A universal molecular mechanism driving aging

Wan Jin[1,2], Jing Zheng[3], Yu Xiao[1,4], Lingao Ju[1], Fangjin Chen[2,5], Jie Fu[6], Hui Jiang[3], Yi Zhang[2,*]

Affiliations:

[1]Department of Biological Repositories, Human Genetic Resources Preservation Center of Hubei Province, Zhongnan Hospital of Wuhan University, Wuhan, China.

[2]Euler Technology, ZGC Life Sciences Park, Beijing, China.

[3]National Institute of Biological Sciences, Beijing, China

[4]Wuhan Research Center, Chinese Academy of Medical Sciences, Wuhan, China.

[5]High Performance Computing Center, the Peking-Tsinghua College of Life Sciences, Peking University, Beijing, China.

[6]Hongkong University of Science and Technology, Hong Kong, China.

[*]Correspondence to Dr. Yi Zhang (zy@eulertechnology.com).

Running title: DNA G-quadruplexes drive aging.

Abstract

How cell replication ultimately results in aging and the Hayflick limit are not fully understood. Here we show that clock-like accumulation of DNA G-quadruplexes (G4s) throughout cell replication drives conserved aging mechanisms. G4 stimulates transcription-replication interactions to delay genome replication and impairs DNA re-methylation and histone modification recovery, leading to loss of heterochromatin. This creates a more permissive local environment for G4 formation in subsequent generations. As a result, G4s gradually accumulate on promoters throughout mitosis, driving clock-like DNA hypomethylation and chromatin opening. In patients and *in vitro* models, loss-of-function mutations in the G4-resolving enzymes WRN, BLM and ERCC8 accelerate the erosion of the epigenomic landscape around G4. G4-driven epigenomic aging is strongly correlated with biological age and is conserved in yeast, nematodes, insects, fish, rodents, and humans. Our results revealed a universal molecular mechanism of aging and provided mechanistic insight into how G-quadruplex processor mutations drive premature aging.

Keywords: Aging; epigenetics; DNA G-quadruplex (G4); mitosis; progeroid.

Introduction

Aging is a fundamental property of life. In his seminal study, Hayflick[1] co-cultured cells of different *in vitro* passage numbers. By tracking their clonal progenies by karyotype counting, Hayflick showed that aged cells produce fewer progenies and are eventually outcompeted from the population. Quantitative analysis of these experiments led to his proposal of the so-called Hayflick limit, namely, that every cell could undergo only finite numbers of mitoses. Since then, modern studies have identified multiple characteristic molecular hallmarks of aged cells[2]. The cell-autonomous features of aging include genomic instability[3,4], telomeric attrition[5], epigenomic landscape erosion[6-11], transcriptional changes[12-14], proteostasis imbalance[15-17], and metabolic switches[18]. However, the exact molecular mechanism driving aging across mitosis leading to the Hayflick limit has not been determined.

While some of the molecular processes that occur during aging, such as DNA damage, are irreversible, aged somatic cells can be "rejuvenated" back to a young ground state by nuclear transfer[19,20], indicating that the genetic composition of aged cells is insufficient for determining the aged state. Forced expression of embryonic stem cell-specific transcription factors (TFs) in somatic cells reprograms them back to an induced pluripotent state[21], not only reversing their differentiation state but also completely resetting their "epigenetic age" [7,8], resulting in an embryonic state that could rederive a young individual animal[21]. These results indicate that age is a property that is carried by non-genetic components in the cell. In agreement with this observation, chemical compounds inducing epigenetic changes could induce pluripotent reprogramming from somatic cells[22-24]. Conversely, programmed attrition of epigenetic modifier enzymes by genetic

54   manipulation results in accelerated aging of animals[25]. Hence, aging is likely to arise from gradual

55   erosion of the established epigenetic landscape.

56

57   Research on human progeroid syndromes (premature aging) has provided insight into the genetic

58   mechanism driving aging. Progeroid syndromes are caused by mutations in three classes of genes.

59   The first class of genes underlying syndromes, such as Hutchinson-Gilford progeria syndrome

60   (LMNA[26-29]) and Marbach-Rustad progeroid syndrome (LEMD2[30]), encode proteins located in the

61   nuclear membrane that are essential for heterochromatin anchoring. Mutations in these genes result

62   in non-specific disruption of heterochromatin structure. The second class of genes, which cause

63   diseases such as Fontaine progeroid syndrome, are essential for maintaining metabolic activity and

64   proliferation, such as the mitochondria phosphate carrier SLC25A24[31,32].

65

66   The last group of genes causes progeroid disease and includes genes associate with XFE syndrome

67   (ERCC4[33,34]), RECON progeroid syndrome (RECQL[35]), Werner syndrome (WRN/RECQL2[4,36-

68   39]), Bloom syndrome (BLM/RECQL3[40]), and Cockayne syndrome (CSB/ERCC6[41] and

69   CSA/ERCC8[34,42]). Although all of these genes function in the DNA repair machinery, they have

70   an additional common feature: they are responsible for the resolution of non-B DNA structures,

71   particularly G-quadruplexes (G4s) [43-46]. G4s are single-stranded DNA structures that can

72   spontaneously form from single-stranded DNA in open chromatin[47-53]. It is particularly enriched

73   at transcription start sites (TSSs) [54,55] and at DNA replication origins[54,56,57] and might serve as a

74   core promoter for both transcription and DNA replication. The G4 loci are enriched in structural

75   variation hotspots in cancer[58-61], suggesting that these loci are particularly vulnerable to DNA

76   breaks. This effect of G4 on DNA damage might be attributed to its ability to induce transcription-

4

77   replication interactions and inhibit DNA replication[45,55,59,61-66]. Furthermore, the formation of G4s

78   regulates chromatin accessibility[55,56,67] and DNA methylation[68]. In addition to its function in

79   promoting DNA damage, how G4s are linked to aging has not been determined.

80

81   Progeroid syndromes resulting from mutations in G4 resolving enzymes, especially ERCC6 and

82   ERCC8, are phenotypically very severe and onset at an early age. Fibroblasts carrying G4

83   resolving enzyme mutation show accelerated decline of replicative potential *in vitro*[69], which could

84   not be solely explained by telomere shortening[70] or increased DNA damage[71]. Aged cells from

85   progeroid syndrome patients carrying WRN[72], BLM[73], or ERCC6[74-76] mutations could still be

86   "rejuvenated" back to iPSC. These data suggest that G4 may mediate non-genetic mechanisms

87   underlying aging.

88

Results

Clock-like formation of G4 foci during cell aging

We analyzed the cellular content of G4s during aging in *in vitro* passaged primary mouse fibroblasts (pMEFs) using scFv-G4-antibody (BG4)-mediated immunofluorescence (Fig. 1a). As previously reported[48], BG4 stained clustered G4 spots within the nuclei (Fig. 1d). We found that the number of G4 spots per nucleus was significantly elevated during cell passage (Fig. 1d-e). We then performed a quantitative BG4 CUT&Tag[51] assay followed by qPCR and sequencing to measure G4 across the genome. We found that G4s tend to gradually accumulate in accessible chromatin (Extended Data Fig. 1), and the fraction of G4 reads within the aging-associated ATAC peak, determined as the ATAC peaks that gradually open during aging[7], linearly increases as the cell number increases (Fig. 1b). Furthermore, the coefficient of variation of the G4 signal increased with cell age (Fig. 1c), indicating that the heterogeneity of G4 decreased during aging. These results indicate that G4s accumulate at specific genomic loci during aging.

G4 accumulation drives clock-like chromatin opening

G4s can only form on opened chromatin. Using chromatin accessibility sequencing (ATAC-seq), we found that at certain genomic loci, chromatin accessibility gradually increases during *in vitro* passage through pMEFs. Chromatin opening during aging is much stronger at TSS of genes harboring a putative G-quadruplex sequence (pGQS) in their promoters (hereinafter termed "pGQSpos genes" for simplicity) than in those without pGQS in their promoters (hereinafter termed "pGQSneg genes") (Fig. 2a). We calculated aging-associated chromatin opening kinetics as the Z-scaled correlation coefficient between ATAC peak reads and sample age[7] (hereinafter

termed the "aging coefficient" for simplicity). The genes that showed a significant correlation between TSS chromatin accessibility and sample age (defined as those with an aging coefficient >= 3) were termed "aging genes", which showed a significantly greater gain of chromatin opening during aging (Fig. 2a). Interestingly, such aging-associated chromatin opening is associated with concurrent gain of G4 signals at the same promoters (Fig. 2b).

To test whether chromatin openings in G4 regions exhibit clock-like behavior, we used EpiTrace[7] to estimate the total number of opened G4 loci in bulk or single-cell ATAC-seq data from *in vitro* passaged pMEFs or primary mouse astrocytes (pASTs). A linear correlation between the total number of opened G4 loci and cell age was found for different cell types across the different assays (Fig. 2c-e). Furthermore, the number of total opened G4 regions in the ATAC-derived sample correlated well with the DNAm-derived biological age of the pMEFs, regardless of the cell cycle phase (Fig. 2f).

The tendency of a particular locus to form a G4 is determined by the energy it releases during folding. We computationally inferred the minimal folding energy (MFE) of each putative G4 sequence by G4Boost[77,78]. In MEFs, the mean MFE of each aging locus correlated with the aging coefficient (Fig. 2g). In an orthogonal analysis, we classified the mouse aging loci by the number of stable (G3)/G repeats they harbored (Extended Data Fig. 2) and found that sites containing >3 (G3)/G repeats had a greater aging coefficient (Fig. 2h). Similarly, in the human genome, scATAC peaks containing (G3)/(G) repeats are thermodynamically more stable than are those containing (G3)/(A/C/T) repeats (Fig. 2i) and have a greater aging coefficient in two scATAC experiments (Fig. 2j-k). Hence, chromatin opening during aging is closely associated with G4 formation.

134

135 If G4 formation is necessary for clock-like chromatin opening, G4s should be enriched at aging

136 loci compared to non-aging loci. To test this hypothesis, we computed loci-wise aging coefficients

137 from mouse, nematode[79], zebrafish[80], Drosophila[81] and human[22] scATAC data and inferred the

138 enriched DNA motives in the aging loci. The sequence motifs most strongly enriched in aging loci

139 were generally G-rich sequences, many of which could form G4 structures (Fig. 3a, e, and

140 Extended Data Fig. 3). In accordance with these findings, the pGQSpos genes had a much greater

141 aging coefficient in their TSS than the pGQSneg genes did (Fig. 3a, e, and Extended Data Fig. 4).

142 Furthermore, comparing to pGQSneg loci, the aging coefficient is increased in pGQSpos loci, and

143 further increased in pGQS loci positive with G4 ChIP-seq signal (Fig. 3b, f) [49,51,82]. Moreover, the

144 pGQSneg locus harboring clock-like differentially methylated loci[8] did not differ from the other

145 pGQSneg loci (Fig. 3b, f). The presence of a G4 ChIP-seq signal correlated with the percentage of

146 opened loci (Fig. 3c, g). These results indicate that G4 formation drives aging-associated, clock-

147 like chromatin opening.

148

149 To test the universality of G4 locus chromatin opening during aging, we estimated the total number

150 of opened G4 loci for accurately timed samples from mouse pMEF *in vitro* passages, from fly

151 embryonic development[81], from zebrafish neural crest development[80], from *C. elegans*

152 development[79], from yeast experimental aging[83], or from human hematopoietic cell development[84]

153 using EpiTrace. In all these cases, the estimated total number of opened G4 loci linearly correlated

154 with known sample age (Fig. 3d, i) or known rank of cell development (Fig. 3h) with high accuracy.

155 These data indicate that gradual opening of G4 loci is a common feature of aging throughout

156 evolution.

157

Pioneer factor primes G4 formation on nucleosome-depleted regions of promoters

159 To understand the molecular mechanism underlying the aging-associated accumulation of G4s and

160 associated chromatin opening, we first searched for potential molecular initiators of G4 formation

161 via bioinformatics. We performed unbiased statistical enrichment for published TF ChIP-seq data

162 in pGQS[57] in the mouse and human genomes[85]. As reported, many TFs are strongly enriched in

163 pGQS[86]. Interestingly, the most strongly pGQS-enriched TFs in mice (Extended Data Fig. 5a) or

164 humans (Extended Data Fig. 5b) are pioneer factors that bind to tightly packed, nucleosome-

165 wrapped DNA and open chromatin[35]. In mice, three out of four "Yamanaka factors" for

166 pluripotency induction[21] and two auxiliary pluripotency TFs (Ell3[87] and Prdm14[88]) composed 71%

167 (5/7) of the selected top pGQS-enriched (top 10% q-value) TFs (Extended Data Fig. 5a). In humans,

168 the known pioneer factors FOXA2, GATA2 and SPI1/PU.1 and ZBTB7A, the homolog of the

169 *Drosophila* pioneer factor GAF[89], are among the top 10 pGQS-enriched TFs (Extended Data Fig.

170 5b).

171

172 We tested the interdependence of pioneer factor binding and G4 formation on pGQS by analyzing

173 ATAC-seq and RNA-seq data from mouse cells with genetic perturbation of the pioneer factors

174 Oct4, Nanog, Sox2[90], and Zbtb7a[91]. ChIP-seq of these TFs revealed the expected enrichment of

175 the pGQSpos gene TSS compared to the pGQSneg gene TSS (Extended Data Fig. 6-8). On the

176 other hand, BG4 CUT&Tag signals are enriched on TSSs with pioneering factor binding sites

177 (Extended Data Fig. 8). Although knocking down Oct4, Sox2, or Nanog does not affect chromatin

178 accessibility at G4 loci (Extended Data Fig. 6), Zbtb7a knockout results in a reduction in chromatin

179 accessibility at pGQSpos (Extended Data Fig. 8). Similarly, Zbtb7a knockout affects the opening

180 of the pGQSpos and pGQSneg TSSs and decreases the 5' RNA expression of their bound targets,

181 regardless of whether they harbor a pGQS in the TSS (Extended Data Fig. 8), indicating decreased

182 occupation of paused RNA polymerase II at these sites. Hence, the pioneer factor Zbtb7a binds to

183 pGQS prior to G4 formation to create a nucleosome-depleted region (NDR), which enables G4

184 formation.

185

186 **G4 stimulates transcription-replication interactions to delay genome replication**

187 The pGQSpos TSS, particularly the pGQSpos TSS of aging genes, is enriched with RNA

188 polymerase II Ser5 phosphorylation and NELF, two essential drivers of transcriptional

189 elongation[92-94] (Fig. 4a and Extended Data Fig. 9). DNA replication origin activity[95] was also

190 enriched on the pGQSpos TSS compared to the pGQSneg TSS (Fig. 4e-f). The application of G4-

191 binding small molecule, which potentially impairs the recruitment of proteins to G4[93,94], negatively

192 affects both transcriptional elongation, as shown by nascent RNA sequencing (Fig. 4b), and DNA

193 replication on the pGQSpos TSS (Fig. 4e), indicating that G4 formation stimulates both

194 transcriptional elongation and DNA replication in its vicinity.

195

196 Strong transcriptional activity coupled with DNA replication activity could enhance transcription-

197 replication interaction (TRI), stimulating the formation of DNA-RNA hybrids that might hinder

198 DNA replication progression[59,63,64,96]. We found that elongating RNA polymerase II-bound newly

199 synthesized DNA[97] is significantly enriched on the pGQSpos TSS (Fig. 4c-d). This enrichment is

200 even more pronounced at the TSS of aging genes associated with pGQSpos (Fig. 4c-d). As a result,

201 significantly more transcription-replication interaction sites were found for the pGQSpos genes

202 (Fig. 4g). In accordance with these results, aging coefficients across the genome are associated

10

203   with the TRI tendency[98] (Extended Data Fig. 10). These data indicate that G4 formation stimulates

204   TRI by simultaneously promoting both transcription elongation and DNA replication.

205

206   The transcription-replication interaction presents a roadblock for DNA replication, which must be

207   resolved before replication resumes. G4-mediated blockade of DNA replication reportedly delays

208   DNA replication in chicken REV-1-deficient cells[99]. To test whether G4-mediated blockade of

209   DNA replication is universal, we analyzed genome replication dynamics in four different mouse

210   and human cell lines with high-resolution Repli-seq[100] data (Fig. 4h-i). In all the cell lines,

211   comparison of the G4 ChIP-seq-negative pGQS loci to the G4 ChIP-seq-positive pGQS loci

212   revealed that the presence of G4 strongly delayed DNA synthesis (Fig. 4h-i). In non-pGQS loci,

213   age-associated chromatin opening is also associated with delayed DNA synthesis (Fig. 4h-i). Taken

214   together, these data indicate that delayed genome replication is a general feature of aging loci and

215   that G4 stimulates local TRI to delay genome replication (Fig. 4j).

216

217   **G4s impair DNA methylation transmission during aging**

218   Delayed genome replication shortens the available time for histone modification recovery and

219   hinders parental histone recycling to the newly synthesized daughter chromosome. Therefore, we

220   hypothesize that impaired epigenomic transmission from the parental to daughter genome results

221   from delayed genome replication at G4 loci.

222

223   We first tested whether the transmission of DNA methylation from the parental to daughter genome

224   is affected by G4. To analyze DNA methylation transmission, we analyzed DNA methylation

225   maintenance between parental and daughter DNA strands in human HeLa cell Hammer-seq data[101].

226   As a positive control of the analysis, DNA methylation maintenance was impaired at the solo-

227   WCGW site compared to the non-WCGW and non-pGQS control CpG sites[101] (Fig. 5a). Although

228   the non-G4 pGQS sites did not differ from the control sites, the G4 sites exhibited significantly

229   delayed DNA methylation maintenance in both the early and late S phases (Fig. 5a). This

230   phenomenon was validated by analyzing another Repli-BS-seq dataset[102] (Extended Data Fig. 11).

231   The difference in DNA methylation maintenance efficiency between the G4 locus and the non-G4

232   pGQS locus diminished upon double knockout of DNMT3A/3B, indicating that G4 formation

233   hinders DNMT3A/3B-mediated DNA methylation (Extended Data Fig. 11).

234

235   To validate whether DNA methylation transmission is impaired by G4 formation during aging, we

236   analyzed DNA methylation levels around G4 loci in mouse and human cells by targeted bisulfite

237   sequencing with anchored PCR (G4AMP or MArchPCR). In mouse and human primary cells, CpG

238   in the vicinity of G4 could undergo hypomethylation during aging (Extended Data Fig. 12a, b, e,

239   h).

240

241   Since not all CpGs in the vicinity of G4 are similarly hypomethylated during aging and since the

242   thermostability of G4 is correlated with age-associated chromatin opening (Fig. 2g-k), we

243   hypothesized that G4 stability might also regulate age-associated DNA hypomethylation. Since

244   G4 is a hub for TF binding[55,86], we tested whether TF activity could regulate G4-associated DNA

245   hypomethylation. We inferred TF activity during pMEF *in vitro* passage using SHARE-seq

246   data[103,104] and determined three TFs (Zfp148, Zfp263, and Klf15) that significantly associated with

247   age (Extended Data Fig. 12c). We then analyzed the DNA methylation level of the same sample

248   CpGs in the vicinity of G4 bound by these TFs. Interestingly, the CpGs falling in these regions

249   showed consistent age-associated DNA hypomethylation (Extended Data Fig. 12d). Hence, G4

250   DNA hypomethylation during aging is regulated by TF activity.

251

252   We then classified G4 regions according to the presence of TF-binding sites (TFBS) and separately

253   analyzed DNA methylation in each group. In both mouse and human cells, most G4-bound TFs

254   are associated with DNA hypomethylation during aging (Extended Data Fig. 12f, g, i, j). Overall,

255   in the vicinity of G4 bound by these TFs, aging was associated with a strong and linear decrease

256   in the DNA methylation level (Fig. 5c-d). Taken together, these results indicate that G4 impairs

257   DNA methylation transmission during aging.

258

259   **G4 impairs histone modification transmission during aging**

260   To analyze the effect of G4 on histone modification transmission, we first analyzed histone

261   modification recovery using mouse ESC H3K27me3 CUT-and-flow data[105] (personal

262   communication). We calculated the frequency of gene-associated H3K27me3 modifications by

263   averaging H3K27me3 CUT&Tag signals from the flanking +/-50 kb regions around the gene. At

264   each locus, the H3K27me3 signal valley across the cell cycle indicates the completion of histone

265   exchange, and its peak indicates the completion of recovery of the parental H3K27me3

266   modification. In general, the pGQSpos gene is more likely to undergo complete histone exchange

267   in the early S phase (Fig. 5b and Extended Data Fig. 13a) and requires slightly more time to recover

268   H3K27me3 modification (Extended Data Fig. 13b-c). As a result, the mean H3K27me3 recovery

269   in the pGQSpos genes was slower than that in the pGQSneg genes (Fig. 5e).

270

271   We validated whether H3K27me3 at G4 loci is reduced during cell aging using single-cell

272   H3K27me3 ChIP-seq (scChIP) data derived from a human breast cancer (BRCA) PDX model[106].

273   In this study, BRCA cells (HBx95) were grafted into mice and subjected to chemotherapy

274   (capecitabine). Clustering of the scChIP profiles of cancer cells from different samples revealed a

275   common ancestor cell population that existed in both the drug-treated and control samples

276   (Extended Data Fig. 14a-c) and two derived cell populations corresponding to either the control,

277   non-drug-resistant, or drug-resistant state (Extended Data Fig. 14c). Compared to those of the

278   ancestor population, the H3K27me3 coverage of the G4 locus decreased (Fig. 5f) regardless of the

279   drug resistance status, and the H3K27me3 coverage of the G4 locus further decreased (Fig. 5f and

280   Extended Data Fig. 14d). These results indicate that G4 impairs histone modification transmission

281   in its vicinity.

282

283   **Aging-associated chromatin opening results from impaired epigenome transmission**

284   Impaired epigenome transmission by G4s during each cell cycle results in gradual DNA

285   hypomethylation and loss of heterochromatin (Fig. 5a-f). This in turn would create a more

286   permissive local environment for G4 formation in subsequent generations. This self-reinforcing

287   cycle of G4 formation not only explains the gradual accumulation of G4s during aging (Fig. 1) but

288   also suggests concurrent clock-like epigenomic changes in the neighboring G4 genomic region.

289

290   Aging peaks were more enriched at the pGQSpos locus than at the pGQSneg locus (Fig. 5g).

291   Analysis of DNA replication and histone modification recovery dynamics revealed that in non-

292   aging sites, DNA replication is completed in the early S-phase, allowing long recovery times for

293   histone modification (H3K27me3; Fig. 5h). This finding is in stark contrast to what has been

294 observed in aging sites, where DNA replication does not complete until late S-phase, and histone

295 modification recovery is delayed until the next cell cycle (Fig. 5h). We then clustered the pGQSpos

296 genes based on their DNA replication and H3K27me3 recovery dynamics using K-means

297 clustering on time series[107] (Fig. 5i) and found a similar scenario: aging sites were depleted from

298 the early replicating clusters and enriched in the late replicating clusters that could not finish

299 H3K27me3 recovery within the cell cycle (Fig. 5j). Taken together, these results indicate that G4

300 impairs epigenome transmission across mitosis to gradually erode the local epigenetic landscape

301 (Fig. 5k).

302

303 **Genetic mutations causing G4 accumulation accelerate the erosion of the epigenetic**

304 **landscape around G4 *in vivo***

305 Mutations in G4-resolving enzymes such as BLM, WRN, ERCC6 and the E3 ligase ERCC8 result

306 in premature aging in human patients. Previously, it was reported that mutations in WRN result in

307 global loss of H3K9me3 in human mesenchymal stem cells (hMSCs) [108], but there is no clear link

308 between biochemical function and phenotype. We reanalyzed the H3K9me3, H3K4me3 and

309 H3K27me3 ChIP-seq data from this study. Strikingly, the WRN mutation resulted in a much

310 greater reduction in H3K9me3 at putative G4 loci than at non-pGQS loci (Fig. 6a). Although the

311 WRN mutation induces only a minor and local reduction in H3K9me3 at non-pGQS loci, it induces

312 a potent reduction in H3K9me3 at pGQS loci (Fig. 6a). This effect is not limited to the local

313 vicinity of pGQS and can extend to distal sites as far as 50 kb away (Fig. 6a). Furthermore, the

314 WRN mutation induces a gain of H3K4me3 specifically at the pGQS locus (Fig. 6b) and decreases

315 H3K27me3 peaks globally (Fig. 6c).

316

317   We analyzed G4-associated DNA methylation by sequencing PBMCs from a group of human

318   progeroid patients with mutations in BLM or ERCC8. Significant aging-associated G4 DNA

319   hypomethylation was detected in the progeroid patients compared to the age-matched non-

320   progeroid donors (Fig. 6d). Furthermore, segregating the ERCC8 patients into groups according

321   to the molecular mutations they carried (Extended Data Fig. 15a) revealed that the aging-

322   associated G4 DNA hypomethylation level was correlated with mutation severity: the DNA

323   hypomethylation phenotype was least severe in patients with compound heterozygous ERCC8

324   mutations with a hypomorphic short nucleotide variation mutation over a structural variation[109]

325   and most severe in a patient with structural variation over a genetically null large chromosomal

326   deletion (Fig. 6e). Compared to that in non-progeroid donors, DNA hypomethylation around G4

327   in these patients accelerated to such an extent that a 10-year-old patient had a DNA methylation

328   level way below that of a 94-year-old non-progeroid donor (Fig. 6f). Linear regression shows that

329   in these patients, the aging-associated G4 DNA hypomethylation was accelerated by a factor of 5

330   compared to controls, which agrees well with the usual life expectancy of 10~20 years of Cockayne

331   and Bloom syndrome patients (Fig. 6f). Analysis of the full pGQS DNA methylation level revealed

332   similar findings (Extended Data Fig. 15b-d). To summarize, these data indicate that pathogenic

333   mutations causing G4 accumulation accelerate the progression of the epigenetic landscape around

334   G4 *in vivo*.

335

Discussion

Aging is a fundamental and universal phenomenon of life. Here, we showed that G4 formation drives aging in eukaryotes by promoting the erosion of the epigenetic landscape (Fig. 7). Pioneer factor primes G4 formation on nucleosome-depleted regions of the transcription start site. Once G4s are formed, they stimulate transcription-replication interactions, which delay genome replication. Delayed genome replication results in impaired epigenome transmission. As a result, G4 triggers the gradual erosion of its neighboring epigenetic landscape.

It is natural to presume that G4s form in open chromatin spontaneously. However, the accumulation of G4s across mitosis is not uniform across the genome and might differ between cell types. Through bioinformatic analysis, we determined that this accumulation is associated with G4 thermostability and might be modulated by protein binding. Identifying the exact molecular mechanism of aging-associated, cell type-specific G4 accumulation is of particular interest for understanding aging-related diseases affecting specific cell types, such as arteriosclerosis, Alzheimer's disease, and amyotrophic lateral sclerosis.

Immunostaining of DNA G4s using a single-chain antibody revealed spot-like G4s in the nucleus[48]. However, G4-associated epigenetic changes, including H3K9me3 erosion and DNA hypomethylation in heterochromatin domains, suggest that these events initially occur near the nuclear lamina. Human mutations in nuclear lamina proteins result in premature aging and inactivation of normally silenced LINE1 retrotransposons in heterochromatin domains[110]. In *Drosophila*, GAF, the *Drosophila* homolog of Zbtb7a[89], binds to a GAGA-rich sequence that is enriched according to ChIP-seq data of the *Drosophila* nuclear pore basket protein Elys[111].

17

359   Interestingly, our results also revealed the GAGA-rich sequence as the most strongly enriched

360   motif at loci involved in aging-associated chromatin opening in the *Drosophila* genome.

361   Furthermore, the pathogenic RNA G4 structure in C9orf72 causes neurodegeneration through

362   blocking nuclear pore trafficking[112,113], indicating an intimate link between the nucleic acid G4

363   structure and the periphery of the nucleus. Although the exact physical location of initial G4

364   formation during aging is currently unknown, we consider this a potentially interesting research

365   direction.

366

367   G4s exist in multiple forms. In particular, pathogenic mutations in progeroid disease might affect

368   different types of G4s: while WRN and BLM mostly act on intrastrand G4s[44], ERCC6 specifically

369   binds to interstrand G4s and might mediate their resolution[43]. Interestingly, WRN-knockout

370   hMSCs exhibit minimal differences in DNA methylation compared with wild-type hMSCs[108],

371   while large shifts in H3K9me3 and H3K4me3 occur around G4 loci. On the other hand, our results

372   showed that compared with those from controls, PBMCs from ERCC8-mutated patients exhibit

373   prominent G4 DNA hypomethylation. It would be interesting to test whether different types of G4s

374   cause different types of epigenetic inheritance deficiency.

375

376   In the mammalian system, inducing constitutive faithful DNA repair alone causes global

377   epigenomic information attrition and organismal aging[25]. G4 can promote replication-dependent

378   genome instability[65] and is enriched in DNA breakpoints in cancer[58]. Since all progeroid-

379   pathogenic G4-resolving helicases function in DNA repair, G4 accumulation may promote aging

380   via a DNA-break-repair mechanism. Although our study does not exclude this possibility, the

381   localized effect of G4 on H3K9me3 and DNA methylation loss in WRN and ERCC8 mutants

382   suggested that replication-dependent epigenomic transmission error is a more likely explanation

383   for these results.

384

385   Nonetheless, our study is not without limitations. The application of G4-binding small molecules

386   might affect both the stability of G4 and the binding of G4 to its interaction partners[93,94]; hence,

387   G4 could be either antagonistic or agonistic in nature. Therefore, the results derived from G4-

388   binding small molecule manipulations could have alternative explanations. Since there are no

389   straightforward methods for affecting G4 stability without simultaneously perturbing its binding

390   with partners, such limitations seem unavoidable. Finally, our work focused on the G4 formation

391   cycle throughout mitosis. However, aging also occurs in cells that do not actively divide, such as

392   neurons. The chromatin accessibility of neurons shifts during aging[114], suggesting that erosion of

393   the epigenetic landscape might also underlie aging in these cells. However, whether G4s might

394   actively regulate aging in these non-dividing cells, and if so, what the underlying molecular

395   mechanism is, is an open question awaiting future investigation.

396

397   In conclusion, our results revealed a highly conserved molecular mechanism driving eukaryotic

398   aging. G4-resolving enzymes were previously proposed to be "mitotic counters" of a cell[69] since

399   cells harboring their mutations have a shortened Hayflick limit. Through our analysis, we

400   established a molecular link to explain how pathogenic mutations in G4-resolving enzymes

401   accelerate mitosis to cause human progeroid disease. While analyzing G4-associated epigenetic

402   changes could serve as a measure of biological age, perturbing G4 formation might be of particular

403   interest for modulating natural and pathological aging.

404

Methods

Human donors, animal samples and cells

This study was conducted in accordance with local legislation, measures of the Declaration of Helsinki, and the Ethics Protocols of the Human Genetic Resource Preservation Center of Hubei Province, China (Hubei Biobank). The human samples used in this study were pretreated and preserved by the Hubei Biobank, approved by the Institutional Ethical Review Board (approval number: 2017038-1). Informed consent was obtained from all individuals. For the mouse experiments, 33 C57BL/6 mice were used for tail fibroblast collection, 5 pregnant C57BL/6 mice were used for primary embryonic fibroblast collection, and 3 dams of C57BL/6 neonatal mice were used for primary astrocyte collection. The animal experiments were carried out at Zhongnan Hospital of Wuhan University and were approved by the Institutional Ethical Review Board (approval number: ZN2023148).

Datasets used in this study

A full list of public datasets used in this study is shown in Extended Data Table 1.

Genotyping of progeroid patients

The patients were initially subjected to whole-exome sequencing, and the candidate pathogenic mutations were confirmed by Sanger sequencing. The ERCC8 patients were further genotyped as described in Wang *et al*., 2017[115]. The ERCC8 exon 4 complex rearrangement was sequenced as described by Ren *et al*., 2003[116]. A full list of the genotypes of these patients is shown in Extended Data Table 2.

Cell culture

Human bone marrow-derived mesenchymal stem cells (BMSCs) were obtained from the Stem Cell Bank, Chinese Academy of Sciences. Human BMSCs were cultured in MSC NutriStem® XF Medium (Biological Industries, 05-200-1A) supplemented with MSC NutriStem® XF Supplement Mix (Biological Industries, 05-201-1U) and 1% penicillin and streptomycin. The cell handling procedures were conducted according to the instruction manual. Primary mouse embryonic fibroblasts (MEFs) were isolated and maintained following a previously described protocol[117]. The FACS-sorting and SHARE-seq sequencing methods used for pMEFs were described in a previous study[7].

BG4 antibody expression and purification

pSANG10-3F-BG4[48] (Addgene #55756) was subsequently transformed into BL21 cells. A single clone of the transformant was precultured in 3 mL of 2xTY + 2% glucose medium at 30°C overnight and subsequently transferred to 1 L of ZYM-5052 autoinduction medium[118] for culture at 25°C overnight. Bacteria were collected by centrifugation. Osmotic extraction was performed by resuspending the bacteria in 20 mL of TES[119] with benzonase (Millipore #E1014) on ice for 1 hr, followed by centrifugation at 10,000xg for 10 mins at 4°C. Supernatants were collected and diluted 1:10 into PBS with 10 mM NaCl, pH 8.0, passed through 0.22 µm filter, captured on column with 1 mL cobalt resin (Takara #635504), washed by 15 column volumes of PBS with 100 mM imidazole and 100 mM NaCl, pH 8.0, then eluted by 3 column volumes of PBS with 200 mM imidazole and 100 mM NaCl, pH 8.0. The collected proteins were concentrated by an Amicon Ultra-3k centrifugal filter (Millipore), exchanged with PBS supplemented with 100 mM NaCl and 50% glycerol, and stored at -80°C.

BG4 immunostaining and automatic quantification

BG4 immunostaining was performed as previously described[48]. Briefly, live MEFs at different passages were grown on glass coverslips, fixed in methanol:acetic acid (3:1) and permeabilized with 0.1% Triton-X100/PBS. After blocking in 2% BSA/PBS, immunofluorescence was performed using standard methods with BG4 (1:50), anti-FLAG (Sigma, F1804, 1:100) and Goat anti-Mouse, Alexa Fluor 594 (Invitrogen, A-11032, 1:1000) antibodies. Coverslips were mounted with Prolong Gold/DAPI (Invitrogen). Quantification was performed in R (4.1.3) using the package EBImage (4.42.0) programatically, without human intervention, to select for ROIs or adjust intensities.

Chromatin accessibility profiling (ATAC-seq) and CUT&Tag sequencing

CUT&Tag and ATAC-seq experiments were performed according to the protocol used in Xiao *et al*., 2022[120]. For the quantitative ATAC-seq experiments, 20,000 human HEK293T cells (kindly provided by Dr. Qi Li, National Institute of Biological Sciences, Beijing, China) were mixed with 50,000 mouse cells prior to the experiment. For quantitative CUT&Tag experiments, 40,000 human HEK293T cells were mixed with 70,000 mouse cells prior to the experiment. The antibodies used in the CUT&Tag experiments were BG4 (homemade), mouse anti-FLAG (Sigma, # F1804), and goat anti-mouse IgG (Sangon, #D111024-0001). Fragmentation after antibody binding was performed with a NovoProtein CUT&Tag 2.0 pAG-Tn5 kit (NovoProtein Ltd., Cat. #N259).

qPCR of G4 enrichment in the opened region

474   CUT&Tag NGS libraries were amplified with (1) a G-quadruplex-specific common primer (NGS-

475   1290, 5'-GGAGGGGAGGGGAGAGGGG-3') and the Illumina P7 adaptor common primer or (2)

476   a pair of Illumina P5/P7 common primers as controls for DNA content in qPCR mix (VAHTS

477   SYBR qPCR master mix (Vazyme, #NQ101) on a Quantagene q225 qPCR machine (Kubo

478   Technologies, Beijing, China). Amplification curves were analyzed by the standard ddCt method.

479

Anchored PCR bisulfite sequencing

481   The multiplex anchored PCR bisulfite sequencing data of mouse primary embryonic fibroblasts

482   generated using the MM285/347 ClockDML primer set were obtained from a previous publication[7].

483   For G4-specific DNA methylation sequencing (G4AMP) of human or mouse samples, genomic

484   DNA was extracted from the proteinase K-digested tail tips of mouse or peripheral blood

485   mononuclear cell (PBMC) of human donors using DNeasy blood tissue kit (Qiagen, #69506). 25

486   ng of genomic DNA was subjected to bisulfite conversion using a commercial kit (Zymo Research,

487   #D5005). The converted product was 3' tailed by dATP with terminal deoxynucleotide transferase

488   (Enzymatics, #P7070L) in the presence of a tail-controlling adaptor, as described for the Tequila 7

489   N bisulfite capture sequencing assay[121], and ligated to the adaptor with *E. coli* DNA ligase (Takara,

490   #2161). The ligated product was amplified with a G-quadruplex-specific common primer (NGS-

491   1290, 5'-GGAGGGGAGGGGAGAGGGG-3') and an adaptor-specific common primer for 12

492   cycles with Phusion U (Thermo Fisher, #F555L). One-tenth of the amplification product was

493   subjected to another round of amplification using Illumina P5/P7 primers. The amplified product

494   was recovered by column purification and sequenced on an Illumina NovaSeq 6000 machine in

495   PE150 format to 5-10 M reads.

496

23

Sequencing data processing

The reference genomes used in this study were hg38 (human), hg19 (human: bisulfite sequencing only), mm10 (mouse, both routine and bisulfite sequencing), hg38-mm10 (human/mouse mixed cells for quantitative ATAC-seq/CUT&Tag), dm6 (*D. melanogaster*), sc3 (*S. cerevisiae)*, ce11 (*C. elegans*), and danRer10 (*D. rerio*). CUT&Tag, CUT&Flow, ChIP-seq, ATAC-seq, G4Access, MNase-seq, and SNS-seq data were aligned to the reference genome using bwa (0.7.17-r1188). RNA sequencing data, including total RNA-seq and TT-seq data, were aligned to the respective reference genomes using STAR (version 2.7.10a_alpha_220818). The bisulfite sequencing data were aligned to the reference genome using bwa-meth (0.2.0). CpG methylation levels were determined using Pile-O-Meth (v0.1.13-3-gca82747). Peaks in the DNA sequencing data were identified by MACS2 (2.2.7.1). The BigWig files were generated by bamCoverage (3.3.0). Coverage profiles were generated by computeMatrix (3.3.0). The mouse genome coverage of the quantitative ATAC-seq/CUT&Tag data was normalized against the reference (human genome) read content in the same library. Prior to calculating the CV in the qBG4 CUT&Tag data, to control for the influence of chromatin accessibility on BG4 CUT&Tag efficiency, the qCUT&Tag signal at a particular locus was divided by the qATAC signal at the same locus.

Enrichment analysis

Region enrichment between a given set of genomic loci and reference loci was performed by LOLA (1.32.0). Motif enrichment in each region was performed by rGADEM (2.44.1). Fisher's exact test was used for statistical comparisons of enrichment. The R package regioneR (1.28.0) was used for the two-region set overlap enrichment significance test.

Estimation of the total opened fraction of G4

Single-cell and bulk ATAC sequencing data were converted into peak-x-cell count tables. Estimation of the total opened fraction of G4 loci was performed with EpiTrace (version 055441a) using the function `EpiTraceAge_Convergence`. The pGQS loci derived from G4Hunter for the corresponding reference genome were used as the reference.

Estimation of the aging coefficient

The aging coefficient was calculated by the function `AssociationOfPeaksToAge` from EpiTrace. The EpiTrace-derived cell age or the known accurate timing of the sample (for the *C. elegans, S. cerevisiae, D. melanogaster, D. rerio,* and mouse pMEF data) was used for correlation against the ATAC coverage of the full peak set. Aging peaks were defined as those with a Z-scale aging coefficient >3. Aging genes were defined as genes that harbored an aging peak in their TSS. To calculate the age-related coefficient profile for a given set of genes, the positions of the ATAC-seq peaks for these genes were converted into relative positions (from the TSS to the TES: 0%-100%), and the smoothened age coefficient was calculated from the age-related coefficients of all the ATAC-seq peaks covering these genes via the `gam` method in R (4.1.3).

Distance between loci and the replication initiation zone

Processed OK-seq data were downloaded from https://github.com/CL-CHEN-Lab/OK-Seq[98]. Replication initiation zones (IZs) in mouse and human cells were taken from published results. The distance between the IZ and G4 loci was calculated by GenomicRanges (1.48.0).

Repli-seq data analysis

543　Processed high-resolution Repli-seq data were downloaded from NCBI (GSE137764). The

544　genomic segments were classified according to whether they overlapped (1) with a G4 ChIP-seq

545　peak, (2) with a pGQS locus, (3) with a non-pGQS aging locus, or (4) none of above. For classwise

546　analysis, Repli-seq signals from the 16 S-phase timepoints for all genomic segments in a particular

547　class were averaged, and the cumulative sum of Repli-seq signals across these time points was

548　calculated. For segmentwise analysis, Repli-seq signals from the 16 S-phase timepoints for each

549　genomic segment were used as a "replication curve". The curves were clustered by K-means with

550　the R package kml (v2.4.6.1).

551

552　**H3K27me3 CUT&Flow data analysis**

553　The H3K27me3 CUT&Flow signal covering each gene was computed by Deeptools

554　computeMatrix (3.3.0). For each gene, the signal was averaged for each sample and normalized

555　from lowest to highest across the samples. The pGQSneg and pGQSpos genes were separately

556　clustered according to the cell cycle phase, exhibiting the lowest normalized H3K27me3 signal

557　("valley"). The H3K27me3 signal of a gene was normalized using an "H3K27me3 modification

558　curve" and was clustered by K-means with kml. To perform joint clustering of the H3K27me3

559　profile and Repli-seq profiles, a DNA replication curve of the genomic segment containing the

560　gene was generated from the Repli-seq data and combined with the H3K27me3 curve. This

561　chimeric curve was then subjected to K-means clustering by kml.

562

563　**H3K27me3 scChIP-seq analysis**

564　A preprocessed count table of H3K27me3 scChIP-seq data was downloaded from NCBI

565　(GSE117309). The reference hg19 pGQS loci (G4Hunter score >=1.52) were downloaded from

26

566   NCBI (GSE187007) and lifted to hg38 using easyLift (0.2.1). To estimate total H3K27me3 peaks

567   across G4 loci for each single cell, EpiTrace was applied to the count matrix, with the lift-over

568   hg38 pGQS locus serving as a reference. Per-G4-locus H3K27me3 ChIP-seq coverage was

569   computed by dividing the total number of H3K27me3 reads in G4 regions by the total number of

570   G4 loci covered. Single-cell clusters were inferred from the total count matrix using Signac (1.6.0).

571   Ancestral clones were inferred to be the single cells in a cluster that were most evenly shared

572   between different derived populations.

573

Prediction of pGQS, G4L1, and G4 thermostability

574

575   G4L1 sequences were searched by an adapted R code following the regex search code provided

576   by Dr. José Arturo Londoño-Vallejo as used in Lombardi *et al*., 2019[122]. The putative G4 sequence

577   was searched for in reference genomes using G4Hunter[57]. To predict G4 thermostability, DNA

578   sequences were extracted for each genomic locus using Biostrings (2.64.0) to generate a FASTA

579   file input for G4Boost[77] (https://github.com/hbusra/G4Boost, version 0de3876).

580

G4 DNA methylation analysis via G4AMP and MArchPCR

581

582   Per-CpG site sequencing coverage and relative frequency of C reads (methylated C) from G4AMP

583   or MArchPCR sequencing results were taken from the Pile-O-meth output. For linear correlation

584   between sample age and CpG-DNA methylation level, a custom script adapted from scAge[123] was

585   used following the protocol described in a previous publication[7]. Maximum likelihood estimation

586   of sample age and determination of the 95% confidence intervals from the G4AMP data were

587   performed with the transplanted scAge[123] function in R based on the CpG loci, which showed a

588   correlation coefficient < -0.7 and p value < 0.01. To determine the TF-associated G4 DNA

589   methylation status, TF motifs were extracted from the JASPAR2022 database (0.99.7) and matched

590   to G4Hunter-predicted pGQS (GSE187007, lifted from mm9 to mm10) using the motifmatchr

591   (1.18.0) function `matchMotifs` with the parameters "cutoff=5e-05" and "width=7". G4 loci

592   matching a particular TF were collected. For each sample, a particular TF-associated DNA

593   methylation level was calculated as the total C reads covering CpG on TF-associated loci divided

594   by the total C+T reads covering CpG on the same set of loci. A linear correlation between TF-

595   associated DNA methylation and sample age was calculated in R (4.1.3). The same group of non-

596   progeroid donors is shown in Fig. 5c and Fig. 6f (for comparison to the progeroid group). The

597   "CTRL" donors in Fig. 6d were <=15 years old from the same non-progeroid group.

598

DNA methylation maintenance analysis via Hammer-seq and Repli-BS-seq

600   Processed stranded CpG methylation read counts of Hammer-seq data were downloaded from

601   NCBI (GSE131098). CpGs were classified according to whether they were within the vicinity of

602   100 bp of the pGQS site or G4 ChIP-seq site or if they corresponded to the solo-WCGW sites in

603   the partially methylated domain (PMD) [9]. The daughter strand was defined as the read strand with

604   less DNA methylation (C-reads). If two strands had similar DNA methylation levels, then a random

605   strand was chosen as the daughter strand. For each pair of parent-daughter strands, the parental

606   methylated loci were defined as "C" bases in the parent strand, and "C" bases in the daughter

607   strand were considered "maintained". The methylation maintenance level of a particular CpG class

608   in a sample was calculated as the mean probability of maintained methylation in daughter strands

609   for these CpGs.

610

611   To analyze methylation maintenance in Repli-BS-seq, processed bisulfite sequencing data were

612   downloaded from NCBI (GSE82045). CpGs were classified as described for the Hammer-seq

613   experiments. Since no strand pair information was available, the methylation maintenance level

614   was calculated as the methylation level normalized between timepoints (0% (min) - 100% (max)).

615

616   **Transcription factor activity inference from scATAC**

617   Transcription factor activity was computed by ChromVAR in ArchR (1.0.1) from the pMEF

618   SHARE-seq dataset from a previous publication[7]. Single-cell ages were estimated from the same

619   dataset using EpiTrace. A linear correlation between TF activity and single-cell age was calculated

620   in R (4.1.3).

Code and data availability

The codes used in this study were deposited on GitHub (link: https://github.com/MagpiePKU/G_quadruplex_aging_paper_2024). Raw mouse sequencing data and processed human sequencing data were deposited on the CNGB GSA.

Author contributions

YZ conceived the concept and led the study. WJ, JZ, YX and LJ performed the experiments. WJ, FC, JF and YZ performed the bioinformatic analysis. KQ and HJ collected the biospecimens. YX, HJ and YZ oversee the study. WJ, YX and YZ wrote the paper, with input from all the authors.

Acknowledgments

643   funders played no role in the study design, data collection and analysis, decision to publish, or

644   preparation of the manuscript.

645

646   **Competing interests**

647   The authors declare no competing interests.

648

References

1   Hayflick, L. The Limited in Vitro Lifetime of Human Diploid Cell Strains. *Exp Cell Res* **37**, 614-636 (1965).

2   Lopez-Otin, C., Blasco, M. A., Partridge, L., Serrano, M. & Kroemer, G. Hallmarks of aging: An expanding universe. *Cell* **186**, 243-278 (2023).

3   Huang, Z. *et al.* Single-cell analysis of somatic mutations in human bronchial epithelial cells in relation to aging and smoking. *Nature genetics* **54**, 492-498 (2022).

4   Sinclair, D. A. & Guarente, L. Extrachromosomal rDNA circles--a cause of aging in yeast. *Cell* **91**, 1033-1042 (1997).

5   Blackburn, E. H., Epel, E. S. & Lin, J. Human telomere biology: A contributory and interactive factor in aging, disease risks, and protection. *Science* **350**, 1193-1198 (2015).

6   Kaeberlein, M., McVey, M. & Guarente, L. The SIR2/3/4 complex and SIR2 alone promote longevity in Saccharomyces cerevisiae by two different mechanisms. *Genes Dev* **13**, 2570-2580 (1999).

7   Wang, X. *et al.* Tracking single cell evolution via clock-like chromatin accessibility. *bioRxiv*, 2022.2005.2012.491736 (2022).

8   Horvath, S. DNA methylation age of human tissues and cell types. *Genome biology* **14**, R115 (2013).

9   Zhou, W. *et al.* DNA methylation loss in late-replicating domains is linked to mitotic cell division. *Nature genetics* **50**, 591-602 (2018).

10   Lu, A. T. *et al.* Universal DNA methylation age across mammalian tissues. *Nat Aging* **3**, 1144-1166 (2023).

11   Kabacik, S. *et al.* The relationship between epigenetic age and the hallmarks of aging in human cells. *Nat Aging* **2**, 484-493 (2022).

12   Debes, C. *et al.* Ageing-associated changes in transcriptional elongation influence longevity. *Nature* **616**, 814-821 (2023).

13   Gyenis, A. *et al.* Genome-wide RNA polymerase stalling shapes the transcriptome during aging. *Nature genetics* **55**, 268-279 (2023).

14   Yokoyama, M., Sasaki, M. & Kobayashi, T. Spt4 promotes cellular senescence by activating non-coding RNA transcription in ribosomal RNA gene clusters. *Cell Rep* **42**, 111944 (2023).

15   Hipp, M. S., Kasturi, P. & Hartl, F. U. The proteostasis network and its decline in ageing. *Nat Rev Mol Cell Biol* **20**, 421-435 (2019).

16   Martinez-Miguel, V. E. *et al.* Increased fidelity of protein synthesis extends lifespan. *Cell metabolism* **33**, 2288-2300 e2212 (2021).

17   Shcherbakov, D. *et al.* Premature aging in mice with error-prone protein synthesis. *Sci Adv* **8**, eabl9051 (2022).

18   Slack, C. *et al.* The Ras-Erk-ETS-Signaling Pathway Is a Drug Target for Longevity. *Cell* **162**, 72-83 (2015).

19   Briggs, R. & King, T. J. Transplantation of Living Nuclei From Blastula Cells into Enucleated Frogs' Eggs. *Proceedings of the National Academy of Sciences of the United States of America* **38**, 455-463 (1952).

20   Gurdon, J. B., Byrne, J. A. & Simonsson, S. Nuclear reprogramming and stem cell creation. *Proceedings of the National Academy of Sciences of the United States of America* **100 Suppl 1**, 11819-11822 (2003).

694 21   Takahashi, K. & Yamanaka, S. Induction of pluripotent stem cells from mouse embryonic
695   and adult fibroblast cultures by defined factors. *Cell* **126**, 663-676 (2006).
696 22   Guan, J. *et al.* Chemical reprogramming of human somatic cells to pluripotent stem cells.
697   *Nature* **605**, 325-331 (2022).
698 23   Ye, J. *et al.* Pluripotent stem cells induced from mouse neural stem cells and small intestinal
699   epithelial cells by small molecule compounds. *Cell Res* **26**, 34-45 (2016).
700 24   Zhao, Y. *et al.* A XEN-like State Bridges Somatic Cells to Pluripotency during Chemical
701   Reprogramming. *Cell* **163**, 1678-1691 (2015).
702 25   Yang, J. H. *et al.* Loss of epigenetic information as a cause of mammalian aging. *Cell* **186**,
703   305-326 e327 (2023).
704 26   Cao, H. & Hegele, R. A. LMNA is mutated in Hutchinson-Gilford progeria (MIM 176670)
705   but not in Wiedemann-Rautenstrauch progeroid syndrome (MIM 264090). *J Hum Genet*
706   **48**, 271-274 (2003).
707 27   Capell, B. C. & Collins, F. S. Human laminopathies: nuclei gone genetically awry. *Nature*
708   *reviews. Genetics* **7**, 940-952 (2006).
709 28   Mounkes, L. C., Kozlov, S., Hernandez, L., Sullivan, T. & Stewart, C. L. A progeroid
710   syndrome in mice is caused by defects in A-type lamins. *Nature* **423**, 298-301 (2003).
711 29   Vigouroux, C., Caux, F., Capeau, J., Christin-Maitre, S. & Cohen, A. LMNA mutations in
712   atypical Werner's syndrome. *Lancet* **362**, 1585; author reply 1586 (2003).
713 30   Marbach, F. *et al.* The Discovery of a LEMD2-Associated Nuclear Envelopathy with Early
714   Progeroid Appearance Suggests Advanced Applications for AI-Driven Facial Phenotyping.
715   *American journal of human genetics* **104**, 749-757 (2019).
716 31   Ehmke, N. *et al.* De Novo Mutations in SLC25A24 Cause a Craniosynostosis Syndrome
717   with Hypertrichosis, Progeroid Appearance, and Mitochondrial Dysfunction. *American*
718   *journal of human genetics* **101**, 833-843 (2017).
719 32   Writzl, K. *et al.* De Novo Mutations in SLC25A24 Cause a Disorder Characterized by
720   Early Aging, Bone Dysplasia, Characteristic Face, and Early Demise. *American journal of*
721   *human genetics* **101**, 844-855 (2017).
722 33   Bogliolo, M. *et al.* Mutations in ERCC4, encoding the DNA-repair endonuclease XPF,
723   cause Fanconi anemia. *American journal of human genetics* **92**, 800-806 (2013).
724 34   Kashiyama, K. *et al.* Malfunction of nuclease ERCC1-XPF results in diverse clinical
725   manifestations and causes Cockayne syndrome, xeroderma pigmentosum, and Fanconi
726   anemia. *American journal of human genetics* **92**, 807-819 (2013).
727 35   Abu-Libdeh, B. *et al.* RECON syndrome is a genome instability disorder caused by
728   mutations in the DNA helicase RECQL1. *J Clin Invest* **132** (2022).
729 36   Nakura, J. *et al.* Homozygosity mapping of the Werner syndrome locus (WRN). *Genomics*
730   **23**, 600-608 (1994).
731 37   Oshima, J. *et al.* Homozygous and compound heterozygous mutations at the Werner
732   syndrome locus. *Hum Mol Genet* **5**, 1909-1913 (1996).
733 38   Sinclair, D. A., Mills, K. & Guarente, L. Accelerated aging and nucleolar fragmentation in
734   yeast sgs1 mutants. *Science* **277**, 1313-1316 (1997).
735 39   Yu, C. E. *et al.* Positional cloning of the Werner's syndrome gene. *Science* **272**, 258-262
736   (1996).
737 40   Ellis, N. A. *et al.* The Bloom's syndrome gene product is homologous to RecQ helicases.
738   *Cell* **83**, 655-666 (1995).
739 41   Mallery, D. L. *et al.* Molecular analysis of mutations in the CSB (ERCC6) gene in patients

with Cockayne syndrome. *American journal of human genetics* **62**, 77-85 (1998).

42 Henning, K. A. *et al.* The Cockayne syndrome group A gene encodes a WD repeat protein that interacts with CSB protein and a subunit of RNA polymerase II TFIIH. *Cell* **82**, 555-564 (1995).

43 Liano, D., Chowdhury, S. & Di Antonio, M. Cockayne Syndrome B Protein Selectively Resolves and Interact with Intermolecular DNA G-Quadruplex Structures. *J Am Chem Soc* **143**, 20988-21002 (2021).

44 Tang, W. *et al.* The Werner syndrome RECQ helicase targets G4 DNA in human cells to modulate transcription. *Hum Mol Genet* **25**, 2060-2069 (2016).

45 Fleming, A. M. *et al.* Human DNA Repair Genes Possess Potential G-Quadruplex Sequences in Their Promoters and 5'-Untranslated Regions. *Biochemistry* **57**, 991-1002 (2018).

46 Wu, W. Q., Hou, X. M., Li, M., Dou, S. X. & Xi, X. G. BLM unfolds G-quadruplexes in different structural environments through different mechanisms. *Nucleic acids research* **43**, 4614-4626 (2015).

47 Varshney, D., Spiegel, J., Zyner, K., Tannahill, D. & Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nat Rev Mol Cell Biol* **21**, 459-474 (2020).

48 Biffi, G., Tannahill, D., McCafferty, J. & Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat Chem* **5**, 182-186 (2013).

49 Hansel-Hertsch, R. *et al.* G-quadruplex structures mark human regulatory chromatin. *Nature genetics* **48**, 1267-1272 (2016).

50 Hansel-Hertsch, R., Spiegel, J., Marsico, G., Tannahill, D. & Balasubramanian, S. Genome-wide mapping of endogenous G-quadruplex DNA structures by chromatin immunoprecipitation and high-throughput sequencing. *Nat Protoc* **13**, 551-564 (2018).

51 Lyu, J., Shao, R., Kwong Yung, P. Y. & Elsasser, S. J. Genome-wide mapping of G-quadruplex structures with CUT&Tag. *Nucleic acids research* **50**, e13 (2022).

52 Muller, S., Kumari, S., Rodriguez, R. & Balasubramanian, S. Small-molecule-mediated G-quadruplex isolation from human cells. *Nat Chem* **2**, 1095-1098 (2010).

53 Zhang, X., Spiegel, J., Martinez Cuesta, S., Adhikari, S. & Balasubramanian, S. Chemical profiling of DNA G-quadruplex-interacting proteins in live cells. *Nat Chem* **13**, 626-633 (2021).

54 Esnault, C. *et al.* G4access identifies G-quadruplexes and their associations with open chromatin and imprinting control regions. *Nature genetics* **55**, 1359-1369 (2023).

55 Lago, S. *et al.* Promoter G-quadruplexes and transcription factors cooperate to shape the cell type-specific transcriptome. *Nature communications* **12**, 3885 (2021).

56 Shen, J. *et al.* Promoter G-quadruplex folding precedes transcription and is controlled by chromatin. *Genome biology* **22**, 143 (2021).

57 Bedrat, A., Lacroix, L. & Mergny, J. L. Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic acids research* **44**, 1746-1759 (2016).

58 Zhang, R. *et al.* G-Quadruplex Structures Are Key Modulators of Somatic Structural Variants in Cancers. *Cancer Res* **83**, 1234-1248 (2023).

59 De Magis, A. *et al.* DNA damage and genome instability by G-quadruplex ligands are mediated by R loops in human cancer cells. *Proceedings of the National Academy of Sciences of the United States of America* **116**, 816-825 (2019).

60 Katapadi, V. K., Nambiar, M. & Raghavan, S. C. Potential G-quadruplex formation at

breakpoint regions of chromosomal translocations in cancer may explain their fragility. *Genomics* **100**, 72-80 (2012).

61   Paeschke, K. *et al.* Pif1 family helicases suppress genome instability at G-quadruplex motifs. *Nature* **497**, 458-462 (2013).

62   Galati, E. *et al.* VID22 counteracts G-quadruplex-induced genome instability. *Nucleic acids research* **49**, 12785-12804 (2021).

63   Lee, C. Y. *et al.* R-loop induced G-quadruplex in non-template promotes transcription by successive R-loop formation. *Nature communications* **11**, 3392 (2020).

64   Miglietta, G., Russo, M. & Capranico, G. G-quadruplex-R-loop interactions and the mechanism of anticancer G-quadruplex binders. *Nucleic acids research* **48**, 11942-11957 (2020).

65   Rider, S. D., Jr. *et al.* Stable G-quadruplex DNA structures promote replication-dependent genome instability. *J Biol Chem* **298**, 101947 (2022).

66   Schiavone, D. *et al.* Determinants of G quadruplex-induced epigenetic instability in REV1-deficient cells. *The EMBO journal* **33**, 2507-2520 (2014).

67   Lawler, N. B. *et al.* G4-DNA formation and chromatin remodelling are interdependent in human cells. *Chem Sci* **14**, 7681-7687 (2023).

68   Mao, S. Q. *et al.* DNA G-quadruplex structures mold the DNA methylome. *Nat Struct Mol Biol* **25**, 951-957 (2018).

69   Faragher, R. G. *et al.* The gene responsible for Werner syndrome may be a cell division "counting" gene. *Proceedings of the National Academy of Sciences of the United States of America* **90**, 12030-12034 (1993).

70   Baird, D. M., Davis, T., Rowson, J., Jones, C. J. & Kipling, D. Normal telomere erosion rates at the single cell level in Werner syndrome fibroblast cells. *Hum Mol Genet* **13**, 1515-1524 (2004).

71   Proietti-De-Santis, L., Drane, P. & Egly, J. M. Cockayne syndrome B protein regulates the transcriptional program after UV irradiation. *The EMBO journal* **25**, 1915-1923 (2006).

72   Kato, H. *et al.* Generation of disease-specific and CRISPR/Cas9-mediated gene-corrected iPS cells from a patient with adult progeria Werner syndrome. *Stem Cell Res* **53**, 102360 (2021).

73   Gatinois, V. *et al.* iPSC line derived from a Bloom syndrome patient retains an increased disease-specific sister-chromatid exchange activity. *Stem Cell Res* **43**, 101696 (2020).

74   Andrade, L. N., Nathanson, J. L., Yeo, G. W., Menck, C. F. & Muotri, A. R. Evidence for premature aging due to oxidative stress in iPSCs from Cockayne syndrome. *Hum Mol Genet* **21**, 3825-3834 (2012).

75   Martins, S. *et al.* Generation of an induced pluripotent stem cell line (IUFi001) from a Cockayne syndrome patient carrying a mutation in the ERCC6 gene. *Stem Cell Res* **55**, 102456 (2021).

76   Wang, S. *et al.* Rescue of premature aging defects in Cockayne syndrome stem cells by CRISPR/Cas9-mediated gene correction. *Protein Cell* **11**, 1-22 (2020).

77   Cagirici, H. B., Budak, H. & Sen, T. Z. G4Boost: a machine learning-based tool for quadruplex identification and stability prediction. *BMC Bioinformatics* **23**, 240 (2022).

78   Lorenz, R. *et al.* ViennaRNA Package 2.0. *Algorithms Mol Biol* **6**, 26 (2011).

79   Janes, J. *et al.* Chromatin accessibility dynamics across C. elegans development and ageing. *eLife* **7** (2018).

80   Fabian, P. *et al.* Lifelong single-cell profiling of cranial neural crest diversification in

832   zebrafish. *Nature communications* **13**, 13 (2022).

833   81   Calderon, D. *et al.* The continuum of Drosophila embryonic development at single-cell
834   resolution. *Science* **377**, eabn5800 (2022).

835   82   Kumegawa, K. *et al.* GRHL2 motif is associated with intratumor heterogeneity of cis-
836   regulatory elements in luminal breast cancer. *NPJ Breast Cancer* **8**, 70 (2022).

837   83   Hendrickson, D. G. *et al.* A new experimental platform facilitates assessment of the
838   transcriptional and chromatin landscapes of aging yeast. *eLife* **7** (2018).

839   84   Corces, M. R. *et al.* Lineage-specific and single-cell chromatin accessibility charts human
840   hematopoiesis and leukemia evolution. *Nature genetics* **48**, 1193-1203 (2016).

841   85   Sheffield, N. C. & Bock, C. LOLA: enrichment analysis for genomic region sets and
842   regulatory elements in R and Bioconductor. *Bioinformatics* **32**, 587-589 (2016).

843   86   Spiegel, J. *et al.* G-quadruplexes are transcription factor binding hubs in human chromatin.
844   *Genome biology* **22**, 117 (2021).

845   87   Ahn, H. J., Cha, Y., Moon, S. H., Jung, J. E. & Park, K. S. Ell3 enhances differentiation of
846   mouse embryonic stem cells by regulating epithelial-mesenchymal transition and apoptosis.
847   *PLoS One* **7**, e40293 (2012).

848   88   Nakaki, F. & Saitou, M. PRDM14: a unique regulator for pluripotency and epigenetic
849   reprogramming. *Trends Biochem Sci* **39**, 289-298 (2014).

850   89   Matharu, N. K., Hussain, T., Sankaranarayanan, R. & Mishra, R. K. Vertebrate homologue
851   of Drosophila GAGA factor. *J Mol Biol* **400**, 434-447 (2010).

852   90   Maresca, M. *et al.* Pioneer activity distinguishes activating from non-activating SOX2
853   binding sites. *The EMBO journal* **42**, e113150 (2023).

854   91   Zhao, Y. *et al.* "Stripe" transcription factors provide accessibility to co-binding partners in
855   mammalian genomes. *Mol Cell* **82**, 3398-3411 e3311 (2022).

856   92   Bowman, E. A. & Kelly, W. G. RNA polymerase II transcription elongation and Pol II CTD
857   Ser2 phosphorylation: A tail of two kinases. *Nucleus* **5**, 224-236 (2014).

858   93   Li, C. *et al.* Ligand-induced native G-quadruplex stabilization impairs transcription
859   initiation. *Genome research* **31**, 1546-1560 (2021).

860   94   Li, C. *et al.* G-quadruplexes sense natural porphyrin metabolites for regulation of gene
861   transcription and chromatin landscapes. *Genome biology* **23**, 259 (2022).

862   95   Prorok, P. *et al.* Involvement of G-quadruplex regions in mammalian replication origin
863   activity. *Nature communications* **10**, 3274 (2019).

864   96   Lim, G. & Hohng, S. Single-molecule fluorescence studies on cotranscriptional G-
865   quadruplex formation coupled with R-loop formation. *Nucleic acids research* **48**, 9195-
866   9203 (2020).

867   97   St Germain, C. P. *et al.* Genomic patterns of transcription-replication interactions in mouse
868   primary B cells. *Nucleic acids research* **50**, 2051-2073 (2022).

869   98   Petryk, N. *et al.* Replication landscape of the human genome. *Nature communications* **7**,
870   10208 (2016).

871   99   Guilbaud, G. *et al.* Local epigenetic reprogramming induced by G-quadruplex ligands. *Nat
872   Chem* **9**, 1110-1117 (2017).

873   100   Zhao, P. A., Sasaki, T. & Gilbert, D. M. High-resolution Repli-Seq defines the temporal
874   choreography of initiation, elongation and termination of replication in mammalian cells.
875   *Genome biology* **21**, 76 (2020).

876   101   Ming, X. *et al.* Kinetics and mechanisms of mitotic inheritance of DNA methylation and
877   their roles in aging-associated methylome deterioration. *Cell Res* **30**, 980-996 (2020).

102 Charlton, J. *et al.* Global delay in nascent strand DNA methylation. *Nat Struct Mol Biol* **25**, 327-332 (2018).

103 Granja, J. M. *et al.* ArchR is a scalable software package for integrative single-cell chromatin accessibility analysis. *Nature genetics* **53**, 403-411 (2021).

104 Schep, A. N., Wu, B., Buenrostro, J. D. & Greenleaf, W. J. chromVAR: inferring transcription-factor-associated accessibility from single-cell epigenomic data. *Nature methods* **14**, 975-978 (2017).

105 Veronezi, G. M. B. & Ramachandran, S. Nucleation and spreading rejuvenate polycomb domains every cell cycle. *bioRxiv*, 2022.2008.2002.502476 (2023).

106 Grosselin, K. *et al.* High-throughput single-cell ChIP-seq identifies heterogeneity of chromatin states in breast cancer. *Nature genetics* **51**, 1060-1066 (2019).

107 Genolini, C., Alacoque, X., Sentenac, M. & Arnaud, C. kml and kml3d: R Packages to Cluster Longitudinal Data. *Journal of Statistical Software* **65**, 1 - 34 (2015).

108 Zhang, W. *et al.* Aging stem cells. A Werner syndrome stem cell model unveils heterochromatin alterations as a driver of human aging. *Science* **348**, 1160-1163 (2015).

109 Ting, T. W. *et al.* Cockayne Syndrome due to a maternally-inherited whole gene deletion of ERCC8 and a paternally-inherited ERCC8 exon 4 deletion. *Gene* **572**, 274-278 (2015).

110 Della Valle, F. *et al.* LINE-1 RNA causes heterochromatin erosion and is a target for amelioration of senescent phenotypes in progeroid syndromes. *Sci Transl Med* **14**, eabl6057 (2022).

111 Pascual-Garcia, P. *et al.* Metazoan Nuclear Pores Provide a Scaffold for Poised Genes and Mediate Induced Enhancer-Promoter Contacts. *Mol Cell* **66**, 63-76 e66 (2017).

112 Freibaum, B. D. *et al.* GGGGCC repeat expansion in C9orf72 compromises nucleocytoplasmic transport. *Nature* **525**, 129-133 (2015).

113 Zhang, K. *et al.* The C9orf72 repeat expansion disrupts nucleocytoplasmic transport. *Nature* **525**, 56-61 (2015).

114 Zhang, Y. *et al.* Single-cell epigenome analysis reveals age-associated decay of heterochromatin domains in excitatory neurons in the mouse brain. *Cell Res* **32**, 1008-1021 (2022).

115 Wang, X. *et al.* Molecular spectrum of excision repair cross-complementation group 8 gene defects in Chinese patients with Cockayne syndrome type A. *Sci Rep* **7**, 13686 (2017).

116 Ren, Y. *et al.* Three novel mutations responsible for Cockayne syndrome group A. *Genes Genet Syst* **78**, 93-102 (2003).

117 Zhu, K. *et al.* WDR82 Negatively Regulates Cellular Antiviral Response by Mediating TRAF3 Polyubiquitination in Multiple Cell Lines. *J Immunol* **195**, 5358-5366 (2015).

118 Studier, F. W. Protein production by auto-induction in high density shaking cultures. *Protein Expr Purif* **41**, 207-234 (2005).

119 Moye, A. L. *et al.* Telomeric G-quadruplexes are a substrate and site of localization for human telomerase. *Nature communications* **6**, 7643 (2015).

120 Xiao, Y. *et al.* Non-invasive diagnosis and surveillance of bladder cancer with driver and passenger DNA methylation in a prospective cohort study. *Clin Transl Med* **12**, e1008 (2022).

121 Xiao, Y. *et al.* Ultrasensitive tumour-agnostic non-invasive detection of colorectal cancer recurrence using ctDNA methylation. *Clin Transl Med* **12**, e1015 (2022).

122 Puig Lombardi, E. *et al.* Thermodynamically stable and genetically unstable G-quadruplexes are depleted in genomes across species. *Nucleic acids research* **47**, 6098-

37

924   6113 (2019).
925   123   Trapp, A., Kerepesi, C. & Gladyshev, V. N. Profiling epigenetic age in single cells. *Nat*
926   *Aging* **1**, 1189-1201 (2021).
927
928

Figure legends

Figure 1. Clock-like formation of DNA G-quadruplex foci during aging.

(a) Schematic of the analysis. DNA G-quadruplexes in mouse primary embryonic fibroblasts (pMEFs) were captured by a BG4 single-chain antibody and subjected to either immunofluorescence (IF) staining or pA-Tn5-mediated CUT&Tag fragmentation followed by qPCR or next-generation sequencing (NGS). **(b)** Fraction of BG4 CUT&Tag reads falling in aging-associated ATAC peaks. **(c)** Coefficient of variation of loci-wise BG4 qCUT&Tag signals in pMEF samples at different *in vitro* passages. **(d)** Representative images of BG4 immunofluorescence staining in young, middle-aged or old pMEFs. **(e)** Quantitative measurement of the BG4 immunofluorescence signal in individual pMEFs. *p*-value: t-test.

Figure 2. Clock-like formation of G4 associates with chromatin opening during aging.

(a) ATAC-seq profile of young, mid-aged, and old pMEFs for genes without pGQS in their promoter (pGQSneg), genes with pGQS in their promoter (pGQSpos), and aging genes with pGQS in their promoter (pGQSpos aging). **(b)** BG4 CUT&Tag profile of young, mid-aged, or old pMEFs on the pGQSneg, pGQSpos, or pGQSpos aging genes. **(c)** Estimated total opened pGQS loci in bulk ATAC-seq of pMEFs. **(d)** Estimated total opened pGQS loci in bulk ATAC-seq of primary mouse astrocytes (pAST) cells. **(e)** Single-cell ATAC reads falling on the pGQS locus in pMEFs. Cells are ranked by their *in vitro* passage number and cell cycle phase (G1, S, and G2/M). **(f)** Correlation of average estimated total opened pGQS loci by EpiTrace from the scATAC data and the scAge-inferred DNA methylation age from the same batch of cells. Samples from the same cell cycle phases were subjected to linear regression. **(g)** Correlation between the G4 minimal folding energy (MFE) in the ATAC-seq peak (X-axis) and the aging coefficient of the same peak

39

952 (Y-axis). **(h)** Aging coefficient of ATAC peaks containing <=3 or >3 G4 repeats. **(i)** MFE of human

953 ATAC-seq peaks containing only G3/(A/C/T) repeats or containing only G3/(G) repeats. **(j)** Aging

954 coefficients of human ATAC-seq peaks containing only G3/(A/C/T) repeats or containing only

955 G3/(G) repeats in CiPSCs induced from fibroblasts (FBs) or mesenchymal stem cells (MSCs).

956

957 **Figure 3. G4 formation is necessary for clock-like chromatin opening.**

958 **(a)** Top enriched sequence motif in aging ATAC-seq peaks in the mouse pMEF scATAC

959 experiment (top) and aging coefficient profiles of the pGQSpos and pGQSneg genes in the same

960 experiment (bottom). **(b)** Aging coefficients of ATAC-seq loci in mouse pMEF scATAC. Loci are

961 classified by whether they harbor a ClockDML, a pGQS, or G4 ChIP-seq signal. **(c)** Relative

962 prevalence of open (age >0, orange) and close (age <0, beige) ATAC-seq loci in mouse pMEF

963 scATAC. Loci are classified by whether they harbor a ClockDML, a pGQS, or G4 ChIP-seq signal.

964 **(d)** Correlation between the known sample age (X-axis) and the average estimated total number of

965 opened G4 loci (Y-axis) in the mouse pMEF scATAC experiment. Samples from different cell

966 cycle phases were separately analyzed. **(e)** Top enriched sequence motifs in aging ATAC-seq peaks

967 in the human CiPSC scATAC experiment (top) and aging coefficient profiles of the pGQSpos and

968 pGQSneg genes in the same experiment (bottom). **(f)** Aging coefficients of ATAC-seq loci in the

969 human breast cancer PDX scATAC cohort. Loci are classified by whether they harbor a ClockDML,

970 a pGQS, or G4 ChIP-seq signal. **(g)** Relative prevalence of open (age >0, orange) and close (age

971 <0, beige) ATAC-seq loci in the human breast cancer PDX scATAC cohort. Loci are classified by

972 whether they harbor a ClockDML, a pGQS, or G4 ChIP-seq signal. **(h)** Correlation of known cell

973 developmental rank (X-axis) and the estimated total number of opened G4 loci (Y-axis) in the

974 human hematopoietic cell ATAC experiment. **(i)** Correlations of known sample times (X-axis) and

40

975 average estimated total number of opened G4 loci (Y-axis) from (1) the aging ATAC experiment,

976 (2) the ATAC-seq data of *C. elegans* samples at known developmental stages, (3) the scATAC data

977 of Drosophila samples at known developmental times, and (4) the scATAC data of zebrafish neural

978 crest lineage cells at known developmental stages.

979

980 **Figure 4. G4 promotes transcription-replication interactions to delay genome replication at**

981 **aging loci.**

982 **(a)** NELF ChIP-seq profiles of control (CTRL) and PDS-treated (PDS+) HEK293T cell samples

983 for pGQSpos genes. **(b)** Nascent RNA (TT-seq) profiles of control (CTRL) and TMPyP4-treated

984 (TMPyP4+) HEK293T cell samples of the pGQSpos gene. **(c)** Profiles of transcription-replication

985 interactions on nascent DNA (TRIPn) on the pGQSneg, pGQSpos, and pGQSpos aging genes in

986 mouse spleen B cells. **(d)** Sitewise TRIPn signal enrichment (over non-BrdU control) of the

987 pGQSneg, pGQSpos, and pGQSpos aging genes in mouse spleen B cells. **(e)** Short replication-

988 associated nucleotide (SNS-seq) profiles of control (CTRL) and PhenDC3-treated (PhenDC3+)

989 mouse MEF cell samples of the pGQSpos gene. **(f)** Short replication-associated nucleotide (SNS-

990 seq) profiles of control (CTRL) and PhenDC3-treated (PhenDC3+) mouse MEF cell samples of

991 the pGQSneg gene. **(g)** Percentage of mouse spleen B cells harboring significant enrichment of

992 the TRIPn signal from the pGQSneg, pGQSpos, and pGQSpos aging genes. **(h)** BrdU pull-down

993 Repli-seq signal across 16 fractions (time points) of S phase cells grouped according to different

994 classes of genes in mouse embryonic stem cells (mESCs), mouse neural progenitor cells (mNPCs),

995 and human embryonic stem cells (hESCs, H1 and H9). **(i)** Cumulative BrdU pull-down Repli-seq

996 signal (normalized to min(0%)-max(100%)) across 16 fractions (time points) of S phase, grouped

997 by different classes of genes, in mouse embryonic stem cells (mESCs), mouse neural progenitor

998   cells (mNPCs), and human embryonic stem cells (hESCs, H1 and H9). **(j)** Schematic of the finding.

999   G4 stimulates transcription-replication interactions by simultaneously promoting both

1000   transcription and replication, resulting in delayed genome replication.

1001

1002   **Figure 5. Delayed genome replication by G4 impairs epigenome transmission during aging.**

1003   **(a)** Methylation maintenance level of early (left) or late (right) S-phase EdU pulse-thymidine-

1004   chase HeLa cells, nonpGQS CpG, PMD solo-WCGW CpG, G4- pGQS CpG, or G4+ pGQS CpG

1005   according to Hammer-seq experiments. The post-EdU chase times are shown on the X-axis (log-

1006   ed). **(b)** Normalized H3K27me3 CUT&Flow signals of pGQSpos (left) and pGQSneg (right) in

1007   mouse mESCs. **(c)** DNA methylation level at age-associated TF-bound G4 peaks in human PBMC

1008   G4AMP samples. Human donor ages are shown on the X-axis. **(d)** DNA methylation level at age-

1009   associated TF-bound G4 peaks in mouse tail fibroblast G4AMP samples. Mouse sample ages are

1010   shown on the X-axis (log-ed). **(e)** Mean H3K27me3 CUT&Flow signal profiles of pGQSpos (left)

1011   and pGQSneg (right) mouse mESCs across different cell cycle phases. **(f)** Single-cell H3K27me3

1012   ChIP-seq signal per G4 locus in human breast cancer PDX cells. The cells were clustered into

1013   common ancestor cells, derived cells in non-drug-resistant sample, and derived cells in the drug-

1014   resistant sample. **(g)** Relative frequency (%) of pGQSneg or pGQSpos ATAC peaks exhibiting an

1015   aging peak. **(h)** Average DNA replication (BrdU, yellow) and normalized H3K27me3 signal

1016   (H3K27me3, blue) across different time points (X-axis) during cell cycle progression. **(i)** Relative

1017   enrichment of aging-related loci in the pGQSpos gene cluster. These pGQSpos genes are clustered

1018   according to their Repli-seq and H3K27me3 CUT&Flow temporal profiles. **(j)** Average DNA

1019   replication (BrdU, yellow) and normalized H3K27me3 signal (H3K27me3, blue) across different

1020   time points (X-axis) for non-aging and aging pGQSpos genes. **(k)** Schematic of the findings.

During DNA replication, G4 might directly hinder the replication fork or create TRI in nascent synthesized DNA, resulting in delayed genome replication. Such delayed genome replication reduced the available time window for histone modification and/or G4-associated DNA methylation recovery.
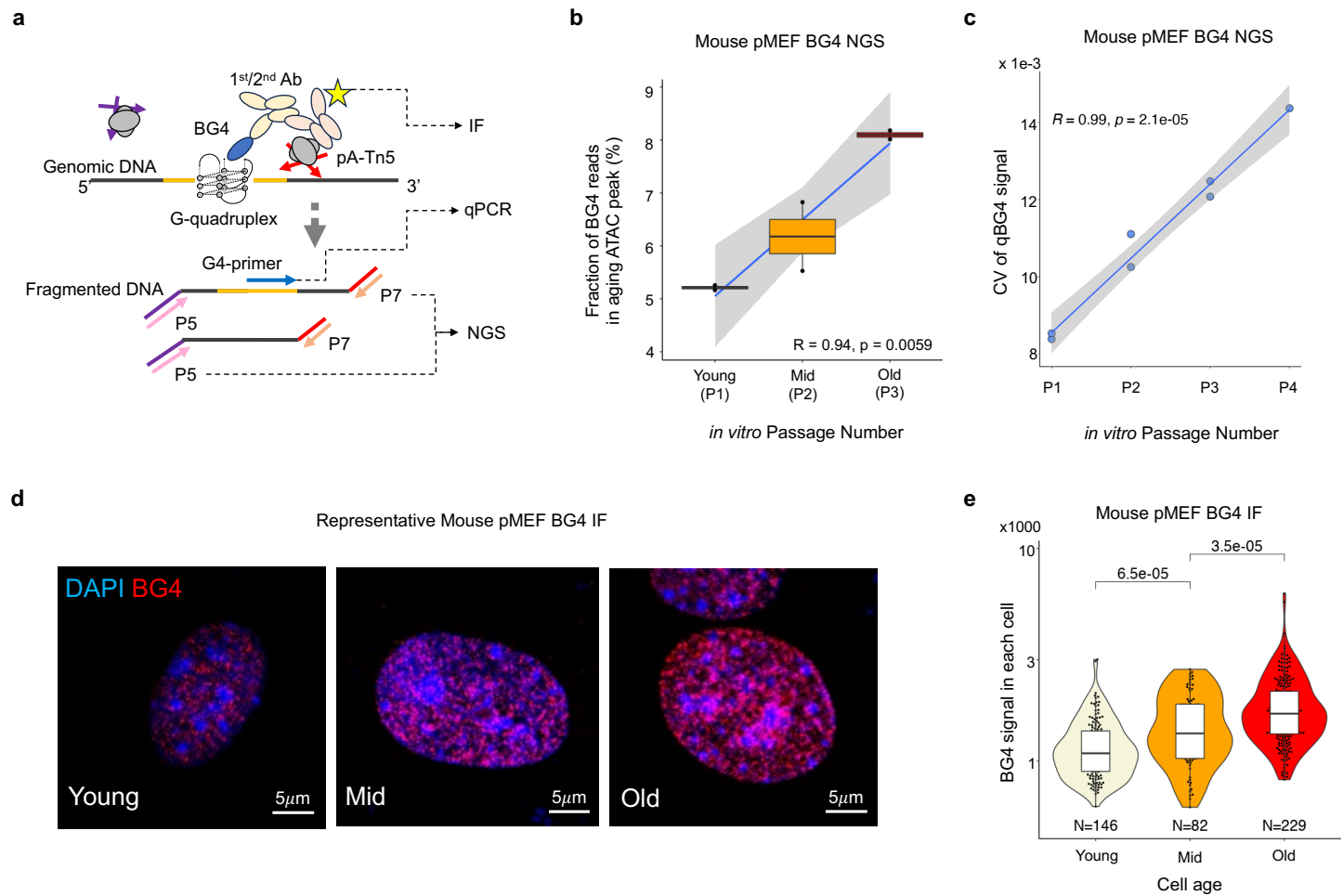
Figure 6. Genetic mutations causing G4 accumulation accelerate the erosion of the epigenetic landscape around G4 *in vivo*.

(a) H3K9me3 profiles across the pGQSpos or pGQSneg sites. Profiles were determined between +/- 50 kb of a particular given site for wild-type (WT) or WRN knockout (WRN KO) hMESC samples. **(b)** H3K4me3 profiles across the pGQSpos or pGQSneg sites. Profiles were determined between +/- 50 kb of a particular gene in wild-type (WT) or WRN knockout (WRN KO) hMESC samples. **(c)** Number of H3K27me3 peaks from wild-type (WT) or WRN knockout (WRN KO) ATAC samples. **(d)** DNA methylation level at age-associated TF-bound G4s peaked in age-matched progeroid and normal control donors. **(e)** DNA methylation level at age-associated TF-bound G4 peaks in ERCC8-mutant progeroid patients segregated by the nature of the ERCC8 mutation (hypomorphic or null). **(f)** Linear regression of DNA methylation levels in age-associated TF-bound G4 peaks in ERCC8-mutant patients and controls (Y-axis) according to donor age (years, X-axis). **(g)** Summary of findings. WRN/BLM/ERCC8 mutations result in accelerated erosion of the epigenetic landscape adjacent to G4.

Figure 7. Summary of G4-driven aging.

During eukaryotic cell aging, putative G-quadruplex regions at transcription start sites can be opened by pioneering factors such as Zbtb7a/GAF. The opened chromatin stochastically folds into

1044   a G-quadruplex. The G-quadruplex induces transcription-replication interactions to delay genome

1045   replication in S phase and impairs epigenetic information transmission locally, resulting in gradual,

1046   clock-like DNA hypomethylation and chromatin opening. G-quadruplex-driven epigenomic aging
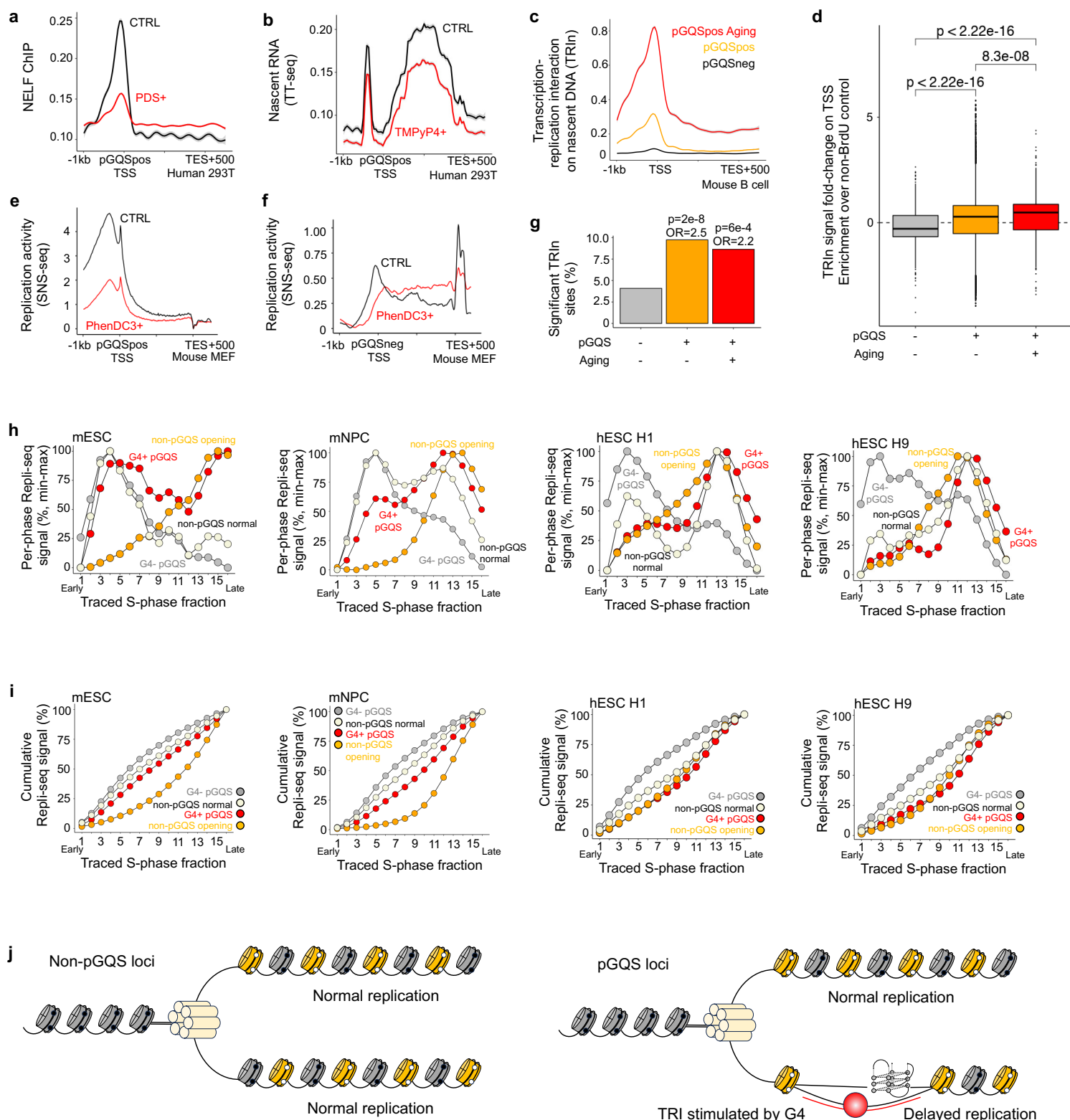
1047   is highly conserved during evolution.

1048

Figure 1. Clock-like formation of DNA G-quadruplex foci during aging.
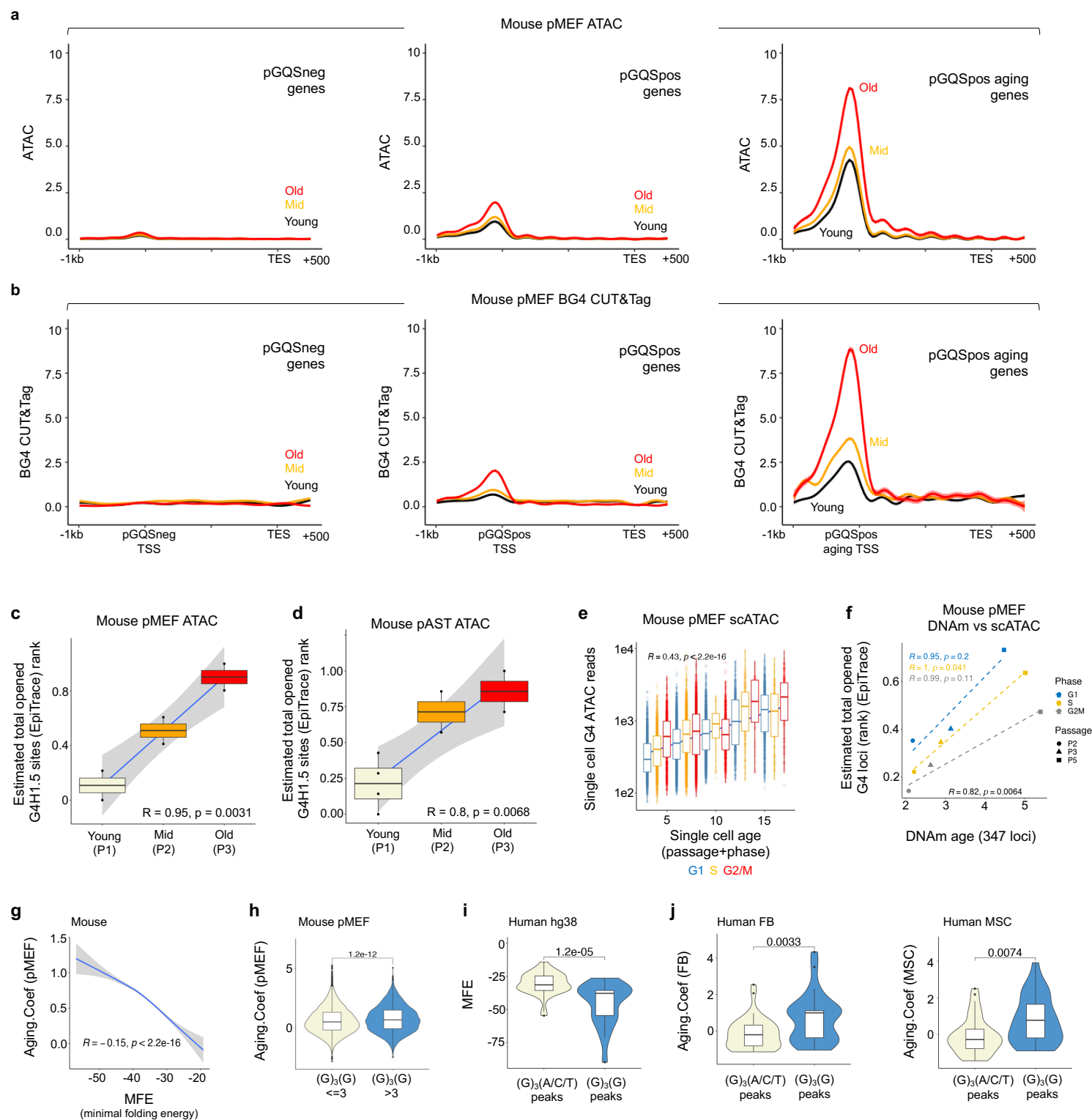
Figure 2. Clock-like formation of G4 associates with chromatin opening during aging.
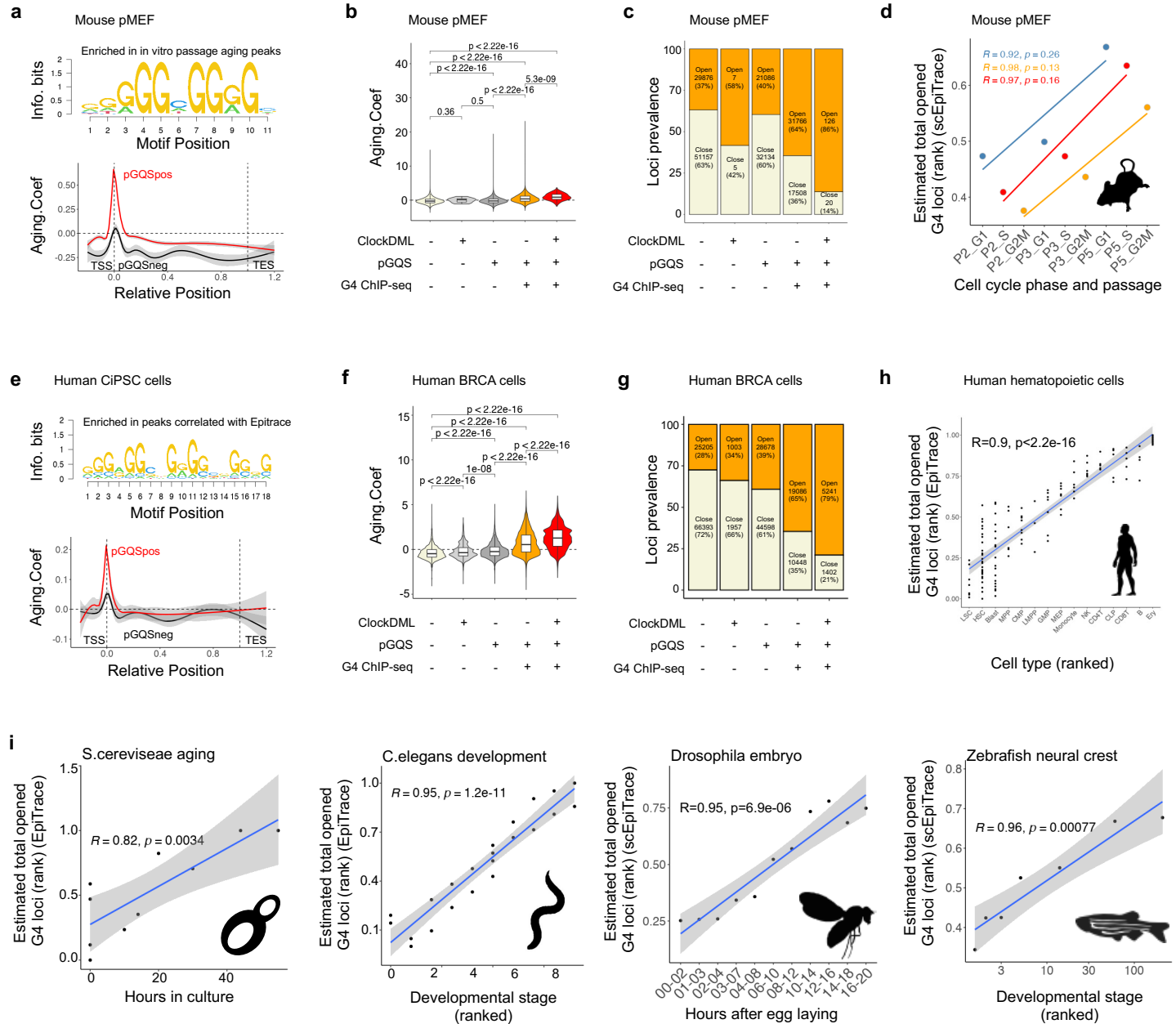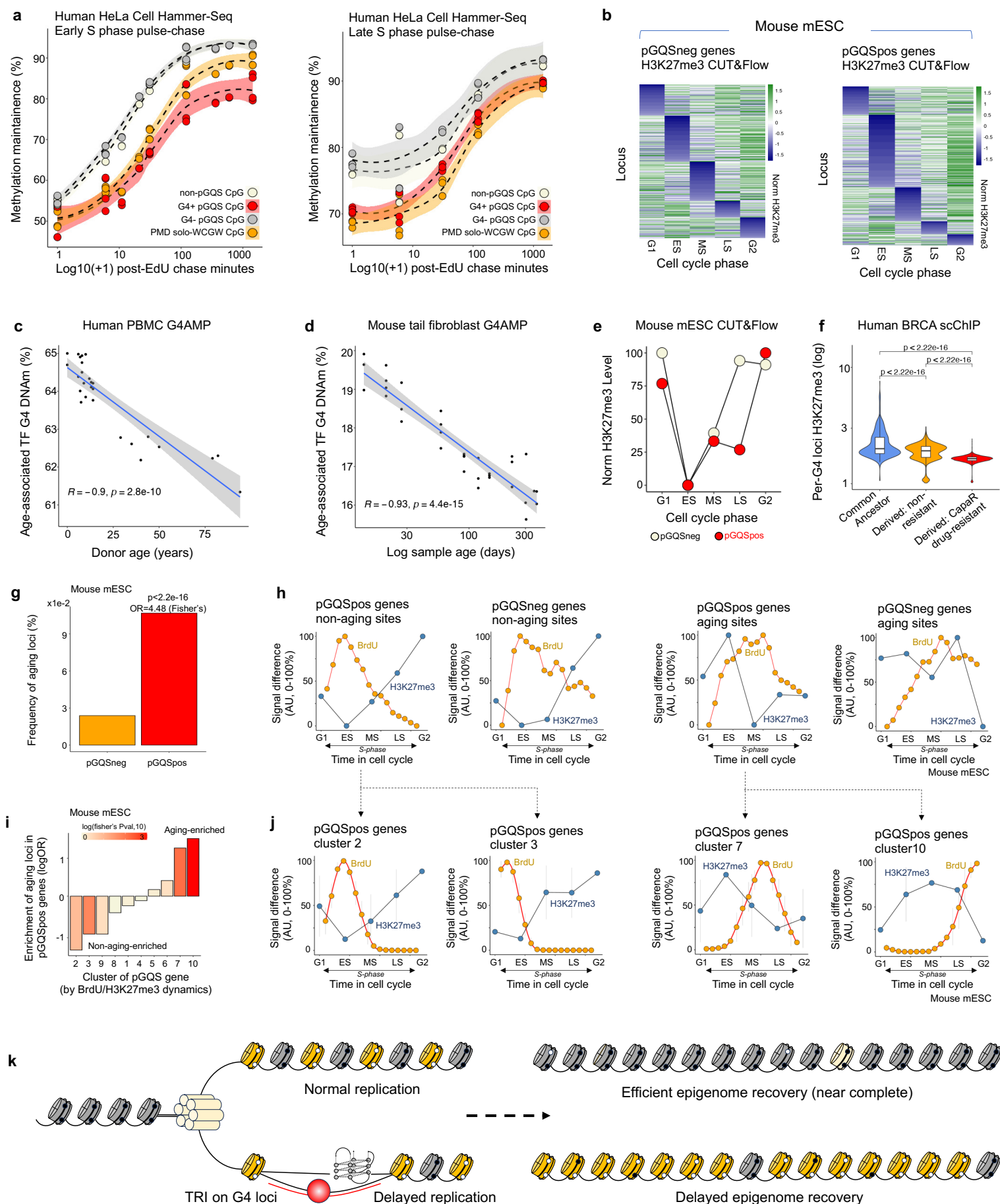
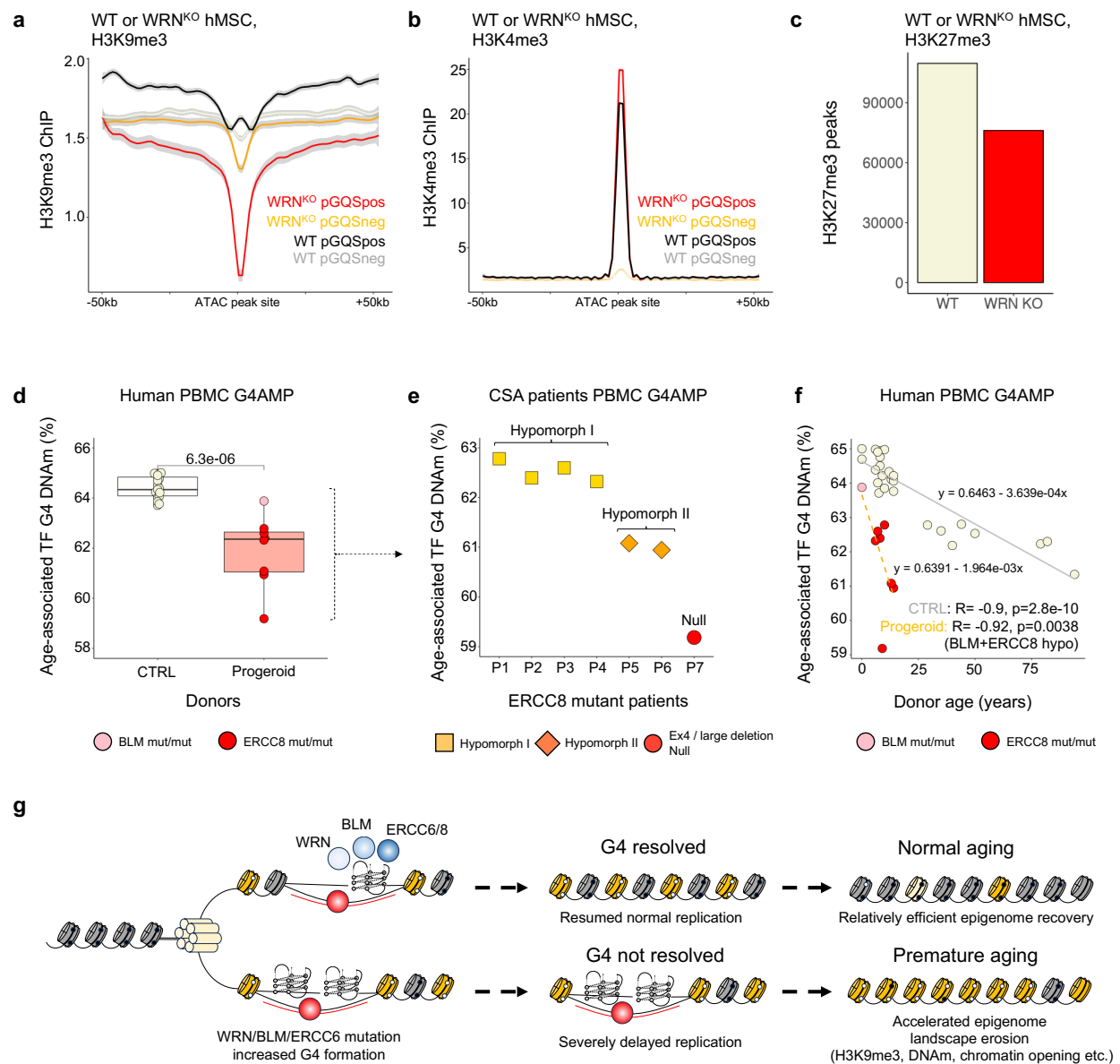**Figure 3. G4 formation is necessary for clock-like chromatin opening.**

Figure 4. G4 promotes transcription-replication interactions to delay genome replication at aging loci.

Figure 5. Delayed genome replication by G4 impairs epigenome transmission during aging.

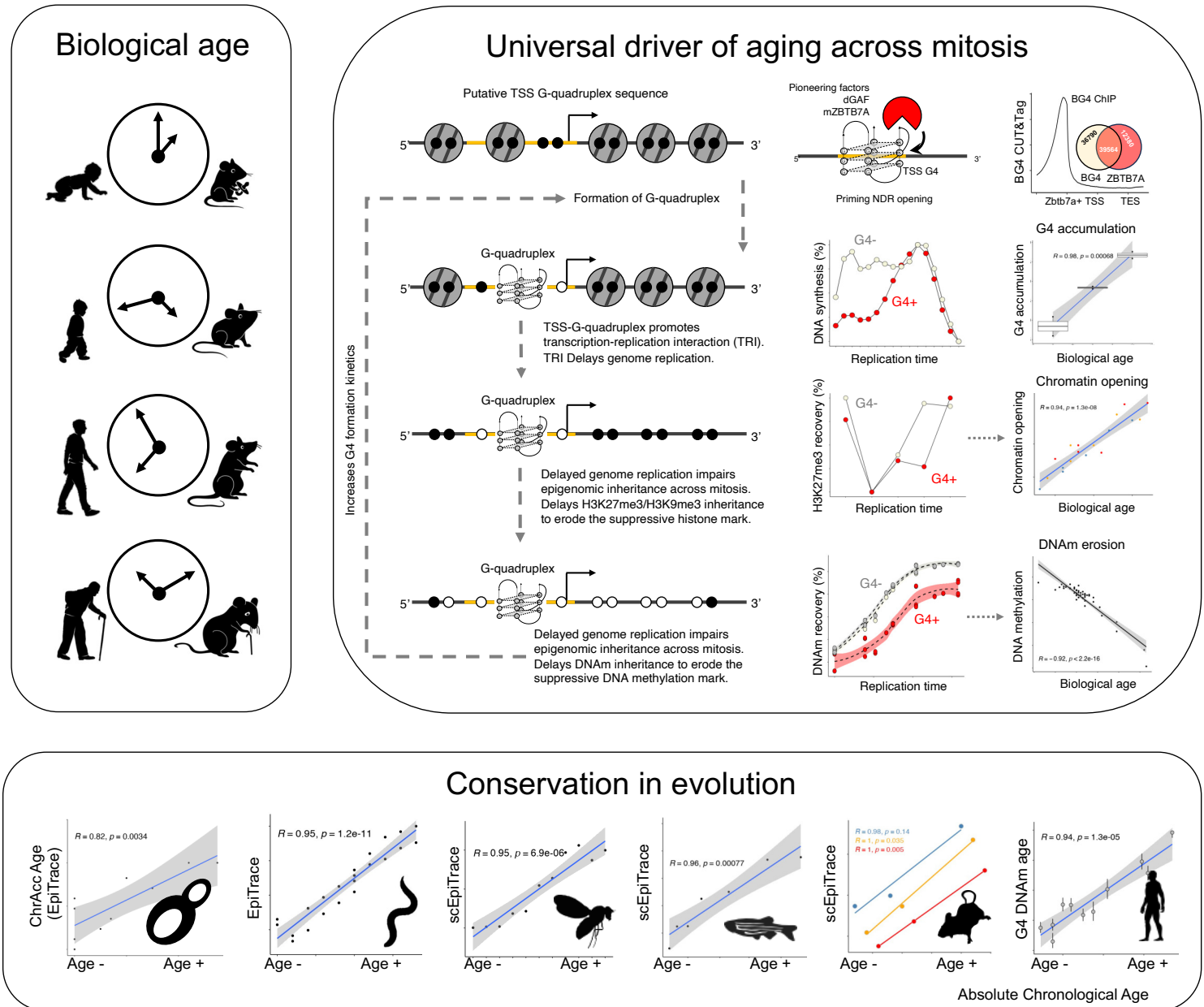**Figure 6. Genetic mutations causing G4 accumulation accelerate the erosion of the epigenetic landscape around G4** *in vivo.*

Figure 7. Summary of G4-driven aging.