #56553	Neuroblast driver
bam-gal4	PMID: 12571107	Germ cell driver
y[1] eq[1]/Df(YS)bb[-]	DGRC #101260	Y chromosome complete rDNA deletion
Df(1)bb158, y[1] / Dp(1;Y)y[+] / C(1); ca[1] awd[K]*	DGRC #106876	X chromosome complete rDNA deletion

553