Comparative genomics reveals signatures of distinct metabolic strategies and gene loss associated with *Hydra* immortality

Kaoruko Nojiri[1¶], Koryu Kin[1,2,3¶], Akimasa Someya[1,4], Tetsuo Kon[5], Koto Kon-Nanjo[5], Hiroshi Shimizu[6], Kazuharu Arakawa[7,8] and Etsuo A. Susaki[1,2,3,9*]

Affiliations:

[1]Faculty of Medicine, Juntendo University, Bunkyo-ku, Tokyo, Japan

[2]Department of Biochemistry and Systems Biomedicine, Juntendo University Graduate School of Medicine, Bunkyo-ku, Tokyo, Japan

[3]Nakatani Biomedical Spatialomics Hub, Juntendo University Graduate School of Medicine, Bunkyo-ku, Tokyo, Japan

[4]Juntendo University Graduate School of Pharmacy and Pharmaceutical Sciences, Urayasu, Chiba, Japan

[5]Department of Neurosciences and Developmental Biology, University of Vienna, Vienna, Austria

[6]National Institute of Genetics, Mishima, Shizuoka, Japan

[7]Institute for Advanced Biosciences, Keio University, Tsuruoka, Yamagata, Japan

[8]Graduate School of Media and Governance, Keio University, Fujisawa, Kanagawa, Japan

[9]Institute for Quantum Life Science, National Institutes for Quantum Science and Technology, Inage, Chiba, Japan

¶These authors contributed equally to this work.

*Author for Correspondence: Etsuo A. Susaki, Department of Biochemistry and Systems Biomedicine, Juntendo University, Tokyo, Japan, +81 358021033 suishess-kyu@umin.ac.jp

Abstract

Hydra is a freshwater cnidarian genus that provides a unique comparative model for aging research, contrasting the immortal *H. vulgaris* with the aging-inducible *H. oligactis.* Here, we report a high-quality, chromosome-level genome assembly of *H. vulgaris* strain AEP.JNIG. Our assembly is comparable in quality to existing resources, facilitating the analysis of genomic diversity across laboratory strains. Epigenomic profiling revealed that gene-body hypermethylation correlates with transcriptional stability and the suppression of spurious transcription in evolutionary conserved genes, suggesting an epigenetic mechanism for genomic integrity. Furthermore, comparative genomics demonstrated that while *Hydra* conserves fundamental *Hallmarks of Aging* pathways, the immortal *H. vulgaris* paradoxically lacks canonical anti-aging genes (e.g., Klotho, NAMPT) found in the aging-inducible *H. oligactis*. Instead, *H. vulgaris* exhibits a distinct metabolic signature related to mitochondrial energy production and NTP synthesis. Collectively, our comparative genomics results suggest multiple potential mechanisms associated with the *H. vulgaris* immortality and the aging traits of *H. oligactis*, providing novel targets for future functional studies.

Key words: *Hydra*, aging, genome assembly, comparative genomics, epigenetic regulation

Significance statement

Why do some organisms age while others appear not to? The freshwater animal *Hydra* provides a unique opportunity to investigate this question, as closely related species display contrasting aging phenotypes. We generated a high-quality genome assembly for a new strain of a non-aging species and conducted comparative analyses with related strains and an aging species. Even closely related strains can accumulate substantial genetic divergence over time, and stable DNA modification patterns were associated with consistent gene activity, suggesting a mechanism that may help maintain cellular balance. Surprisingly, several well-known longevity genes are present in the aging species but absent in the non-aging one. This suggests that extended lifespan may not simply depend on possessing more “anti-aging” genes, but instead may reflect differences in how core biological processes are organized. Our study provides new insights into the genetic basis of aging and highlights *Hydra* as a powerful model for understanding longevity.

Introduction

Hydra is a freshwater cnidarian that can be maintained in the laboratory in mass culture (Klimovich et al. 2019; Galliot 2012; Loomis & Lenhoff 1956). Based on its distinctive traits including potent regenerative capacity and both sexual and asexual modes of reproduction, *Hydra* has been widely studied in fields such as stem cell biology, regeneration, and aging (Holstein 2023; Vogg et al. 2019; Tomczyk et al. 2015). In particular, *Hydra vulgaris* is one of the few organisms in which negligible senescence has been demonstrated by long-term population monitoring of mortality and fertility (Martínez 1998; Schaible et al. 2015). Aging organisms exhibit progressive decline in physiological function and accumulation of pathological changes over time, characterized by an increase in mortality rate (Gompertz curve) (Finch et al. 1990; Jones et al. 2014). On the other hand, *H. vulgaris* maintains a flat Gompertz curve (Schaible et al. 2015), suggesting its immortality. This remarkable trait is thought to be supported by the continuous self-renewal capacity of its stem cell populations, regulated by genes including FoxO and TCF/β-catenin (Boehm et al. 2012; Hemmrich et al. 2012; Gufler et al. 2018). Interestingly, *H. oligactis*, another species within the *Hydra* genus, shows increased mortality and physiological deterioration after sexual reproduction when transferred to 10 °C (Yoshida et al. 2006). This inducible aging phenotype has been attributed to the depletion of interstitial stem cells (ISCs) (Sun et al. 2020; Tomczyk et al. 2020).

Hydra also constitutes an apparent exception to evolutionary theories of aging. These theories generally propose that organismal aging reflects an evolutionary trade-off that favors an energy-saving strategy, in which ensuring germline accuracy is essential, whereas maintaining somatic cell integrity is often a dispensable expenditure (Kirkwood 1977; Hamilton 1966). Therefore, whereas many long-lived species delay reproduction until later developmental stages, a pattern consistent with such theoretical predictions, *Hydra* attains sexual maturity early in life yet retains an extended lifespan (Schaible et al. 2015).

What might be the origin of these unique aging traits in *Hydra*? While organismal aging is generally considered a complex process involving both genetic and environmental factors, empirical evidence suggests that genetic components represent the primary determinant. Multiple findings support this proposition. First, maximum lifespan and aging patterns described by Gompertz curves show species-specific differences (Jones et al. 2014). Second, multiple aging-related genes and pathways have been identified, such as Klotho, progeroid syndrome-related genes including DNA repair and nuclear integrity, and NAD^+-related pathways (Nabeshima 2002; Lombard et al. 2005). The *Hallmarks of aging* provides a comprehensive framework that summarizes genes and biological processes in aging (López-Otín et al. 2023). Third, tissue and organismal aging is closely associated with the resilience of tissue stem cells, which is also determined by genomic information (Rando & Jones 2021). Lastly, a recent study demonstrated that the heritability of intrinsic life span is approximately 50%, suggesting that genetic control over longevity is comparable to other complex traits (Shenhar et al. 2026). Therefore, comparative genomic analysis may provide critical insights into the genetic elements underlying the immortality observed in *H. vulgaris.*

While high-quality whole-genome assembly for the *H. vulgaris* strain AEP (North American strain) (Cazet et al. 2023), originally established in North America (Martínez et al. 2010), provides useful references for comparative genomics, they do not capture the full extent of intra-species genomic variation. Such variation may arise from spontaneously occurring mutants in laboratory cultures. Indeed, independent maintenance of laboratory populations or colonies can generally lead to genomic divergence across species (Chebib et al. 2021; Farslow et al. 2015). In *H. vulgaris* 105 (formerly *H. magnipapillata*) maintained by asexual budding, the stem cells have accumulated mutations at even slightly higher rates than those in mammals, with some under positive selection (Sahm et al. 2024). Distinct types of transposable elements, such as the CR1 LINE element, are inserted into stem cell genomes, further contributing to genomic divergence (Kon-Nanjo et al. 2025; Wong et al. 2019). These suggest that the *Hydra*

genome can change over long-term cultures in laboratories, even when the animals reproduce asexually.

To avoid risks of overlooking population-specific variants that can affect phenotypic traits, at least two high-quality reference genomes from independently maintained *H. vulgaris* populations are desirable. To address this, we sequenced and assembled a chromosome-level genome for the *H. vulgaris* AEP strain maintained at the National Institute of Genetics, Mishima, Japan (hereafter *H. vulgaris* strain AEP.JNIG) (Kawaida et al. 2010), thereby expanding available genomic resources for this species. We verified that *H. vulgaris* AEP.JNIG forms a monophyletic group with other *H. vulgaris* strains, including AEP and 105, and is the most closely related to AEP, while chromosomal analyses revealed substantial structural differences, confirming genome-level divergence among populations. Building on these data, we conducted a comparative analysis of aging-related genes and pathways using the genomes of three *H. vulgaris* strains and the aging species *H. oligactis* (Cazet et al. 2023). By comparing signaling pathways and genes tied to the *Hallmarks of Aging* (López-Otín et al. 2023), we identified gene sets that differ between aging and non-aging species. From these comparisons, we characterized genomic differences between *H. vulgaris* and representative aging animals, including *H. oligactis*, and identified candidate genes potentially associated with *H. vulgaris* immortality.

145 # Results

146 ## Chromosome-level genome assembly of *H. vulgaris* AEP.JNIG

147

148 The *H. vulgaris* population maintained at the National Institute of Genetics, Mishima, Japan (*H.*
149 *vulgaris* AEP.JNIG, Fig 1A) was investigated. We independently sequenced and assembled
150 the *H. vulgaris* AEP.JNIG genome as part of the Japanese *Hydra* resource; during the course
151 of our work, a genome assembly of an *H. vulgaris* AEP strain maintained in the US was
152 published (Cazet et al. 2023). The availability of both assemblies allowed us to evaluate the
153 quality of the AEP.JNIG genome and to perform population-level comparisons of genomic
154 variation.

155

156 The whole genome sequence of *H. vulgaris* AEP.JNIG genome was initially collected as the
157 assemblies of Illumina paired-end reads and Nanopore long reads. We obtained
158 82,804,328,738 bp of paired-end short reads and 168,380,136,564 bp of Nanopore long reads
159 with an N50 read length of 3,493 bp (Table 1). *De novo* assembly was performed using these
160 sequences with Phased Error Correction and Assembly Tool (PECAT) (Nie et al. 2024) (see
161 Materials and Methods for details). The initial genome assembly had a total size of
162 954,340,183 bp, consisting of 248 contigs. The longest contig was 61,504,050 bp, the contig
163 N50 was 9,577,232 bp, the contig L50 was 30, and the GC content was 27.94%. (Table 2).

164

165 To construct a chromosome-level genome assembly, 398,191,666 Hi-C read pairs were
166 generated from the same clone. We scaffolded the initial genome assembly using the Hi-C
167 reads with HiC-pro (Servant et al. 2015) and YaHS (Zhou et al. 2022) (see Materials and
168 Methods for details). Fifteen chromosome-level scaffolds were obtained, each showing strong
169 intra-chromosomal contact signals in the Hi-C contact map (Fig 1B). The total size of the final
170 assembly was 922,069,947 bp, the largest scaffold was 86,117,336 bp, scaffold N50 was
171 63,895,935 bp, the contig L50 was 7, and the GC content was 27.88% (Table 2).

We evaluated the accuracy and completeness of the final genome assembly by searching for sets of BUSCO (v 5.8.3) (Manni et al. 2021) using the metazoa odb10 database (Kriventseva et al. 2019). We identified 904 complete BUSCO genes out of 954 (94.7%), consisting of 879 single-copy genes (92.1%) and 25 duplicated genes (2.6%). 11 BUSCO genes were identified as fragmented (1.2%), and 39 BUSCO genes were missing (4.1%) (Fig 1C and Table 3). These assembly results are comparable to those for the *H. vulgaris* AEP genome (Cazet et al. 2023) (Tables 2 and 3), supporting the sufficient quality of our genome assembly dataset. We further created gene models based on this chromosome-level assembly and the transcriptome using BREAKER2 (Brůna et al. 2021) and evaluated BUSCO scores. This analysis identified 891 complete BUSCO genes out of 954 (93.4%), consisting of 772 single-copy genes (80.9%) and 119 duplicated genes (12.5%). 14 BUSCO genes were fragmented (1.5%), and 49 BUSCO genes were missing (5.1%). Again, these results are also comparable to those for the AEP genome (Cazet et al. 2023) (Table 4).

To further confirm the chromosome-level integrity of our AEP strain, we also performed karyotyping (Fig 1D-F). Our microscopic observations predominantly indicate the karyotype of 2n = 30, with a few exceptions likely due to artifacts during sample preparations (Fig 1E, S1 Fig). Importantly, the absence of any karyotypes with 2n > 32 strongly suggests that *H. vulgaris* AEP.JNIG does not possess extra chromosome pairs compared to other *H. vulgaris* strains. Therefore, we conclude that the *H. vulgaris* AEP.JNIG has a karyotype of 2n = 30 (Fig 1F).

Comparison of *H. vulgaris* AEP.JNIG genome with AEP genome revealed overall conservation of chromosomal structure

To determine the phylogenetic relationship between our strain and the other *Hydra* strains, we performed species tree reconstruction using a large-scale phylogenomic dataset (a

199 multiple sequence alignment of 981,632 amino acid positions from 954 proteins) from 31
200 metazoan organisms including those of two other *H. vulgaris* strains and 2 holozoan protists
201 (Cazet et al. 2023; Kon-Nanjo et al. 2025; Chapman et al. 2010) (see Materials and Methods
202 for details) (Fig 2A). The resulting species tree robustly placed *H. vugaris* AEP.JNIG as the
203 closest relative of the *H. vulgaris* AEP strain. Furthermore, in order to see whether there is
204 any strain closer to AEP.JNIG than AEP, we extracted four genes (*COI*, *COII*, *EF1α*, *Cnnos1*)
205 from the AEP.JNIG genome and made a phylogenetic tree with other strains using a dataset
206 previously published (Kawaida et al. 2010). In all cases, *H. vulgaris* AEP.JNIG formed a
207 monophyletic group with *H. vulgaris* AEP (S2 Fig), further corroborating a close phylogenetic
208 relationship between *H. vulgaris* AEP.JNIG and *H. vulgaris* AEP.
209
210 To characterize the degree of genetic divergence between AEP and AEP.JNIG, we compared
211 their genomes using D-GENIES (Cabanettes & Klopp 2018) and GENESPACE (Lovell et al.
212 2022) (Fig 2B and 2C). The dot plot from the genome alignment between these two genome
213 sequences by D-GENIES revealed prominent diagonal lines, indicating a high degree of
214 collinearity between them (Fig 2B). The ribbon plot from the gene syntenic analysis result by
215 the GENESPACE demonstrated overall correspondence at the chromosomal level (Fig 2C).
216 We then used plotsr (Goel & Schneeberger 2022) and SyRI (Goel et al. 2019) to visualize the
217 detailed genomic structural variations such as inversions, translocations, and duplications,
218 and to quantify both structural and sequence-level differences including SNPs and indels. In
219 total, 954 syntenic regions were identified, covering 796 Mb of the AEP assembly and 789 Mb
220 of the AEP.JNIG assembly. In addition, 360 inversions (31.5 Mb) and 660 translocations (31.5
221 Mb) were also detected, along with numerous duplication events (239 in the AEP and 2,073
222 in the AEP.JNIG) (Fig 2D and S3A Fig). At the sequence level, 1,833,472 SNPs, 648,211
223 insertions (5.4 Mb in total), and 261,207 deletions (4.7 Mb in total) were identified (S3A Fig).
224 Together, these results indicate that overall chromosomal organization is largely conserved
225 between the two strains, despite substantial accumulation of sequence-level variation.
226 Although we observed structural variations between the AEP and AEP.JNIG assemblies, we

227 cannot entirely rule out the possibility that some of these differences arise from misassemblies
228 in either genome. However, given the strong signals observed in the Hi-C contact maps for
229 both assemblies, such artifacts are likely minimal.

230

231 We also examined the transposable element (TE) landscape of the *H. vulgaris* AEP.JNIG
232 assembly using Earl Grey (v6.3.3) (Baril et al. 2024). Repetitive elements comprised 78.4%
233 of the total genome (Table 5). DNA transposons were the most abundant class, representing
234 35.8% of the genome, and LINEs accounted for 15.2%. By contrast, LTR and Rolling-circle
235 elements were relatively rare (1.0% and 1.0%, respectively). For direct comparison, we
236 applied the same analysis to the *H. vulgaris* AEP assembly. The overall repeat content
237 (77.7%) and element composition were similar between the two populations, although LINEs
238 and Peneopes were slightly more abundant in AEP.JNIG (Table 6), indicating general
239 consistency between the two assemblies. Despite slight differences in subfamily composition,
240 TE landscape comparisons between the AEP.JNIG and AEP assemblies showed overall
241 consistency (S3B and S3C Figs).

242

243 DNA Methylation within gene bodies contributes to transcriptional 244 stability (homeostasis) of conserved genes in *H. vulgaris*

245

246 Leveraging the methylation signal in the Nanopore reads, we examined genome-wide mCpG
247 distribution in *H. vulgaris* AEP.JNIG using Dorado (v0.7.1) and Modkit (v0.4.3) (see Materials
248 and Methods). A mosaic methylation pattern was observed across the genome, with stably
249 methylated and unmethylated regions interspersed. Centromeric regions identified by quarTeT
250 (v1.2.5) (Lin et al. 2023) had low DNA methylation (Fig 3A).

251

252 We next compared methylation levels across genomic features. DNA methylation levels in
253 promoters, exons, introns, and intergenic regions exhibited bimodal distributions, with distinct

254 peaks corresponding to highly and weakly methylated regions. The methylation levels within
255 gene regions were higher on average than those within promoter and intergenic regions (Fig
256 3B and S4A Fig). Repetitive elements also exhibited a bimodal methylation pattern. Among
257 repeat classes, LINEs and LTRs tended to be more highly methylated, a trend that is clearer
258 in the boxplots **(**Fig 3C and S4B Fig).
259
260 We further analyzed the relationship between DNA methylation and gene expression in our
261 dataset. We focused on conserved genes, defined as orthologs shared across the 33
262 holozoans (see Materials and Methods), because conserved genes have been reported to be
263 hypermethylated in multiple species, including cnidarians (Zhang & Jacobs 2022; Sarda et al.
264 2012). Consistent with these reports, conserved genes in our data showed higher methylation
265 than other genes (Fig 3D). Conserved genes tended to be more highly expressed and exhibited
266 less variation in expression levels (Fig 3E and 3F), suggesting a link between methylation and
267 stable gene regulation. Finally, the ratio of transcription levels between exon 1 and downstream
268 exons, a measure of potential spurious transcription (Li et al. 2018), was negatively correlated
269 with exon methylation, and this correlation was more pronounced for conserved genes than
270 for other genes (Fig 3G).
271
272 These results suggest that genomic methylation, particularly within gene bodies, may help
273 stabilize higher expression levels and suppress spurious transcription, especially in conserved
274 genes. This observation is consistent with findings in other cnidarians, such as coral (Li et al.
275 2018).
276

277 Aging-related orthologs are conserved in *Hydra* comparable to other 278 aging model species

279

280 We then investigated aging-related orthologs in immortal *H. vulgaris* strains (105 (Chapman et
281 al. 2010), AEP (Cazet et al. 2023), and AEP.JNIG (this study)) and the aging species *H.*
282 *oligactis* (Cazet et al. 2023) in comparison with other model organisms, including *Homo*
283 *sapiens*, *Mus musculus*, *Drosophila melanogaster*, and *Caenorhabditis elegans*. To reduce the
284 impact of incomplete genome assemblies or gene annotations, we integrated orthogroup
285 information across three *H. vulgaris* strains, treating an orthogroup as present when it was
286 detected in any one of the strains.
287
288 First, we manually compiled a list of aging-related genes based on the *Hallmarks of Aging*
289 framework (López-Otín et al. 2023). Specifically, genes explicitly mentioned in the original
290 paper were directly included. For hallmarks described at the level of signaling pathways rather
291 than individual genes, constituent genes were identified primarily using the KEGG Pathway
292 database (Kanehisa & Goto 2000). When pathways were not clearly defined or fully
293 represented in KEGG, additional genes were identified using the Gene Ontology (GO)
294 database and relevant literature. The complete list of genes is summarized in S1 Table.
295 Subsequently, we used OrthoFinder (Emms & Kelly 2019) to construct orthogroups based on
296 predicted protein sequences derived from genome assemblies, supplementing the analysis
297 with BLAST searches. Finally, we examined the orthogroups annotated with these aging-
298 related genes. The presence of a sequence from a given species within an orthogroup was
299 taken as evidence that the species possesses the corresponding gene.
300
301 Our analysis revealed that *Hydra* (both *H. vulgaris* and *H. oligactis*) conserve a substantial
302 proportion of genes associated with key aging-related pathways, such as autophagy,
303 proteostasis, DNA damage repair, mitochondrial quality control, stem cell regulation, insulin
304 receptor signaling, NAD signaling, and cGAS-STING signaling (Fig 4A and S5 Fig). For
305 instance, approximately 93% of the genes in the KEGG autophagy pathway (107/115) and
306 93% in the base excision repair pathway (39/42) were identified in *H. vulgaris* (see S2 Table
307 for details). These findings are consistent with previous reports (Pascual-Torner et al. 2022;

308 Barve et al. 2021; Chera et al. 2009) and suggest that the fundamental mechanisms of aging
309 mediated by these genes are likely conserved across species, including the immortal *H.*
310 *vulgaris* (Fig 4A).
311
312 Furthermore, a cross-species comparison of the total number of aging-related orthologs
313 revealed that *Hydra* shares a larger number of conserved orthologs with humans than do *D.*
314 *melanogaster* and *C. elegans* (Fig 4B). This indicates that *Hydra* has retained a highly
315 conserved repertoire of aging-related genes, despite its significant evolutionary distance from
316 traditional model organisms. Our results are supported by previous work using Reciprocal Best
317 Hit (RBH) analysis, which similarly showed that *H. vulgaris* possesses a greater number of
318 human orthologs than either *D. melanogaster* or *C. elegans* (Wenger & Galliot 2013).
319
320 Therefore, these results substantiate the potential of *Hydra* as a model organism for aging
321 research, offering a unique opportunity to investigate conserved mechanisms of organismal
322 aging across the metazoan tree of life.
323

324 Comparative Analysis between *H. vulgaris* and *H. oligactis* reveals
325 aging-related orthologs

326
327 Finally, we conducted a comparative orthogroup analysis between *H. vulgaris* and *H. oligactis*,
328 particularly aiming at the identification of gene sets related to their aging or immortal traits. As
329 was in Fig 4, we used the integrated dataset for *H. vulgaris* orthogroups. To confirm the
330 absence of these orthogroups in each species, a BLAST search (Altschul et al. 1990) was
331 performed (see Materials and Methods). This analysis revealed that *H. vulgaris* possessed 33
332 unique orthogroups that were not detected in *H. oligactis*, while *H. oligactis* contained 65
333 unique orthogroups (Fig 5A and 5B; S3 Table).
334

335 To explore their biological functions, we performed Gene Ontology (GO) enrichment, KEGG
336 pathway, and Reactome pathway analyses using clusterProfiler (v4.16.0) (Wu et al. 2021)
337 (Figs 5C and 5D; S6 Fig; S4 Table). Intriguingly, terms related to oxidative phosphorylation
338 (OXPHOS) were enriched in both species (e.g., adjusted p-value = 4.60E-5 for *H. oligactis* and
339 9.72E-4 for *H. vulgaris* in GO-biological processes (GO-BP); S4 Table). However, the specific
340 genes contributing to these terms are contrasted. Unique genes identified in the *H. oligactis*
341 GO-BP terms included α-synuclein (SNCA), mitochondrial deoxyguanosine kinase (DGUOK),
342 and components of complexes I (NDUFA10, MT-ND3), III (MT-CYB), IV (MT-CO2/CO3), and
343 V (MT-ATP6). In contrast, *H. vulgaris* uniquely possessed different components of complexes
344 I (NDUFA3/B7), III (UQCR10/11), and V (ATP5F1E). The analysis also revealed species-
345 specific terms regarding nucleic acid metabolic processes. For instance, *H. oligactis*-specific
346 orthogroups were enriched for GO-BP terms such as "purine deoxyribonucleoside
347 monophosphate metabolic process" (adjusted p-value = 8.29E-7), "purine nucleotide metabolic
348 process" (adjusted p-value = 2.99E-6), and "amide metabolic process" (adjusted p-value =
349 3.71E-5) (Fig 5C and S4 Table). Conversely, in *H. vulgaris*, major GO-BP terms related to NTP
350 synthesis were enriched, such as "ATP biosynthetic process" (adjusted p-value = 9.72E-4) and
351 "purine nucleoside triphosphate biosynthetic process" (adjusted p-value = 1.10E-3) (Fig 5D
352 and S4 Table). These contrasting results suggest species-specific characteristics in energy
353 production and metabolic processes.
354
355 We further narrowed our analysis to aging-related genes and pathways unique to each species.
356 By intersecting the species-specific orthogroups with the aging-related genes associated with
357 the *Hallmarks of Aging* (Fig 4 and S1 Table), we identified twelve [SNCA, FUT1/2, NAMPT,
358 BLVRB, SCD/SCD5, DPT, RBP3, IRG family (IRGC, IRGM), MPG, POLR2L, CD14, Klotho
359 family (KLB, LCT, LCTL, KL, GBA3)] and five [MON1A/1B, TOMM7, SS18 family (SS18,
360 SS18L1/2), PCLAF, RPS29] unique orthogroups for *H. oligactis* and *H. vulgaris*, respectively
361 (Fig 5B). Extending our comparative analysis to 29 other holozoan species revealed that the
362 twelve aging-related orthogroups identified in *H. oligactis* are largely conserved across these

363 taxa. Notably, both humans and mice possess the complete set of these twelve orthogroups
364 (Fig 5E). Conversely, the five orthogroups initially identified as unique to the immortal *H.*
365 *vulgaris* were found to be present in the other 29 species (Fig 5E). This suggests that these
366 orthogroups are unlikely to be the primary drivers of the immortal phenotype. Consequently,
367 we focused our subsequent investigation on the twelve *H. oligactis*-specific orthogroups.
368
369 These twelve orthogroups possess well-known functions in aging processes. α-synuclein
370 (SNCA) is a key molecule implicated in Parkinson's disease and mitochondrial oxidative
371 phosphorylation (OXPHOS) deficiency (Miwa et al. 2022). Nicotinamide adenine dinucleotide
372 (NAD) is widely recognized for its association with aging (Imai & Guarente 2014; Imai 2016),
373 and upregulation of genes involved in NAD^+ biosynthesis accompanies lifespan extension
374 (Tyshkovskiy et al. 2023). Nicotinamide phosphoribosyltransferase (NAMPT) functions as a
375 rate-limiting enzyme in NAD^+ biosynthesis, while biliverdin reductase B (BLVRB) is an
376 NAD(P)H-dependent reductase involved in antioxidant defense. Stearoyl-CoA desaturase
377 (SCD/SCD5) has been linked to brain aging in humans and mouse models of Alzheimer's
378 disease through its role in the synthesis of unsaturated fatty acids (McNamara et al. 2008;
379 Hamilton et al. 2022). N-methylpurine DNA glycosylase (MPG) is involved in the base excision
380 repair (BER) pathway, potentially contributing to the DNA damage response associated with
381 aging, and is known to be upregulated in long-lived mammals (Tyshkovskiy et al. 2023; Xu et
382 al. 2008). The presence of RNA polymerase II, I, and III subunit L (POLR2L) may be related
383 to aging-associated changes in transcriptional integrity (Debès et al. 2023; Papantonis et al.
384 2025). Finally, it is intriguing that Klotho family molecules are absent in *H. vulgaris*. *Klotho* is a
385 well-known anti-aging gene whose knockout in mice causes premature aging (Kuro-o 2009).
386 Although the absence of Klotho in the immortal *H. vulgaris* appears inconsistent with the
387 findings in mice, it suggests that the dependence on Klotho-mediated pathways, such as
388 phosphate and vitamin D metabolism, may be essential for conferring the aging traits.
389

390 Collectively, these comparative analyses offer insights to address fundamental biological
391 questions: specifically, how *H. vulgaris* acquired the immortal trait (or conversely, lost the aging
392 trait) and what constitutes the essential genes and pathways underlying organismal aging.
393

394 # Discussion

395

396 The purpose of this study was to establish a chromosome-level genome assembly and
397 accompanying epigenomic resources for *H. vulgaris* strain AEP.JNIG and offer insights into *H.*
398 *vulgaris* immortality through comparative genome analysis.

399

400 We established a chromosome-level genome assembly for *H. vulgaris* AEP.JNIG with quality
401 metrics comparable to the previously reported AEP strain (BUSCO: 94.7%; Scaffold N50: 63.9
402 Mb). Using the new high-quality genome assembly and accompanying gene annotations, we
403 clarified the relationship of our strain to other *H. vulgaris* strains. Our analyses demonstrate
404 that *H. vulgaris* AEP.JNIG forms a monophyletic group with *H. vulgaris* AEP. This close
405 relationship was consistently supported by both a phylogenomic dataset including multiple *H.*
406 *vulgaris* strains as well as single-gene analyses incorporating a broader taxonomic sampling
407 of *Hydra* laboratory strains and natural isolates.

408

409 Despite the phylogenetic proximity of the *H. vulgaris* AEP.JNIG and *H. vulgaris* AEP, through
410 detailed comparison between the two genomes, we detected a substantial accumulation of
411 SNPs and indels, along with numerous inversions and duplications. These genomic variations
412 likely reflect "laboratory drift" driven by long-term asexual reproduction, consistent with recent
413 findings that asexual lineages accumulate somatic mutations at high rates (Kershenbaum et
414 al. 2025; Sahm et al. 2024) and undergo structural changes such as Robertsonian
415 translocations (Kon-Nanjo et al. 2025). Together, these findings highlight substantial genomic
416 diversity within the species, establishing the AEP.JNIG assembly as a crucial resource that
417 complements the existing *H. vulgaris* genomes.

418

419 In addition to the genome assembly, we provide foundational resources for epigenomic
420 analyses, including single-nucleotide 5-methylcytosine (5mC) calls and transposable element

421 (TE) annotations. Our Nanopore signal-based analysis identified a global methylation level of
422 19%, which is lower than the 27% previously estimated by bisulfite sequencing (Ying et al.
423 2022). This difference likely reflects the improved accuracy of signal-based detection, which is
424 particularly advantageous for analyzing repeat-rich genomes like *Hydra* (Doshi et al. 2025).
425 We confirmed that the *H. vulgaris* genome exhibits a characteristic methylation feature of
426 invertebrates (Suzuki & Bird 2008), where gene bodies are hypermethylated while promoters
427 and putative centromeres remain hypomethylated. Further, we found the potential functional
428 link between DNA methylation and transcriptional stability. We observed that hypermethylation
429 within gene bodies, particularly in evolutionary conserved genes, correlates with high
430 expression levels and low transcriptional noise (spurious transcription). This finding suggests
431 that gene body methylation serves as a mechanism to maintain transcriptional fidelity and
432 homeostasis in *Hydra*, aligning with findings in other cnidarians such as corals (Li et al. 2018;
433 Dixon et al. 2018). The robust methylation-based stabilization of the transcriptome, together
434 with the loss of an RNA polymerase component (POLR2L) (Fig 5), may represent a key
435 machinery for *Hydra*'s negligible senescence. Given that epigenetic alterations relates to the
436 progress of aging, and that species-specific epigenetic signatures have recently been shown
437 to determine maximum lifespan in mammals (Li et al. 2024), *Hydra* serves as a unique model
438 for investigating how an organism can escape these epigenetic constraints.
439
440 Finally, to interrogate the genetic basis of *Hydra* immortality, we employed a comparative
441 genomics approach contrasting the immortal *H. vulgaris* with the aging-inducible *H. oligactis*,
442 consistent with recent frameworks emphasizing comparisons between exceptional species
443 and related controls (Rechsteiner et al. 2025). Our orthology analysis confirmed that *Hydra*
444 shares a larger repertoire of aging-related orthologs with humans than do traditional
445 invertebrate models (Wenger & Galliot 2013) and highly conserves genes associated with the
446 *Hallmarks of Aging*. This supports the view that aging mechanisms are ancestral (Lemoine
447 2021), suggesting *Hydra* achieves immortality by robustly regulating these pathways rather
448 than lacking them.

449

450 Intriguingly, however, we identified a counter-intuitive divergence: potential anti-aging genes
451 (e.g., Klotho, NAMPT, MPG) are unique to the aging *H. oligactis* and absent in the immortal *H.*
452 *vulgaris*. Additionally, metabolic signatures differed, with *H. vulgaris* enriched for ATP
453 synthesis pathways and *H. oligactis* for amide/purine metabolism pathways. We hypothesize
454 that *H. vulgaris*, inhabiting thermally stable environments (Khalturin et al. 2009), has optimized
455 an efficient energy metabolism that minimizes cellular damage, such as the synthesis of
456 reactive oxygen species. In contrast, the presence of metabolic regulators like Klotho in *H.*
457 *oligactis*, living in a cooler environment and sensitive to heat (Khalturin et al. 2009; Bosch et
458 al. 1988), implies complexity and trade-offs in the metabolic system, consistent with an
459 evolutionary aging theories such as antagonistic pleiotropy (Kirkwood 1977; Hamilton 1966).
460 Thus, *H. vulgaris* may maintain immortality by relying on a simpler homeostatic system that
461 avoids dependence on signaling pathways prone to age-dependent deterioration and damage
462 accumulation. From the viewpoint of evolutionary aging theories, it would be intriguing to
463 investigate mechanisms by which sexual reproduction triggers senescence in *H. oligactis*,
464 particularly within the context of the metabolic trade-offs imposed by its distinct genetic
465 background.

466

467 Several limitations of this study should be noted. First, our comparative analysis relies on
468 orthology with human aging-related genes defined by the *Hallmarks of Aging*. This approach
469 may overlook lineage-specific aging mechanisms unique to cnidarians. Additionally, despite
470 rigorous BLAST validation (e-value < 10^{-5}), the large evolutionary distance may lead to
471 potential undetected orthologs due to sequence divergence. Second, and most importantly,
472 the functional roles of the identified candidate genes (e.g., *H. oligactis*-specific genes) and the
473 causal link between DNA methylation and transcriptional stability remain to be experimentally
474 verified. Future studies employing functional assays, such as gene knockdown or transgenic
475 approaches, are essential to confirm the biological significance of these genomic findings for
476 *Hydra* aging and immortal traits.

477

478 In conclusion, our study provides an additional genomic resource comprising a new high-
479 quality, chromosome-level assembly of *H. vulgaris* AEP.JNIG, which facilitates the analysis of
480 genomic variation across independent laboratory populations. We also present a robust DNA
481 methylation profile derived from Nanopore sequencing signals, offering a foundation for future
482 epigenetic studies in cnidarians. Furthermore, through comparative genomics, we offer novel
483 insights into the potential genomic origins of *H. vulgaris* immortality, highlighting distinct
484 metabolic strategies and the evolutionary trade-offs associated with anti-aging genes.
485 Collectively, these resources and findings further establish *Hydra* as a robust and unique
486 model for advancing the comparative biology of aging. To fully interrogate the questions
487 regarding immortality in *Hydra*, more extensive analyses are required, including investigations
488 into individual gene functions, the system-level input-output characteristics of entire pathways,
489 and genetic loss- and gain-of-function assays.

490 # Materials and Methods

491

492 ## *Hydra* strain and culture

493 A *Hydra vulgaris* strain maintained as AEP at the Japan National Institute of Genetics (JNIG)
494 was acquired and cultured at Juntendo University using standard procedures. Briefly, the
495 animals were kept at 20ºC in the *Hydra* medium (1 mM Tris-HCl pH7.6, 1 mM $CaCl_2$, 1 mM
496 NaCl, 0.1 mM KCl, 0.1 mM $MgSO_4$) and were fed brine shrimp (*Artemia franciscana*) three
497 times a week.

498

499 ## Genome and RNA-Seq sequencing

500 For clean genomic DNA extraction, 100 polyps of *H. vulgaris* AEP.JNIG polyps were starved
501 for 10 days in the presence of antibiotics (50 µg/mL of ampicillin, rifampicin, neomycin, and
502 streptomycin) and herbicide (3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU), 5 µM) to
503 eliminate possible algal contaminants. Whole genomic DNA was then extracted using WiZard
504 HMW DNA Extraction Kit (Promega, Cat. No. A2920) and sent for sequencing.

505

506 Nanopore sequencing was done using the instrument PromethION with the flowcell version
507 R.9.4.1. The basecalling was done using Guppy version 5.0
508 (https://nanoporetech.com/ja/software/other/guppy/). The qualities of the reads were assessed
509 using NanoPlot (https://github.com/wdecoster/NanoPlot) (De Coster & Rademakers 2023).
510 Short-read genome sequencing libraries were also prepared from the extracted genomic DNA
511 using the Truseq DNA PCR-Free kit and 150 bp paired-end reads were sequenced on an
512 Illumina NovaSeq 6000 platform.

513

514 The Hi-C sequencing library was prepared from 100 polyps of *H. vulgaris* AEP.JNIG with
515 EpiTect Hi-C Kit (Qiagen) according to manufacturer's instructions and was sequenced on
516 NovaSeq X Plus (Illumina) as 150 bp paired-end reads.

517

518 ## Karyotyping

519 Karyotyping was performed with minor modifications to a previously described protocol for
520 various invertebrates (Guo et al. 2018). To increase the number of proliferating cells (David &
521 Campbell 1972), *H. vulgaris* AEP.JNIG polyps were heavily fed in the morning and, 12 hours
522 later, treated with 0.1% nocodazole for 1 hour at room temperature. The organisms were then
523 fixed in Carnoy's fixative (methanol : acetic acid = 3 : 1) for 30 minutes on ice. Fixed *Hydra*
524 bodies were transferred to glass slides and incubated with a few drops (~30 µL) of 60% acetic
525 acid for 5 minutes at room temperature. Siliconized coverslips prepared with Sigmacote
526 (Sigma-Aldrich, Cat. No. SL2-25ML) were placed on top of the specimens, and the tissues
527 were gently squashed by applying constant pressure for approximately 3 seconds to obtain a
528 single layer of nuclei (Guo et al. 2018). Slides were incubated overnight at 4 °C and
529 subsequently chilled on dry ice for 10 minutes. Coverslips were removed using a blade while
530 the slides remained on dry ice. After returning to room temperature, slides were gently rinsed
531 twice with 1x phosphate-buffered saline (PBS), stained with 1 µg/mL 4′,6-diamidino-2-
532 phenylindole (DAPI) for 5 minutes at room temperature, rinsed again with PBS, and mounted
533 using ProLong Diamond (Invitrogen, Cat. No. P36965). Samples were imaged using an in-
534 house spinning-disk confocal microscope (Nikon Eclipse Ti2-E equipped with a spinning-disk
535 scanner with 100-µm-wide pinholes aligned with a Nipkow disk (CSU-MPϕ100; Yokogawa
536 Electric) (Otomo et al. 2015) and an oil-immersion lens (SR HP Apo TIRF 100XC Oil, numerical
537 aperture [NA] = 1.49, Nikon)).

538

539 ## Genome assembly and annotation

540 The fastq files were assembled using the Phased Error Correction and Assembly Tool
541 (PECAT) (Nie et al. 2024) with medaka (https://github.com/nanoporetech/medaka) as a
542 polishing tool in the final phase of assembly.

543

544 For Hi-C scaffolding in *H. vulgaris* AEP.JNIG basecalled and demultiplexed reads totaling 119
545 Gbp were first filtered using fastp v0.19.4 (Chen 2023) with -c and -g options, and the filtered
546 reads were processed with HiC Pro v3.1.0 (Servant et al. 2015). Qualified reads were used to
547 scaffold the assembly using YAHS v1.2 (Zhou et al. 2022) with default parameters. Scaffolds
548 with length over 10 kbp were retained. Contact map was generated using Juicer 2.20 pipeline
549 (Durand et al. 2016). The Hi-C contact map was curated using the Juicebox v1.11.08 (Durand
550 et al. 2016). The final Hi-C contact map and genome assembly were generated using the 'run-
551 asm-pipeline-post-review.sh' script from the 3D-DNA pipeline (Dudchenko et al. 2017). The
552 completeness of filtered genome assembly was evaluated by BUSCO v5.7.1 (Manni et al.
553 2021) with the metazoa_odb10 dataset.
554
555 Prior to gene annotation, we soft-masked the assembled genome using RepeatModeler v2.0.2
556 and RepeatMasker v4.1.1 (Flynn et al. 2020). The soft-masked genomes were used to
557 annotate protein-coding genes using BRAKER v2.1.6 (Brůna et al. 2021). We used *H. vulgaris*
558 AEP.JNIG whole-body RNA-seq datasets as hints for annotation, mapped with HISAT2 v.2.2.1
559 (Kim et al. 2019).
560
561 The gene models annotated for the initial genome assembly were transferred to the final
562 assembly using Liftoff v1.6.3 (Shumate & Salzberg 2021). Nine transcripts that were not
563 transferred properly with the initial Liftoff run were manually mapped to the new assembly using
564 minimap2 (Li 2018) with the splice mode, and the ORFs were predicted using Transdecoder
565 (Haas).
566

567 Phylogenetic tree reconstruction

568 We constructed a proteome dataset of 31 metazoan and 2 unicellular holozoan organisms. A
569 workflow of constructing multiple sequence alignments based on the one described in (Liu et
570 al. 2024) was followed. Briefly, BUSCO v.5.8.0 (Manni et al. 2021) (Manni et al. 2021)was

571 performed against each proteome using metazoa_odb10. Single-copy genes were extracted
572 from the BUSCO output of each species and collected into individual FASTA files. These
573 FASTA files were filtered first by the number of species present in the dataset, with the
574 threshold being 23 (i.e., less than 10 missing species). PREQUAL v1.02 (Whelan et al. 2018)
575 was run with a posterior probability threshold of 0.95. Then MAFFT v7.525 (Katoh & Standley
576 2013) was used with the "-linsi" mode, without any other optional parameters. Divvier v1.01
577 (Ali et al. 2019) was used to remove regions of uncertain alignment (with the "-mincol 4 -
578 divvygap" option), followed by trimming using trimAl v1.5 (Capella-Gutiérrez et al. 2009) (with
579 the "-automated1" option). All the individual alignments were concatenated to create the final
580 supermatrix used for the subsequent phylogenetic analyses. A maximum-likelihood
581 phylogenetic tree was reconstructed using IQ-TREE2 v2.3.6 (Minh et al. 2020) under the
582 "LG+C60+F+G" model, with branch support assessed by the ultrafast bootstrap method (1,000
583 replicates).
584
585 Molecular phylogenetic trees of *Hydra* strains based on mitochondrial and nuclear genes (*COI*,
586 *COII*, *EF1α*, and *Cnnos1*) were reconstructed using the dataset used in (Kawaida et al. 2010).
587 Corresponding genes were retrieved from the *H. vulgaris* AEP.JNIG and the *H. vulgaris* AEP
588 genomes and added to the original dataset. All trees were reconstructed with IQ-TREE2 under
589 the best-fitting substitution models selected by ModelFinder (MFP option), with branch support
590 assessed using the ultrafast bootstrap method.
591

592 Detection and visualization of genomic rearrangements between *H.* 593 *vulgaris* AEP.JNIG and AEP assemblies

594 A dot plot was generated with D-GENIES (Cabanettes & Klopp 2018) using the genomes of *H.*
595 *vulgaris* AEP.JNIG and *H. vulgaris* AEP. Gene synteny plot was created using GENESPACE
596 (Lovell et al. 2022). Structural rearrangements between the two assemblies were then detected
597 by SyRI v1.7.1 (Goel et al. 2019) using *H. vulgaris* AEP as the reference and *H. vulgaris*

598 AEP.JNIG as the query, and the results were visualized with plotsr (Goel & Schneeberger
599 2022). Gene density was calculated based on *H. vulgaris* AEP gene models. Centromere
600 positions were estimated from the repeat annotations identified by EDTA v2.2.2 (Ou et al.
601 2019) and processed with quarTeT (v1.2.5) (Lin et al. 2023), while repeat density was
602 calculated using the output from Earl Grey (v6.3.3) (Baril et al. 2024).
603

604 Transposable element prediction

605 Transposable elements (TEs) in the *H. vulgaris* AEP and AEP. JNIG genomes were identified
606 using Earl Grey v6.3.3 (Baril et al. 2024).
607

608 Methylation landscape analysis

609 The distribution of DNA methylation (mCpG) was detected from Nanopore raw signal data
610 using Dorado (v0.7.1) (Oxford Nanopore Technologies: Dorado.
611 https://github.com/nanoporetech/dorado.) and Modkit (v0.4.3) (Oxford Nanopore
612 Technologies: Modkit. https://github.com/nanoporetech/modkit.). In Fig 3A the methylation
613 frequency values output by Modkit were visualized using karyoploteR v1.32.0 (Gel & Serra
614 2017). Based on the gene model of *H. vulgaris* AEP.JNIG and the genomic locations of each
615 TE class detected by Earl Grey v6.3.3 (Baril et al. 2024), mCpG/CpG ratios calculated for each
616 genomic region using the “modkit stats” command in Fig 3B and 3C. Promoter regions were
617 defined as the 2,000 bp upstream of the transcription start site (TSS), and regions that were
618 neither within genes nor promoters were defined as intergenic regions. In Fig 3D and 3E,
619 conserved genes were defined as orthologs shared across 33 species based on the output
620 from Orthofinder. Fig 3F displays a scatter plot in which the x-axis represents the gene body–
621 specific mCpG/CpG ratios obtained from “modkit stats”, and the y-axis shows the
622 corresponding RNA-seq expression levels (logTPM). The conserved genes shown in Fig 3E
623 correspond to orthologs shared among all 33 species analyzed in Fig 5.
624

625 ## Spurious transcription analysis

626 Paired-end RNA-seq reads were aligned to the *H. vulgaris* AEP.JNIG genome using HISAT2
627 v2.2.1 (Kim et al. 2019) with default parameters. The resulting SAM files were converted to
628 BAM format and sorted using SAMtools v1.23 (Danecek et al. 2021). Base-level read coverage
629 across the genome was calculated using bedtools2 v2.31.1 (Quinlan & Hall 2010) with the
630 command “bedtools genomecov -d”. Exon-level mean coverage was then calculated by
631 mapping the exon BED file to the per-base coverage BED file using “bedtools map -o mean”.
632 For each gene, exon-level coverage values were normalized by dividing by the coverage of
633 the first exon (exon 1) of the same gene, ratios were log-transformed using the natural
634 logarithm for downstream comparisons. Exons with zero or undefined coverage values were
635 excluded. Genes were further annotated based on two criteria: DNA methylation status and
636 evolutionary conservation. Methylated genes were defined as genes with mCpG/CpG ratios
637 greater than 60%, whereas conserved genes were defined as genes belonging to orthogroups
638 shared across all 33 species. This produced four groups: conserved–hypermethylated,
639 conserved–hypomethylated, nonconserved–hypermethylated, and nonconserved–
640 hypomethylated. For each exon position (exons 1–6), the mean log-transformed coverage ratio
641 was calculated separately for each of the four gene groups.

642

643 ## Conservation of aging related orthologous comparative Analysis of
644 ## Orthologous Groups in *Hydra oligactis* and *Hydra vulgaris*

645 Orthologs were identified using OrthoFinder(v2.5.5) (Emms & Kelly 2019). Protein sequences
646 from 33 species were used as input data, and the topology of the phylogenetic tree was based
647 on the study by (DeBiasse et al. 2022; Schultz et al. 2023), which was then confirmed by our
648 own phylogenomic analysis. Details are provided on the official GitHub repository
649 (https://github.com/nojiruko/H_vulgaris_AEP.JNIG). Each orthogroup was annotated using
650 human protein sequences as references, and the correspondence table of gene name and
651 Protein stable ID obtained from Ensembl BioMart

652 (https://www.ensembl.org/info/data/biomart/index.html) was used. Additionally, we compared
653 the presence or absence of orthogroups between *Hydra oligactis* and *Hydra vulgaris, made a*
654 *list of ortholog groups differentially present,* and performed GO, KEGG, and Reactome
655 enrichment analyses. Enrichment analyses were conducted using ClusterProfiler v4.14.4 (Wu
656 et al. 2021) and ReactomePA v1.50.0 (Yu & He 2016), with the human database referenced
657 from org.Hs.eg.db v3.20.0 (Carlson et al. 2019).
658 Genes involved in aging-related pathways were determined based on López-Otín, C. *et al*.
659 (2023) (López-Otín et al. 2023). First, we manually compiled a list of aging-related genes based
660 on the *Hallmarks of Aging* framework (López-Otín et al. 2023). Specifically, genes explicitly
661 mentioned in the original paper were directly included. For hallmarks described at the level of
662 signaling pathways rather than individual genes, constituent genes were identified primarily
663 using the KEGG Pathway database (Kanehisa & Goto 2000). When pathways were not clearly
664 defined or fully represented in KEGG, additional genes were identified using the Gene
665 Ontology (GO) database and relevant literature. The complete list of genes is summarized in
666 S1 Table. For orthogroups showing differential conservation between *H. oligactis* and *H.*
667 *vulgaris*, we performed BLASTP (protein query vs protein database) searches to verify whether
668 homologous sequences were present in the proteome of the other species, using an e-value
669 threshold of 10^{-5}. Specifically, for orthogroups identified in *H. oligactis* but absent in *H. vulgaris*,
670 *H. oligactis* protein sequences were used as BLASTP queries against the locally created
671 proteome databases of *H. vulgaris* AEP and *H. vulgaris* AEP.JNIG. Conversely, for ortholog
672 groups found in *H. vulgaris* but absent in *H.oligactis*, *H. vulgaris* AEP.JNIG protein sequences
673 were queried against the locally created proteome database of *H. oligactis*. Orthogroups
674 present in at least one *H. vulgaris* strain but not detected in *H. oligactis* were defined as *H.*
675 *vulgaris*-unique. Orthogroups absent from all three *H. vulgaris* strains but present in *H. oligactis*
676 were defined as *H. oligactis*-unique.
677

678 # Data Availability

679 The genome assembly of *H. vulgaris* AEP.JNIG and the associated sequence reads are
680 available at GenBank under BioProject PRJNA1413991 (BioSample SAMN54890366). Raw
681 Nanopore reads, Illumina genomic reads, and Illumina Hi-C reads have been deposited in
682 the Sequence Read Archive (SRA) under accession numbers SRR37338222,
683 SRR37338221, and SRR37338220, respectively.
684

685 # Acknowledgements

686 We thank the lab members at DBSB Juntendo, particularly Y. Kumagai, Q. Su, K. Suzuki, A.
687 Kanemitsu, K. Matsumoto, Y. Saito, and M. Ota for maintaining *Hydra* polyps and supporting
688 the related experiments; K. Otomo for supporting chromosome imaging and karyotyping; Y.
689 Wada and T.K. Suzuki for supporting computational resources. We also thank GeneBay, Inc.
690 (Yokohama, Japan) for sequencing the *H. vulgaris* AEP.JNIG genome; K. Ikeo (National
691 Institute of Genetics, Japan) for maintaining and providing Japanese hydra resource at NIG,
692 S. Kuraku (National Institute of Genetics, Japan), R. Nakato (The University of Tokyo, Tokyo,
693 Japan), and A. Watanabe (CyberomiX Inc., Kyoto, Japan) for fruitful suggestions about
694 genome assembly and analysis. We would also like to thank S. Hamada, K. Hamada, and O.
695 Koizumi (Fukuoka Women's University, Fukuoka, Japan) for his/her essential guidance in
696 establishing *Hydra* in our laboratory and for providing protocols for its successful husbandry.
697 We also extend our gratitude to Y. Takai (Institute for Advanced Biosciences, Keio University)
698 for her technical assistance in Hi-C library preparation. We acknowledge the use of large
699 language models (ChatGPT and Gemini) for English language proofreading and stylistic
700 editing of the manuscript.
701 This work was supported by Japan Society for the Promotion of Science (JSPS) KAKENHI
702 Grant-in-Aid for Scientific Research (B) [grant number JP22H02824 to E.A.S.]; JSPS
703 KAKENHI for International Leading Research [grant number 23K20044 to E.A.S.]; Japan
704 Science and Technology agency (JST) Core Research for Evolutionary Science and

705 Technology (CREST) [grant number JPMJCR23B7 to E.A.S.]; Japan Agency for Medical
706 Research and Development (AMED) Project for Promotion of Cancer Research and
707 Therapeutic Evolution (P-PROMOTE) [grant number JP25ama221610 to E.A.S.]; Grants-in-
708 Aid from Nakatani Foundation [to E.A.S.]; JSPS KAKENHI Gran-in-Aid for Research Activity
709 Start-up [24K23213 to K.K.]; intramural Grant-in-Aid and Operating Costs Subsidies for Private
710 Universities [to E.A.S.]; research funds from the Yamagata Prefectural Government and
711 Tsuruoka City, Japan [to K.A.]; and the JSPS Overseas Research Fellowships [to T.K.].
712
713

714 Author contributions

715 E.A.S., K.N., K.K., and A.S. conceived and directed the study. H.S. maintained and provided
716 *H. vulgaris* AEP.JNIG resource. A.S. prepared the *H. vulgaris* AEP.JNIG genomic DNA and
717 RNA samples for sequencing. K.A. performed the Hi-C sequencing and analysis. K.K., T.K.,
718 K.K-N., and K.A. assembled the genome data. K.N. and K.K. performed the genomic
719 comparison between AEP.JNIG and AEP, whole-genome methylation analysis, and
720 comparative genome analysis. K.K. performed the karyotyping experiment. E.A.S., T.K., and
721 K.A. acquired funding. E.A.S., K.N., and K.K. mainly wrote the manuscript. All authors
722 discussed the results and commented on the manuscript.
723

724 COI disclosure

725 There is no COI of all the authors regarding this study.
726

Figure legends

Figure 1. Chromosome-level genome assembly of *H. vulgaris* AEP.JNIG.

(A) Picture of cultured *H. vulgaris* AEP.JNIG. Scale bar: 1 mm.

(B) Hi-C contact map, showing the 15 chromosome-level scaffolds of *H. vulgaris* AEP.JNIG. Darker red indicates genomic regions with higher contact frequencies, reflecting closer spatial proximity.

(C) Snail plot demonstrating the quality of our genome assembly.

(D) A representative image of metaphase chromosomes in *H. vulgaris* AEP.JNIG stained with DAPI and observed with a scanning disk confocal microscope. Scale bar: 5 μm.

(E) Histogram of karyotypes observed in this study. Chromosome counts show a distinct peak at 2n = 30. The absence of nuclei with 2n > 32 supports a chromosome number of 30, consistent with our and other genome assembly data.

(F) Karyogram of *Hydra vulgaris* AEP.JNIG showing 15 chromosomes.

Alt text: Composite figure with panels A–F illustrating the chromosome-level genome assembly and karyotype of *Hydra vulgaris* AEP.JNIG. A is a photograph of a cultured *H. vulgaris* polyp. B is a Hi-C contact map displayed in red and white colors showing 15 chromosome-scale scaffolds as strong diagonal blocks indicating high contact frequencies within chromosomes. C is a circular “snail plot” summarizing genome assembly statistics, including BUSCO completeness of 94.7%. D is a fluorescence microscopy image of metaphase chromosomes stained with DAPI, showing a diploid set of chromosomes. E is a histogram of chromosome counts with a clear peak at 2n = 30. F is a karyogram showing 15 chromosome pairs arranged from longest to shortest.

Figure 2. Comparison of *H. vulgaris* AEP and *H. vulgaris* AEP.JNIG genome assemblies

753 **(A)** Maximum likelihood phylogenetic analysis of a phylogenomic dataset constructed from 33
754 holozoan proteomes shows that *H. vulgaris* AEP and *H. vulgaris* AEP.JNIG form a
755 monophyletic clade. Nodes marked with black dots (●) indicate 100% ultrafast bootstrap
756 support, while other nodes were labelled with their corresponding ultrafast bootstrap support
757 values. The scale bar represents 0.5 amino acid substitutions per site.
758 **(B)** A dot-plot created with D-GENIES shows that the genome of *H. vulgaris* JNIG (on the y-
759 axis) aligns well with that of *H. vulgaris* AEP genome (Cazet et al. 2023) (on the x-axis,
760 indicating a high genomic similarity between the two strains. The color indicates the identity
761 score as shown in the legend on the right.
762 **(C)** Gene synteny plot generated by GENESPACE showing conserved synteny of orthologous
763 genes between the AEP and AEP.JNIG genomes. The AEP.JNIG chromosomes are shown at
764 the top and the AEP chromosomes at the bottom. Chromosome numbering and sequence
765 strand orientation were standardized to match previously published genome assemblies.
766 **(D)** Visualization of syntenic regions and structural rearrangements between chromosomes
767 from *H. vulgaris* AEP and *H. vulgaris* AEP.JNIG genomes. Annotation tracks for four genomic
768 features (number of SNPs, genes, repetitive regions and centromeres) were included using
769 optional parameters above the lines showing the chromosomes of each strain. In the
770 annotation tracks, higher values indicate greater density.
771
772 **Alt text**: Composite figure with panels A–D comparing the genome assemblies of *Hydra*
773 *vulgaris* AEP and AEP.JNIG. A is a phylogenetic tree based on a dataset of 33 holozoan
774 proteomes showing that AEP and AEP.JNIG form a monophyletic clade, with bootstrap support
775 values indicated at the nodes. B is a genome alignment dot plot with the AEP genome on the
776 x-axis and the AEP.JNIG genome on the y-axis, showing strong diagonal alignments indicating
777 high genomic similarity. C is a gene synteny plot showing conserved syntenic relationships
778 between chromosomes of the AEP.JNIG genome and the AEP genome. D is a visualization of
779 syntenic regions and structural rearrangements between chromosomes of the two genomes,

780 with annotation tracks showing the density of SNPs, genes, repetitive regions, and
781 centromeres.
782
783 **Figure 3.** DNA Methylation within Gene Bodies Contributes to Transcriptional Stability
784 (Homeostasis) of Conserved Genes in *H. vulgaris*
785 **(A)** Global pattern of DNA methylation in *H. vulgaris* AEP.JNIG genome showed that
786 centromeric regions had low DNA methylation. The top track shows the methylation frequency
787 (the probability that each base is methylated). The second track indicates the predicted
788 centromere positions in red, which was identified by quarTeT (v 1.2.5) (Lin et al. 2023) .
789 **(B)** Methylation level across each genomic region. The gene region was further analyzed by
790 dividing exon and intron regions.
791 **(C)** Methylation level of each repetitive region. The colors correspond to those used in Fig 3A.
792 **(D)** Boxplots of gene-body methylation levels in conserved (purple) and other (gray) genes.
793 Differences were significant (Mann-Whitney U test, $p < 1\times10^{-4}$).
794 **(E)** Boxplots showing expression levels in the conserved (purple) and the other (gray) genes.
795 Differences were significant (Mann-Whitney U test, $p < 1\times10^{-4}$).
796 **(F)** Relationship between DNA methylation levels in gene bodies and gene expression. Purple
797 dots indicate conserved genes, and gray dots indicate other genes. Values indicate the
798 coefficient variation of each group's logTPM.
799 **(G)** Spurious transcription in gene bodies is significantly lower in the conserved and highly
800 methylated genes (normalized by exon length for the logFC). The y axis shows the natural
801 logarithm of the coverage fold change (FC) of exons 1 to 6 versus exon 1, with lower values
802 meaning less spurious transcription.
803
804 **Alt text**: Composite figure with panels A–G illustrating that gene-body DNA methylation is
805 associated with conserved genes and contributes to transcriptional stability in *Hydra vulgaris*.
806 A is a genome-wide visualization of DNA methylation frequency showing low methylation levels
807 in centromeric regions, with predicted centromere positions indicated along the chromosomes.

808 B is a comparison of methylation levels across genomic regions, showing gene-body
809 methylation patterns in exons and introns. C is a comparison of methylation levels across
810 different classes of repetitive elements. D is a boxplot comparing gene-body methylation levels
811 in conserved genes (purple) and other genes (gray), showing significantly higher methylation
812 in conserved genes. E is a boxplot comparing gene expression levels between conserved and
813 other genes, with conserved genes showing higher expression. F is a scatter plot showing the
814 relationship between gene-body methylation and gene expression variability, with conserved
815 genes (purple) showing more stable expression patterns. G is a plot measuring spurious
816 transcription within gene bodies, indicating that conserved and highly methylated genes show
817 lower levels of spurious transcription.
818
819 **Figure 4.** Conservation of aging-related orthologs in *Hydra*
820 **(A)** Conservation of orthologs related to the *Hallmarks of Aging*. The vertical axis lists
821 representative aging-related pathways (defined as KEGG terms). Each pathway was extracted
822 from the 12 hallmarks of aging proposed in 2023 (López-Otín et al. 2023). Genes including
823 each pathway were listed as described in the Materials and Methods and summarized in S4
824 Table. Numbers inside the circles indicate the number of conserved orthologs, and the circle
825 size corresponds to the number of conserved orthologs. The *Hydra* genus is highlighted in
826 purple.
827 (**B)** Bar plot showing the total number of conserved aging-related orthologs identified in each
828 species. The y-axis represents the number of orthogroups. The *Hydra* genus is highlighted in
829 purple.
830
831 **Alt text**: Composite figure with panels A–B illustrating the conservation of aging-related genes
832 across species. A is a bubble plot showing the conservation of orthologs associated with aging-
833 related pathways defined by the Hallmarks of Aging. Pathways are listed on the vertical axis,
834 and circle size indicates the number of conserved orthologs. The *Hydra* genus is highlighted
835 in purple. B is a bar plot comparing the total number of conserved aging-related orthologs

836 identified in each species, with *Hydra* species highlighted in purple, showing *Hydra* species
837 have more aging-related orthologs than flies or nematodes.
838
839 **Figure 5.** Comparative Analysis between *H. vulgaris* and *H. oligactis*
840 (**A)** Schematic diagram illustrating the method used to identify *H. vulgaris*- and *H. oligactis*-
841 specific aging-related genes. A part of figure materials was prepared using biorender.com
842 (permission No. JK2989SSUB).
843 **(B)** Twelve genes are present in *H. oligactis* but absent in *H. vulgaris*, whereas five genes are
844 present in *H. vulgaris* but not in *H. oligactis*. A total of 7,684 genes is shared between the two
845 species. Note that only genes annotated based on homology to human genes were included
846 in this analysis. We referred to the original article (López-Otín et al. 2023) for the factors of
847 *Hallmarks of Aging*. Icons for the analyzed species were referred to from PhyloPic
848 (https://www.phylopic.org/) under Creative Commons license.
849 **(C)** Gene Ontology (GO) enrichment analysis for biological processes was performed using 65
850 orthogroups present in *H. oligactis* but absent in *H. vulgaris*. The y-axis shows the top 13
851 significantly enriched GO terms. The x-axis represents the gene ratio, defined as the proportion
852 of input genes associated with each term. The color intensity of the bars indicates the adjusted
853 p-value (p.adj), with darker colors representing higher statistical significance.
854 **(D)** The same analysis was performed for 33 orthogroups present in *H. vulgaris* but absent in
855 *H. oligactis*.
856 **(E)** Orthogroups specific to *H. oligactis* (65 OGs) and *H. vulgaris* (33 OGs) were screened
857 across 33 organisms to identify those involved in aging-related pathways defined in this study
858 based on the *Hallmarks of Aging*. The final list of genes is shown on the x-axis. The y-axis
859 represents the species included in this analysis, ordered by phylogeny. Dots indicate the
860 presence of the corresponding gene in each organism, while blanks indicate absence. The
861 colors at the top represent the *Hallmarks of Aging* categories assigned to each gene. The
862 names of each category and the specific terms included in them are shown on the right.
863

Alt text: Composite figure with panels A–E illustrating comparative analyses of aging-related genes between *Hydra vulgaris* and *Hydra oligactis*. A is a schematic diagram showing the analytical workflow used to identify species-specific aging-related genes in the two *Hydra* species. B is a diagram summarizing shared and species-specific genes, showing 7,684 genes shared between the species, 65 genes present only in *H. oligactis* (12 of them aging-related), and 33 genes present only in *H. vulgaris* (5 of them aging-related). C is a bar plot showing Gene Ontology enrichment results for orthogroups present in *H. oligactis* but absent in *H. vulgaris*, with the y-axis listing enriched biological processes and the x-axis showing the gene ratio. D is a similar Gene Ontology enrichment plot for orthogroups present in *H. vulgaris* but absent in *H. oligactis*. E is a presence–absence matrix showing the distribution of species-specific orthogroups across 33 organisms, with dots indicating gene presence and blank cells indicating absence.

Table 1. Summary of the genome sequencing

Table 2. Assembly statistics of the *H. vulgaris* genome.

Table 3. BUSCO score from the genome assembly

Table 4. BUSCO score from the gene model

Table 5. Composition of TEs identified by Earl Grey in the genomes of *H. vulgaris* AEP. JNIG.

Table 6. Composition of TEs identified by Earl Grey in the genomes of *H. vulgaris* AEP.

885 # Supplementary figure legends

886

887 **Figure S1.** Karyotyping Data

888 Karyotype of *H. vulgaris* AEP.JNIG. A total of 30 nuclei samples were examined, and the
889 number of chromosomes were counted. The number in yellow indicates the karyotype (2n) of
890 the corresponding image. Scale Bar = 5 μm.

891

892 **Figure S2**. Molecular phylogenetic trees of *Hydra* strains based on mitochondrial and nuclear
893 genes.

894 Trees were reconstructed using IQ-TREE2 with best-fitting models selected by ModelFinder
895 (MFP option). Branch support was evaluated using the ultrafast bootstrap method. *H. vulgaris*
896 AEP.JNIG and *H. vulgaris* AEP sequences are shown in red and blue, respectively. *H. vulgaris*
897 AEP sequences from (Kawaida et al. 2010) ("*Hydra vulgaris* AEP") and those derived from the
898 *H. vulgaris* AEP genome (Cazet et al. 2023) ("*Hydra vulgaris* AEP_genome") (Cazet et al.
899 2023)are indicated by distinct labels. Numbers at nodes represent ultrafast bootstrap support
900 values. Scale bars represent 0.05 nucleotide substitutions per site.

901 **(A)** A molecular phylogenetic tree based on *COI*.

902 **(B)** A molecular phylogenetic tree based on *COII*.

903 **(C)** A molecular phylogenetic tree based on *EF1α*.

904 **(D)** A molecular phylogenetic tree based on *Cnnos1*.

905

906 **Figure S3.** Comparison of the transposable elements between AEP.JNIG and AEP

907 **(A)** Quantitatively assessed structural and sequence variation identified by SyRI. Structural
908 and sequence variations were identified using SyRI with the AEP assembly as the reference
909 and the AEP.JNIG assembly as the query. Sequence lengths are shown in megabases (Mb).

910 **(B)** Transposable elements (TEs) landscape showing the distribution of repetitive elements in
911 the genome based on Kimura distance. The x-axis represents the Kimura 2-Parameter

912 distance, where lower scores indicate more recent insertions into the genome, while higher
913 scores correspond to older insertions. The y-axis shows the total number of base pairs
914 contributed by each class of transposable elements at each divergence level. The pie chart
915 shows the components and proportions of repetitive sequences in the *H. vulgaris* AEP.JNIG
916 genome.
917 **(C)** Transposable elements (TEs) landscape of *H. vulgaris* AEP, as in B.
918
919 **Figure S4.** The methylation level of *H. vulgaris* AEP depicted as box plots, related to Fig 3B
920 and 3C.
921 **(A)** The box plot of the methylation level across each genomic region, corresponding to Fig 3B.
922 **(B)** The box plot of the methylation level of each repetitive region, corresponding to Fig 3C.
923
924 **Figure S5.** Conservation of aging-related orthologs in *Hydra*, related to Fig 4.
925 Conservation of orthologs related to the *Hallmarks of Aging*, as in Fig 4A. The vertical axis lists
926 representative aging-related pathways (defined as GO terms).
927
928 **Figure S6.** Enrichment analysis results of *H. oligactis*-specific and *H. vulgaris*-specific
929 orthogroups using multiple annotation databases.
930 **(A)** Results of Gene Ontology – Cellular Component (GO-CC) terms.
931 **(B)** Results of Gene Ontology – Molecular Function (GO-MF) terms.
932 **(C)** Results of KEGG pathway terms.
933 **(D)** Results of Reactome pathway terms.
934 In each panel, *H. oligactis*-specific orthogroups are shown on the left, whereas *H. vulgaris*-
935 specific orthogroups are shown on the right.
936

Table S1. Complete list of aging-related genes included in this study. This table provides the full list of aging-related genes identified in this study. For each gene, the second column indicates the associated aging-related pathways, and the third column classifies the gene according to the *Hallmarks of Aging*.

Table S2. Summary of the conservation of aging-related genes for each pathway, related to Fig 4 and Fig S5**.**

This table provides an overview of the conservation of aging-related genes identified through GO, KEGG and literature-based searches based on the *Hallmarks of Aging* (López-Otín et al. 2023). The first column lists aging-related pathways, and the subsequent columns show the number of conserved orthogroups in each species.

Table S3. Lists of orthogroups specific to *H. oligactis* (65 OGs) and *H. vulgaris* (33 OGs). Each row corresponds to one orthogroup. The second column shows annotations with the corresponding human genes. The third column shows the pathway annotation for each aging-related gene.

Table S4. Enrichment analysis of the 33 *H. vulgaris* specific orthogroups and 65 *H. oligactis* specific orthogroups. Significantly enriched terms (adjusted p-value < 0.05) are shown. Ten files are provided, corresponding to all combinations of species-specific gene sets (*H. oligactis*-specific and *H. vulgaris*-specific) and annotation databases (GO biological process, GO cellular component, GO molecular function, KEGG, and Reactome).

References

Ali RH, Bogusz M, Whelan S. 2019. Identifying clusters of high confidence homologies in multiple sequence alignments. Mol. Biol. Evol. 36:2340–2351.

Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ. 1990. Basic local alignment search tool. J. Mol. Biol. 215:403–410.

Baril T, Galbraith J, Hayward A. 2024. Earl Grey: A fully automated user-friendly transposable element annotation and analysis pipeline. Mol. Biol. Evol. 41. doi: 10.1093/molbev/msae068.

Barve A, Galande AA, Ghaskadbi SS, Ghaskadbi S. 2021. DNA repair repertoire of the enigmatic hydra. Front. Genet. 12:670695.

Boehm A-M et al. 2012. FoxO is a critical regulator of stem cell maintenance in immortal Hydra. Proc. Natl. Acad. Sci. U. S. A. 109:19697–19702.

Bosch TC, Krylow SM, Bode HR, Steele RE. 1988. Thermotolerance and synthesis of heat shock proteins: these responses are present in Hydra attenuata but absent in Hydra oligactis. Proc. Natl. Acad. Sci. U. S. A. 85:7927–7931.

Brůna T, Hoff KJ, Lomsadze A, Stanke M, Borodovsky M. 2021. BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genom. Bioinform. 3:lqaa108.

Cabanettes F, Klopp C. 2018. D-GENIES: dot plot large genomes in an interactive, efficient and simple way. PeerJ. 6:e4958.

Capella-Gutiérrez S, Silla-Martínez JM, Gabaldón T. 2009. trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 25:1972–1973.

Carlson M, Falcon S, Pages H, Li N. 2019. org. Hs. eg. db: Genome wide annotation for Human. R package version. 3:3.

Cazet JF et al. 2023. A chromosome-scale epigenetic map of the Hydra genome reveals conserved regulators of cell state. Genome Res. 33:283–298.

Chapman JA et al. 2010. The dynamic genome of Hydra. Nature. 464:592–596.

Chebib J, Jackson BC, López-Cortegano E, Tautz D, Keightley PD. 2021. Inbred lab mice are not isogenic: genetic variation within inbred strains used to infer the mutation rate per nucleotide site. Heredity (Edinb.). 126:107–116.

Chen S. 2023. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta. 2:e107.

Chera S, Buzgariu W, Ghila L, Galliot B. 2009. Autophagy in Hydra: a response to starvation and stress in early animal evolution. Biochim. Biophys. Acta. 1793:1432–1443.

Danecek P et al. 2021. Twelve years of SAMtools and BCFtools. Gigascience. 10. doi: 10.1093/gigascience/giab008.

David CN, Campbell RD. 1972. Cell cycle kinetics and development of Hydra attenuata. I. Epithelial cells. J. Cell Sci. 11:557–568.

998 De Coster W, Rademakers R. 2023. NanoPack2: population-scale evaluation of long-read
999 sequencing data. Bioinformatics. 39:btad311.

1000 Debès C et al. 2023. Ageing-associated changes in transcriptional elongation influence
1001 longevity. Nature. 616:814–821.

1002 DeBiasse MB et al. 2022. A cnidarian phylogenomic tree fitted with hundreds of 18S leaves.
1003 bioRxiv. doi: 10.1101/2022.10.03.510641.

1004 Dixon G, Liao Y, Bay LK, Matz MV. 2018. Role of gene body methylation in acclimatization
1005 and adaptation in a basal metazoan. Proc. Natl. Acad. Sci. U. S. A. 115:13342–13346.

1006 Doshi R, Kinnear E, Chatterjee S, Guha P, Liu Q. 2025. Reliable investigation of DNA
1007 methylation using Oxford nanopore technologies. Sci. Rep. 15:15900.

1008 Dudchenko O et al. 2017. De novo assembly of the Aedes aegypti genome using Hi-C yields
1009 chromosome-length scaffolds. Science. 356:92–95.

1010 Durand NC et al. 2016. Juicer provides a one-click system for analyzing loop-resolution Hi-C
1011 experiments. Cell Syst. 3:95–98.

1012 Emms DM, Kelly S. 2019. OrthoFinder: phylogenetic orthology inference for comparative
1013 genomics. Genome Biol. 20:238.

1014 Farslow JC et al. 2015. Rapid Increase in frequency of gene copy-number variants during
1015 experimental evolution in Caenorhabditis elegans. BMC Genomics. 16:1044.

1016 Finch CE, Pike MC, Witten M. 1990. Slow mortality rate accelerations during aging in some
1017 animals approximate that of humans. Science. 249:902–905.

1018 Flynn JM et al. 2020. RepeatModeler2 for automated genomic discovery of transposable
1019 element families. Proc. Natl. Acad. Sci. U. S. A. 117:9451–9457.

1020 Galliot B. 2012. Hydra, a fruitful model system for 270 years. Int. J. Dev. Biol. 56:411–423.

1021 Gel B, Serra E. 2017. karyoploteR: an R/Bioconductor package to plot customizable
1022 genomes displaying arbitrary data. Bioinformatics. 33:3088–3090.

1023 Goel M, Schneeberger K. 2022. Plotsr: Visualizing structural similarities and rearrangements
1024 between multiple genomes. Bioinformatics. 38:2922–2926.

1025 Goel M, Sun H, Jiao W-B, Schneeberger K. 2019. SyRI: finding genomic rearrangements
1026 and local sequence differences from whole-genome assemblies. Genome Biol. 20:277.

1027 Gufler S et al. 2018. β-Catenin acts in a position-independent regeneration response in the
1028 simple eumetazoan Hydra. Dev. Biol. 433:310–323.

1029 Guo L et al. 2018. An adaptable chromosome preparation methodology for use in
1030 invertebrate research organisms. BMC Biol. 16:25.

1031 Haas BJ. https://github.com/transdecoder/transdecoder.

1032 Hamilton LK et al. 2022. Stearoyl-CoA Desaturase inhibition reverses immune, synaptic and
1033 cognitive impairments in an Alzheimer's disease mouse model. Nat. Commun. 13:2061.

1034 Hamilton WD. 1966. The moulding of senescence by natural selection. J. Theor. Biol. 12:12–
1035 45.

1036 Hemmrich G et al. 2012. Molecular signatures of the three stem cell lineages in hydra and
1037 the emergence of stem cell function at the base of multicellularity. Mol. Biol. Evol. 29:3267–
1038 3280.

1039 Holstein TW. 2023. The Hydra stem cell system - Revisited. Cells Dev. 174:203846.

1040 Imai S-I. 2016. The NAD World 2.0: the importance of the inter-tissue communication
1041 mediated by NAMPT/NAD+/SIRT1 in mammalian aging and longevity control. NPJ Syst. Biol.
1042 Appl. 2:16018.

1043 Imai S-I, Guarente L. 2014. NAD+ and sirtuins in aging and disease. Trends Cell Biol.
1044 24:464–471.

1045 Jones OR et al. 2014. Diversity of ageing across the tree of life. Nature. 505:169–173.

1046 Kanehisa M, Goto S. 2000. KEGG: kyoto encyclopedia of genes and genomes. Nucleic
1047 Acids Res. 28:27–30.

1048 Katoh K, Standley DM. 2013. MAFFT multiple sequence alignment software version 7:
1049 improvements in performance and usability. Mol. Biol. Evol. 30:772–780.

1050 Kawaida H, Shimizu H, Fujisawa T, Tachida H, Kobayakawa Y. 2010. Molecular
1051 phylogenetic study in genus Hydra. Gene. 468:30–40.

1052 Kershenbaum S et al. 2025. Variation in phenotype, genotype, and somatic diversity among
1053 asexual Schmidtea mediterranea planarians. iScience. 28:113035.

1054 Khalturin K, Hemmrich G, Fraune S, Augustin R, Bosch TCG. 2009. More than just orphans:
1055 are taxonomically-restricted genes important in evolution? Trends Genet. 25:404–413.

1056 Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment
1057 and genotyping with HISAT2 and HISAT-genotype. Nat. Biotechnol. 37:907–915.

1058 Kirkwood TB. 1977. Evolution of ageing. Nature. 270:301–304.

1059 Klimovich A, Wittlieb J, Bosch TCG. 2019. Transgenesis in Hydra to characterize gene
1060 function and visualize cell behavior. Nat. Protoc. 14:2069–2090.

1061 Kon-Nanjo K et al. 2025. The dynamic genomes of Hydra and the anciently active repeat
1062 complement of animal chromosomes. Genome Biol. 26:186.

1063 Kriventseva EV et al. 2019. OrthoDB v10: sampling the diversity of animal, plant, fungal,
1064 protist, bacterial and viral genomes for evolutionary and functional annotations of orthologs.
1065 Nucleic Acids Res. 47:D807–D811.

1066 Kuro-o M. 2009. Klotho and aging. Biochim. Biophys. Acta. 1790:1049–1058.

1067 Lemoine M. 2021. The evolution of the hallmarks of aging. Front. Genet. 12:693071.

1068 Li CZ et al. 2024. Epigenetic predictors of species maximum life span and other life-history
1069 traits in mammals. Sci. Adv. 10:eadm7273.

1070 Li H. 2018. Minimap2: pairwise alignment for nucleotide sequences. Bioinformatics.
1071 34:3094–3100.

1072 Li Y et al. 2018. DNA methylation regulates transcriptional homeostasis of algal
1073 endosymbiosis in the coral model Aiptasia. Sci. Adv. 4:eaat2142.

1074 Lin Y et al. 2023. quarTeT: a telomere-to-telomere toolkit for gap-free genome assembly and
1075 centromeric repeat identification. Hortic. Res. 10:uhad127.

1076 Liu H et al. 2024. A taxon-rich and genome-scale phylogeny of Opisthokonta. PLoS Biol.
1077 22:e3002794.

1078 Lombard DB et al. 2005. DNA repair, genome stability, and aging. Cell. 120:497–512.

1079 Loomis WF, Lenhoff HM. 1956. Growth and sexual differentiation of hydra in mass culture. J.
1080 Exp. Zool. 132:555–573.

1081 López-Otín C, Blasco MA, Partridge L, Serrano M, Kroemer G. 2023. Hallmarks of aging: An
1082 expanding universe. Cell. 186:243–278.

1083 Lovell JT et al. 2022. GENESPACE tracks regions of interest and gene copy number
1084 variation across multiple genomes. Elife. 11. doi: 10.7554/eLife.78526.

1085 Manni M, Berkeley MR, Seppey M, Simão FA, Zdobnov EM. 2021. BUSCO update: Novel
1086 and streamlined workflows along with broader and deeper phylogenetic coverage for scoring
1087 of eukaryotic, prokaryotic, and viral genomes. Mol. Biol. Evol. 38:4647–4654.

1088 Martínez DE. 1998. Mortality patterns suggest lack of senescence in hydra. Exp. Gerontol.
1089 33:217–225.

1090 Martínez DE et al. 2010. Phylogeny and biogeography of Hydra (Cnidaria: Hydridae) using
1091 mitochondrial and nuclear DNA sequences. Mol. Phylogenet. Evol. 57:403–410.

1092 McNamara RK, Liu Y, Jandacek R, Rider T, Tso P. 2008. The aging human orbitofrontal
1093 cortex: decreasing polyunsaturated fatty acid composition and associated increases in
1094 lipogenic gene expression and stearoyl-CoA desaturase activity. Prostaglandins Leukot.
1095 Essent. Fatty Acids. 78:293–304.

1096 Minh BQ et al. 2020. IQ-TREE 2: New models and efficient methods for phylogenetic
1097 inference in the genomic era. Mol. Biol. Evol. 37:1530–1534.

1098 Miwa S, Kashyap S, Chini E, von Zglinicki T. 2022. Mitochondrial dysfunction in cell
1099 senescence and aging. J. Clin. Invest. 132. doi: 10.1172/JCI158447.

1100 Nabeshima Y-I. 2002. Klotho: a fundamental regulator of aging. Ageing Res. Rev. 1:627–
1101 638.

1102 Nie F et al. 2024. De novo diploid genome assembly using long noisy reads. Nat. Commun.
1103 15:2964.

1104 Otomo K et al. 2015. Multi-point scanning two-photon excitation microscopy by utilizing a
1105 high-peak-power 1042-nm laser. Anal. Sci. 31:307–313.

1106 Ou S et al. 2019. Benchmarking transposable element annotation methods for creation of a
1107 streamlined, comprehensive pipeline. Genome Biol. 20:275.

1108 Papantonis A, Antebi A, Partridge L, Beyer A. 2025. Age-associated changes in
1109 transcriptional elongation and their effects on homeostasis. Trends Cell Biol. 35:645–650.

1110 Pascual-Torner M et al. 2022. Comparative genomics of mortal and immortal cnidarians
1111 unveils novel keys behind rejuvenation. Proc. Natl. Acad. Sci. U. S. A. 119:e2118763119.

1112 Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic

1113 features. Bioinformatics. 26:841–842.

1114 Rando TA, Jones DL. 2021. Regeneration, rejuvenation, and replacement: Turning back the
1115 clock on tissue aging. Cold Spring Harb. Perspect. Biol. 13:a040907.

1116 Rechsteiner C, Morandini F, Kim SJ, Seluanov A, Gorbunova V. 2025. Unlocking longevity
1117 through the comparative biology of aging. Nat. Aging. 5:1686–1703.

1118 Sahm A et al. 2024. Hydra has mammal-like mutation rates facilitating fast adaptation
1119 despite its nonaging phenotype. Genome Res. 34:2217–2228.

1120 Sarda S, Zeng J, Hunt BG, Yi SV. 2012. The evolution of invertebrate gene body
1121 methylation. Mol. Biol. Evol. 29:1907–1916.

1122 Schaible R et al. 2015. Constant mortality and fertility over age in Hydra. Proc. Natl. Acad.
1123 Sci. U. S. A. 112:15701–15706.

1124 Schultz DT et al. 2023. Ancient gene linkages support ctenophores as sister to other
1125 animals. Nature. 618:110–117.

1126 Servant N et al. 2015. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing.
1127 Genome Biol. 16:259.

1128 Shenhar B et al. 2026. Heritability of intrinsic human life span is about 50% when
1129 confounding factors are addressed. Science. 391:504–510.

1130 Shumate A, Salzberg SL. 2021. Liftoff: accurate mapping of gene annotations.
1131 Bioinformatics. 37:1639–1643.

1132 Sun S et al. 2020. Inducible aging in Hydra oligactis implicates sexual reproduction, loss of
1133 stem cells, and genome maintenance as major pathways. GeroScience. 42:1119–1132.

1134 Suzuki MM, Bird A. 2008. DNA methylation landscapes: provocative insights from
1135 epigenomics. Nat. Rev. Genet. 9:465–476.

1136 Tomczyk S et al. 2020. Deficient autophagy in epithelial stem cells drives aging in the
1137 freshwater cnidarian Hydra. Development. 147:dev177840.

1138 Tomczyk S, Fischer K, Austad S, Galliot B. 2015. Hydra, a powerful model for aging studies.
1139 Invertebr. Reprod. Dev. 59:11–16.

1140 Tyshkovskiy A et al. 2023. Distinct longevity mechanisms across and within species and their
1141 association with aging. Cell. 186:2929-2949.e20.

1142 Vogg MC, Galliot B, Tsiairis CD. 2019. Model systems for regeneration: Hydra. Development.
1143 146:dev177212.

1144 Wenger Y, Galliot B. 2013. Punctuated emergences of genetic and phenotypic innovations in
1145 eumetazoan, bilaterian, euteleostome, and hominidae ancestors. Genome Biol. Evol.
1146 5:1949–1968.

1147 Whelan S, Irisarri I, Burki F. 2018. PREQUAL: detecting non-homologous characters in sets
1148 of unaligned homologous sequences. Bioinformatics. 34:3929–3930.

1149 Wong WY et al. 2019. Expansion of a single transposable element family is associated with
1150 genome-size increase and radiation in the genus Hydra. Proc. Natl. Acad. Sci. U. S. A.
1151 116:22915–22917.

1152 Wu T et al. 2021. clusterProfiler 4.0: A universal enrichment tool for interpreting omics data.
1153 Innovation (Camb.). 2:100141.

1154 Xu G, Herzig M, Rotrekl V, Walter CA. 2008. Base excision repair, aging and health span.
1155 Mech. Ageing Dev. 129:366–382.

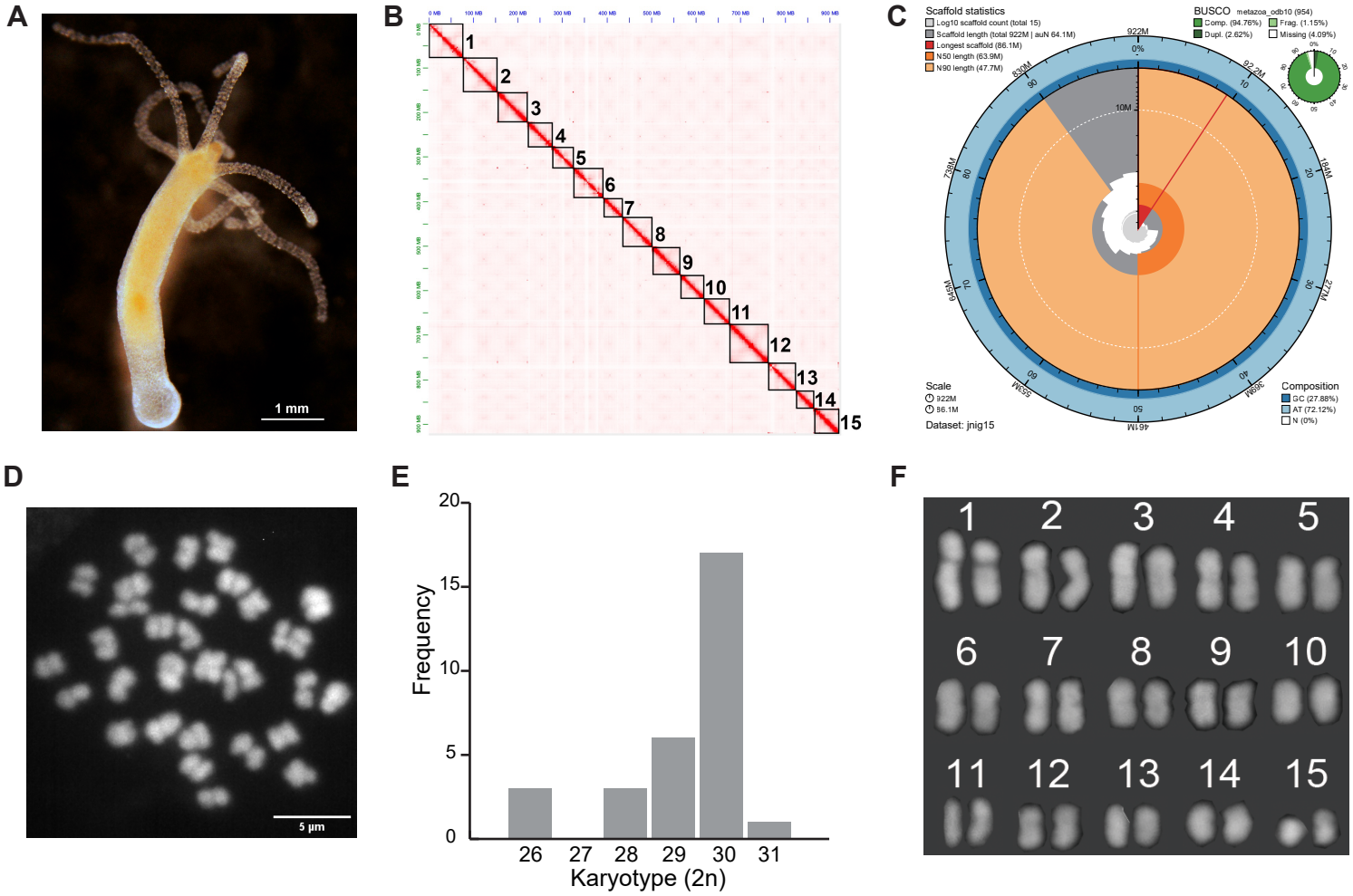
1156 Ying H et al. 2022. The role of DNA methylation in genome defense in Cnidaria and other
1157 invertebrates. Mol. Biol. Evol. 39:msac018.

1158 Yoshida K, Fujisawa T, Hwang JS, Ikeo K, Gojobori T. 2006. Degeneration after sexual
1159 differentiation in hydra and its relevance to the evolution of aging. Gene. 385:64–70.

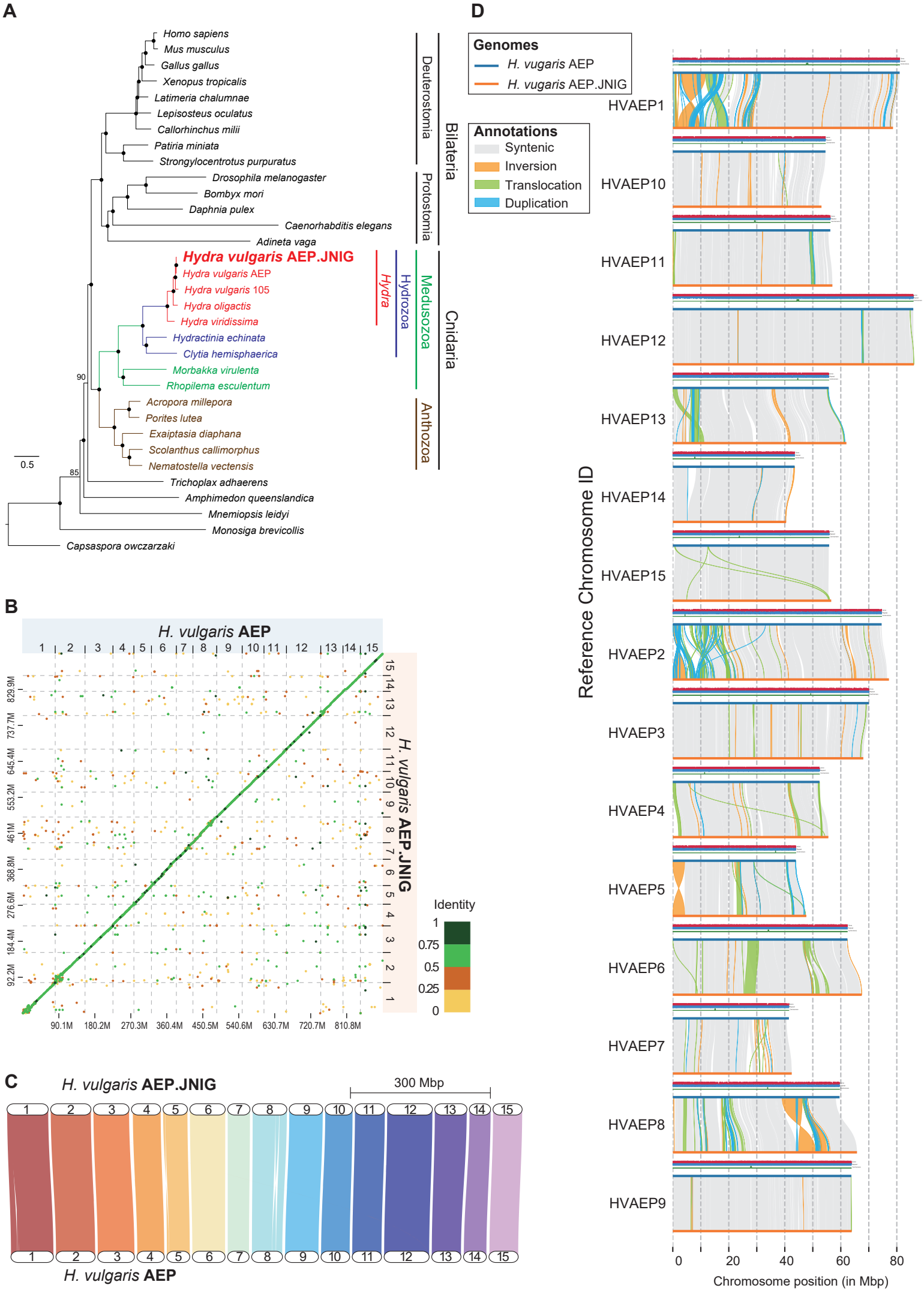
1160 Yu G, He Q-Y. 2016. ReactomePA: an R/Bioconductor package for reactome pathway
1161 analysis and visualization. Mol. Biosyst. 12:477–479.

1162 Zhang X, Jacobs D. 2022. A broad survey of gene body and repeat methylation in Cnidaria
1163 reveals a complex evolutionary history. Genome Biol. Evol. 14. doi: 10.1093/gbe/evab284.

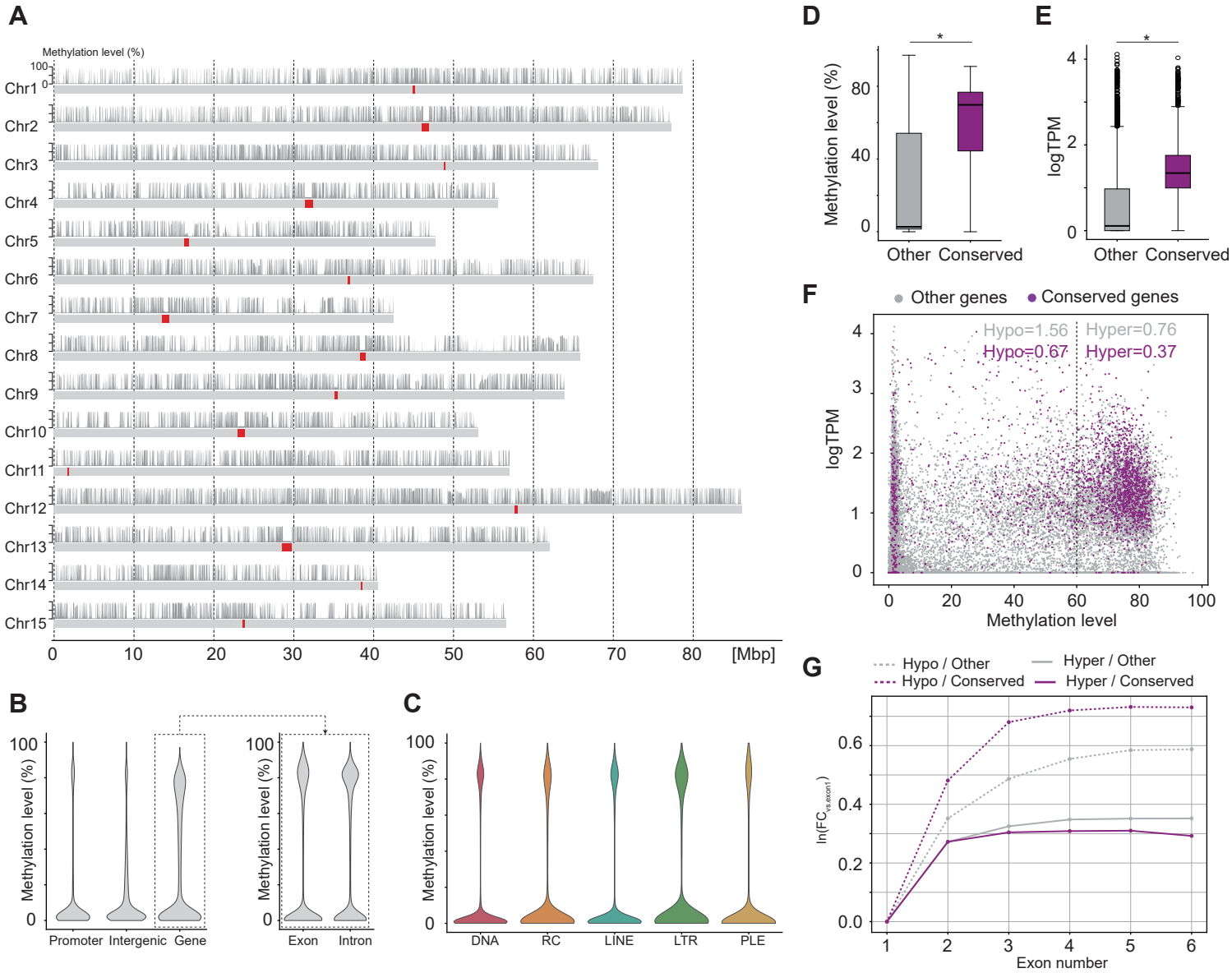
1164 Zhou C, McCarthy SA, Durbin R. 2022. YaHS: yet another Hi-C scaffolding tool.
1165 Bioinformatics. 39. doi: 10.1101/2022.06.09.495093.



Nojiri, Kin *et al*. Figure 1



Nojiri, Kin *et al*. Figure 2

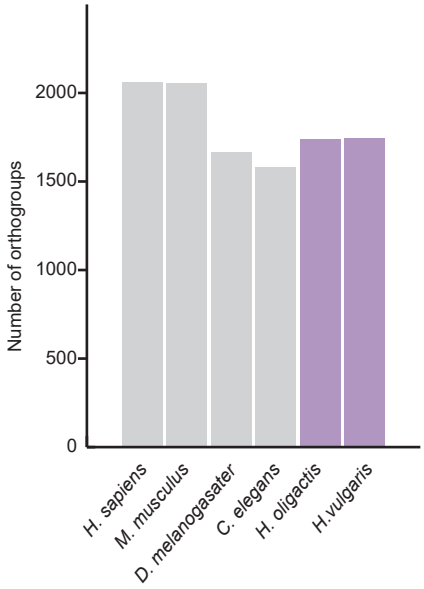


Nojiri, Kin *et al*. Figure 3

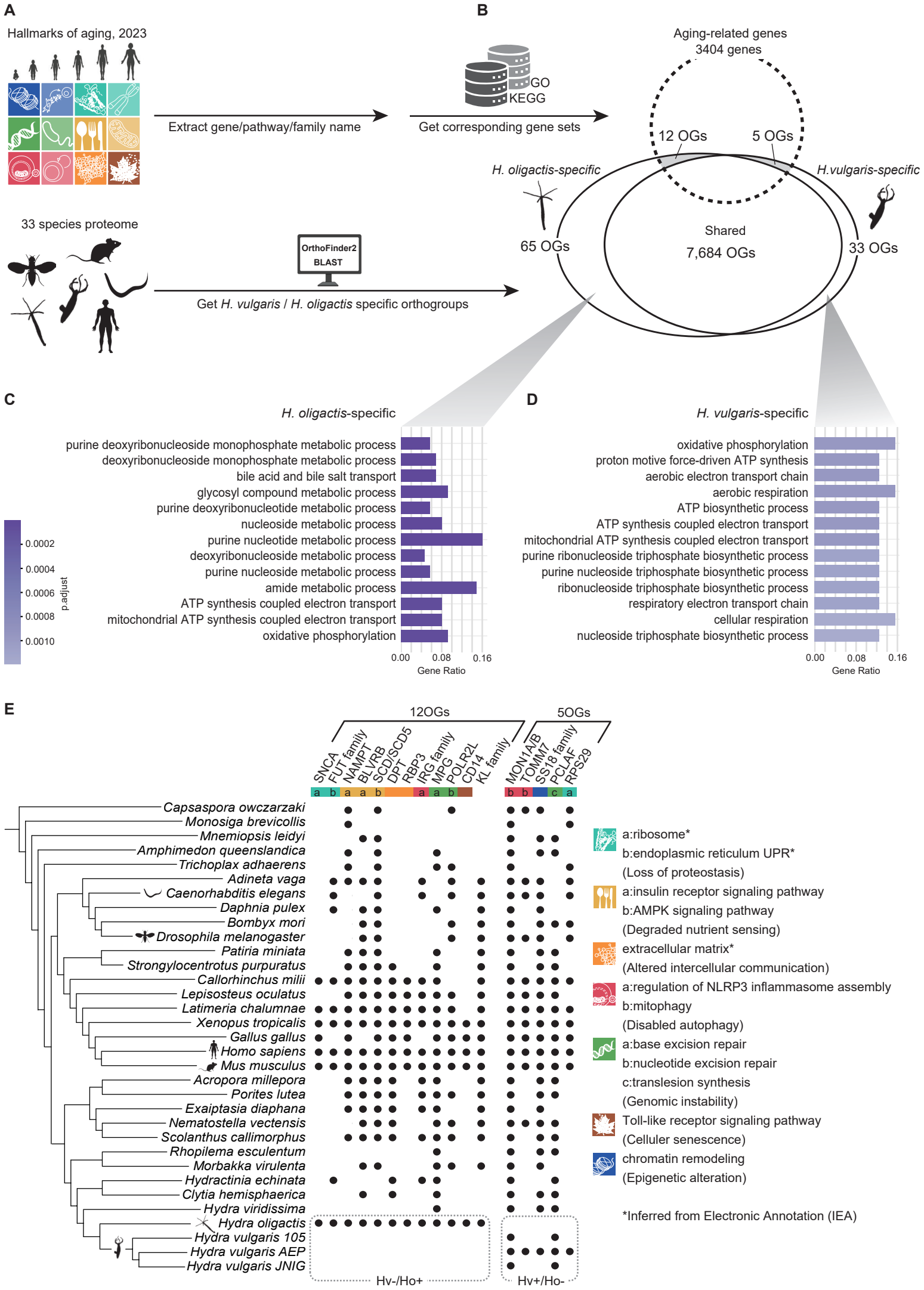
A

	H. sapiens	M. musculus	D. melanogasater	C. elegans	H. oligactis	H.vulgaris
hsa03050: proteasome	38	38	36	35	36	36
hsa03083: polycomb repressive complex	40	40	39	34	37	37
hsa03410: base excision repair	42	42	32	27	39	38
hsa03420: nucleotide excision repair	56	56	53	52	55	54
hsa03430: Mismatch repair	22	22	20	19	21	21
hsa03440: Homologous recombination	37	37	26	25	33	33
hsa03450: Non-homologous end-joining	11	11	8	8	11	11
hsa04010: MAPK signaling pathway	134	134	113	105	113	112
hsa04014: Ras signaling pathway	104	104	91	86	92	92
hsa04120: Ubiquitin mediated proteolysis	105	104	100	96	103	103
hsa04137: mitophagy	67	67	62	61	62	64
hsa04140: autophagy	115	115	108	101	107	107
hsa04142: lysosome	90	90	79	77	84	84
hsa04151: PI3K-Akt signaling pathway	173	173	125	119	128	128
hsa04152: AMPK signaling pathway	72	72	68	64	69	68
hsa04211: longevity regulating pathway	51	51	51	46	50	49
hsa04218: Cellular senescence	82	82	74	68	75	75
hsa04310: Wnt signaling pathway	86	86	78	70	80	80
hsa04330: Notch signaling pathway	31	31	28	28	27	27
hsa04340: Hedgehog signaling pathway	35	35	31	30	34	34
hsa04350: TGFβ signaling pathway	58	58	46	48	52	52
hsa04550: Signaling pathways regulating pluripotency of stem cells	60	60	59	58	58	58
hsa04620: Toll-like receptor signaling pathway	53	53	32	31	37	36

B



Nojiri, Kin *et al*. Figure 4



Nojiri, Kin *et al*. Figure 5

Table 1. Summary statistics of the genome sequencing

	Total read	Total base	Read length (N50)
WGS (Illumina)	548,373,038	82,804,238,738	2x150
WGS (Nanopore)	63,239,530	168,380,136,564	3493.0
Hi-C (Illumina)	796,383,332	119,457,499,800	2x150

Table 2. Assembly statistics of the *H. vulgaris* genomes

QUAST v5.3.0	Initial assembly AEP.JNIG	Final Assembly AEP.JNIG	Published Assembly AEP (Cazet *et al*. 2023)
Total length	954,340,183	922,069,947	900,935,055
Scaffolds	248	15	15
Longest scaffold	61,504,050	86,117,336	85,887,368
N50	9,577,232	63,895,935	59,555,380
L50	30	7	7
GC content	27.94	27.88	27.77

Table 3. BUSCO scores of the *H. vulgaris* genome assemblies

BUSCO v5.8.3	Initial assembly AEP.JNIG	Final assmbly AEP.JNIG	Published Assembly AEP (Cazet *et al*. 2023)
Complete BUSCOs	889 (93.2%)	904 (94.7%)	906 (95%)
Complete and single-copy BUSCOs	820 (86.0%)	879 (92.1%)	893 (93.6%)
Complete and duplicated BUSCOs	69 (7.2%)	25 (2.6%)	13 (1.4%)
Fragmented BUSCOs	26 (2.7%)	11 (1.2%)	11 (1.2%)
Missing BUSCOs	39 (4.1%)	39 (4.1%)	37 (3.9%)
Total BUSCO group searched	954	954	954

Table 4. BUSCO scores of the gene models based on the chromosome-level genome assemblies

BUSCO v5.8.3	Final assmbly AEP.JNIG	Published Assembly AEP (Cazet *et al*. 2023)
Complete BUSCOs	891 (93.4%)	905 (94.9%)
Complete and single-copy BUSCOs	772 (80.9%)	705 (73.9%)
Complete and duplicated BUSCOs	119 (12.5%)	200 (21.0%)
Fragmented BUSCOs	14 (1.5%)	14 (1.5%)
Missing BUSCOs	49 (5.1%)	35 (3.7%)
Total BUSCO group searched	954	954

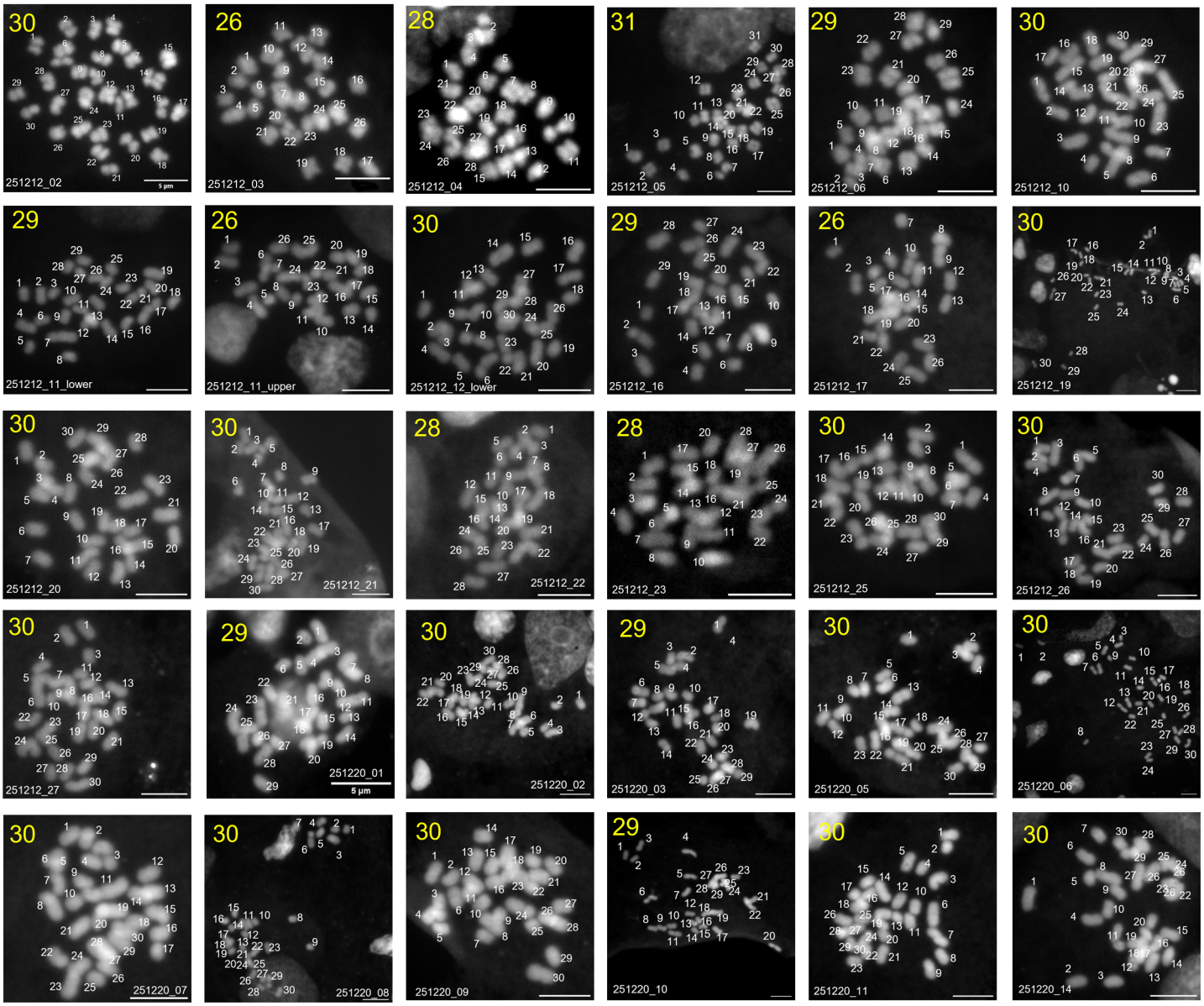
Table 5. Composition of TEs identified by Earl Grey in the genome of *H. vulgaris* AEP.JNIG

TE Classification	Coverage (bp)	Copy number	% Genome coverage	Genome size	TE family count
DNA	330,098,105	597,308	35.8	922,069,947	597,307
Rolling Circle	8,793,916	19,836	0.954	922,069,947	19,836
Penelope	23,760,322	68,248	2.58	922,069,947	68,248
LINE	140,003,727	267,317	15.2	922,069,947	267,317
LTR	9,088,772	10,622	0.986	922,069,947	10,622
Other (Simple Repeat, Microsatellite, RNA)	54,617,458	421,636	5.92	922,069,947	421,636
Unclassified	156,717,139	504,751	17.0	922,069,947	504,750
Repeat	723,079,439	NA	78.4	922,069,947	NA
Non-Repeat	198,990,508	NA	21.6	922,069,947	NA

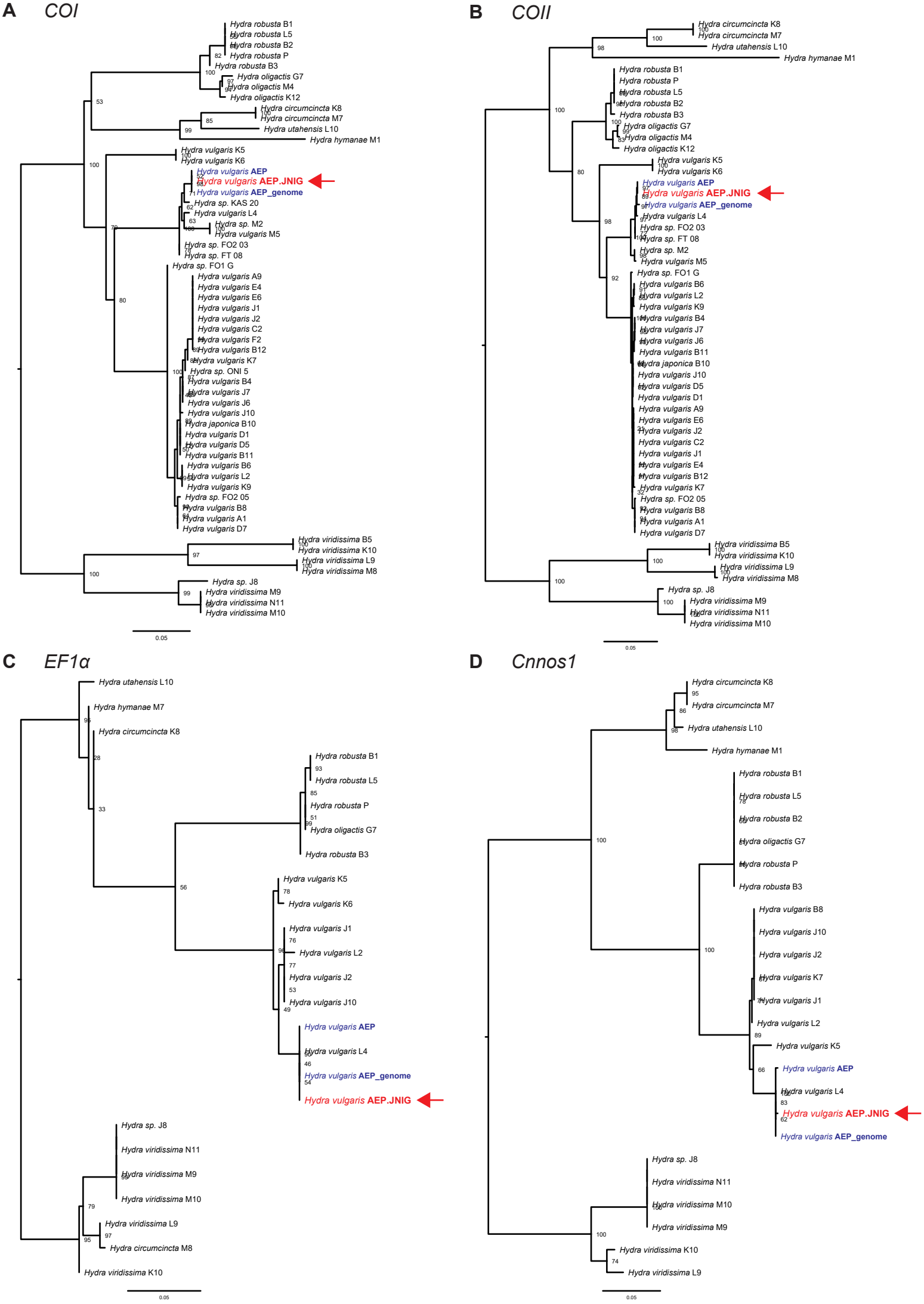
Table 6. Composition of TEs identified by Earl Grey in the genome of *H. vulgaris* AEP

TE Classification	Coverage (bp)	Copy number	% Genome coverage	Genome size	TE family count
DNA	325,041,912	583,293	36.1	900,935,055	583,293
Rolling Circle	8,786,864	18,840	0.975	900,935,055	18,840
Penelope	20,973,039	55,171	2.33	900,935,055	55,171
LINE	138,121,823	244,363	15.3	900,935,055	244,363
LTR	9,180,297	12,122	1.02	900,935,055	12,122
Other (Simple Repeat, Microsatellite, RNA)	51,236,873	418,879	5.69	900,935,055	418,877
Unclassified	146,461,074	459,730	16.3	900,935,055	459,730
Repeat	699,801,882	NA	77.7	900,935,055	NA
Non-Repeat	201,133,173	NA	22.3	900,935,055	NA

Nojiri, Kin *et al*. Tables



Nojiri, Kin *et al.*, Figure S1



Nojiri, Kin *et al*. Figure S2

A

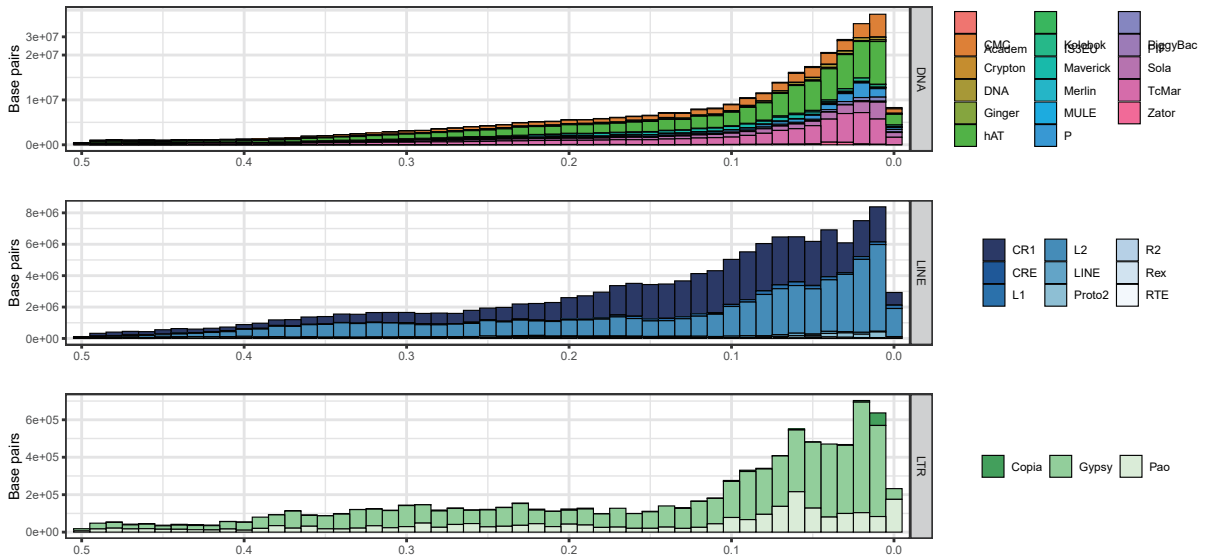
Structual annotations

Variation type	Counts	Length_ref	Length_qry
Syntenic regions	954	795720877	788696964
Inversions	360	31484124	33524687
Translocations	660	31504935	32468566
Duplications (reference)	239	5965793	-
Duplications (query)	2073	-	14204764
Not aligned (reference)	1591	60624245	-
Not aligned (query)	3516	-	57781185

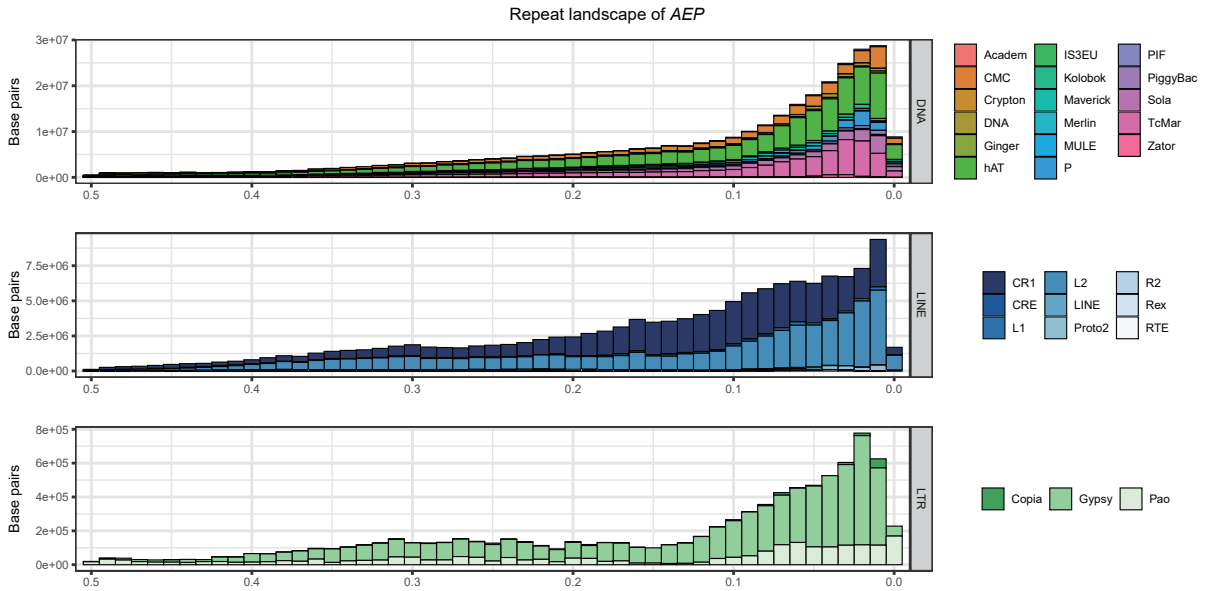
Sequence annotations

Variation type	Counts	Length_ref	Length_qry
SNPs	1833472	1833472	1833472
Insertions	648211	-	5429694
Deletions	261207	4662625	-
Copy gains	119	-	2375861
Copy losses	129	1112463	-
Highly diverged	27555	125424357	119430308
Tandem repeats	369	87143	77078

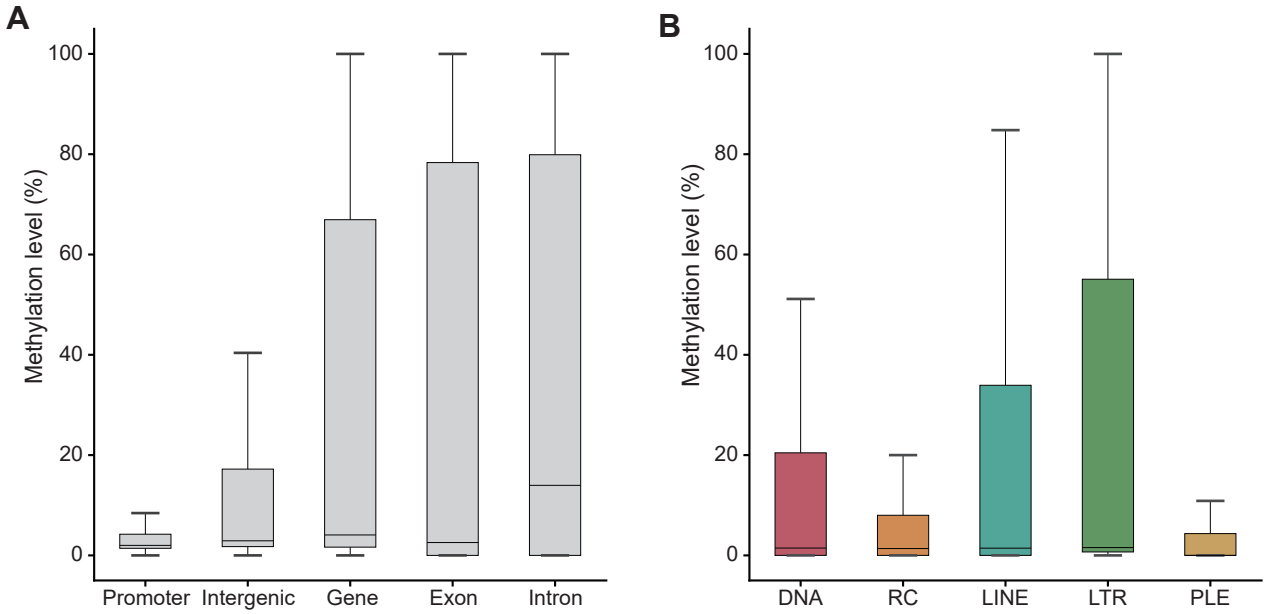
B



C



Nojiri, Kin *et al*. Figure S3



Nojiri, Kin *et al*. Figure S4

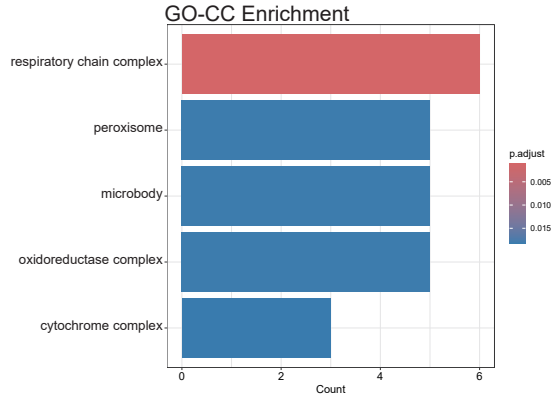
	H. sapiens	M. musculus	D. melanogasater	C. elegans	H. oligactis	H.vulgaris
GO:0004402 histone acetyltransferase activity	29	29	27	26	28	28
GO:0004407 histone deacetylase activity	11	11	11	10	11	11
GO:0005125 cytokine activity	134	133	24	30	36	35
GO:0005840 ribosome	206	206	200	191	189	189
GO:0006338 chromatin remodeling	434	434	378	335	392	393
GO:0008286 insulin receptor signaling pathway	61	61	58	55	58	56
GO:0008459 chondroitin 6-sulfotransferase activity	1	1	1	0	1	1
GO:0009008 DNA-methyltransferase activity	4	4	4	3	4	4
GO:0019985 translesion synthesis	23	23	19	18	21	22
GO:0030968 endoplasmic reticulum unfolded protein response	59	59	52	49	55	54
GO:0031012 extracellular matrix	237	237	174	169	189	187
GO:0031445 regulation of heterochromatin formation	9	9	8	8	9	9
GO:0031577 spindle checkpoint signaling	29	29	22	23	28	28
GO:0031931 TORC1 complex	6	6	6	5	6	6
GO:0032200 telomere organization	101	101	82	72	92	92
GO:0032452 histone demethylase activity	13	13	13	11	13	13
GO:0038084 vascular endothelial growth factor signaling pathway	13	13	13	10	13	13
GO:0040029 epigenetic regulation of gene expression	200	200	169	147	175	175
GO:0042054 histone methyltransferase activity	35	35	33	32	34	34
GO:0045814 negative regulation of gene expression	128	128	107	94	113	113
GO:0046902 regulation of mitochondrial membrane permeability	43	43	34	31	36	36
GO:0097472 cyclin dependent protein kinase activity	12	12	12	12	12	12
GO:0140807 NAD protein glutamate ADP ribosyltransferase activity	6	6	6	5	6	6
GO:0140896 cGAS STING signaling pathway	28	28	26	26	26	26
GO:0141137 positive regulation of gene expression	48	48	44	38	44	44
GO:1900225 regulation of NLRP3 inflammasome complex assembly	45	45	37	38	39	38

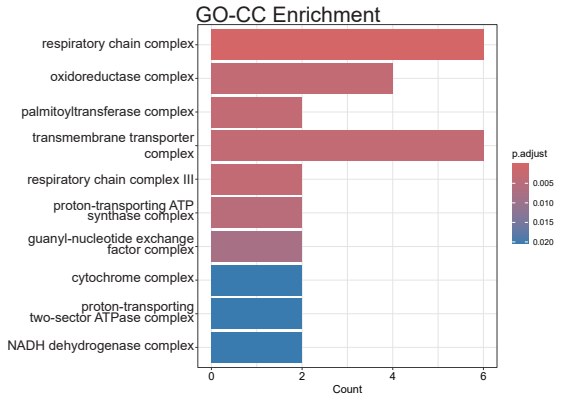
Nojiri, Kin *et al*. Figure S5

H. oligactis specifc orthogroups

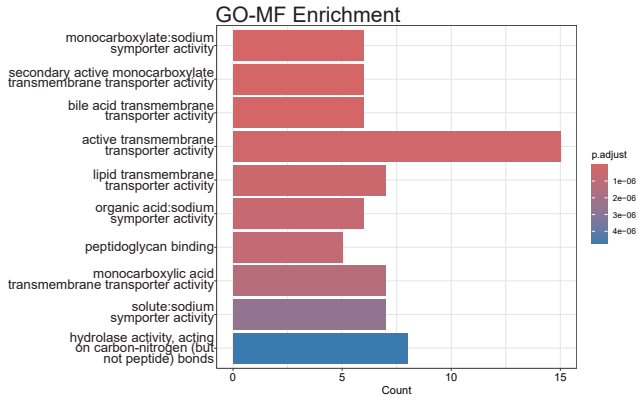
H. vulgaris specifc orthogroups

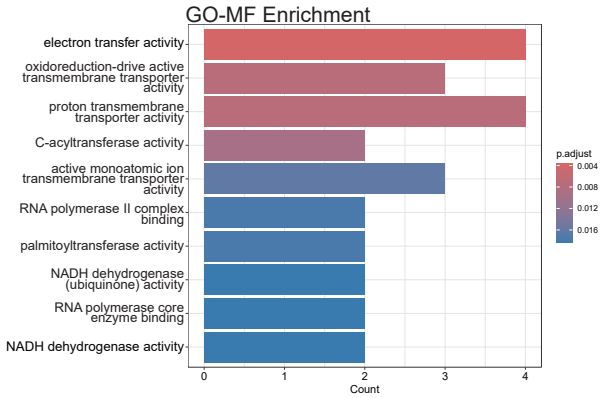
A



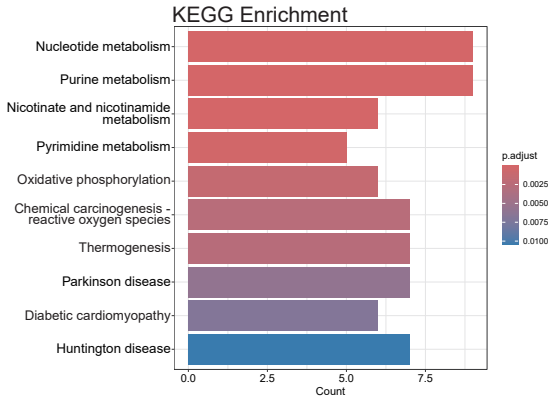


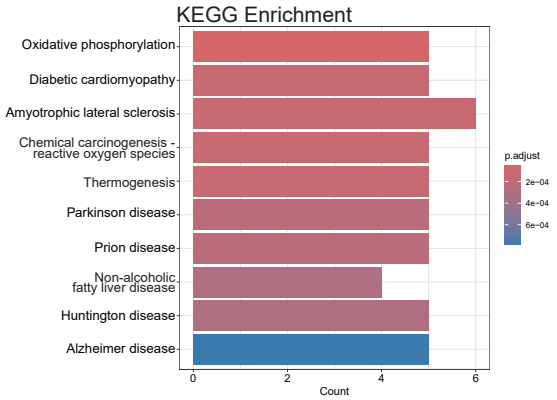
B



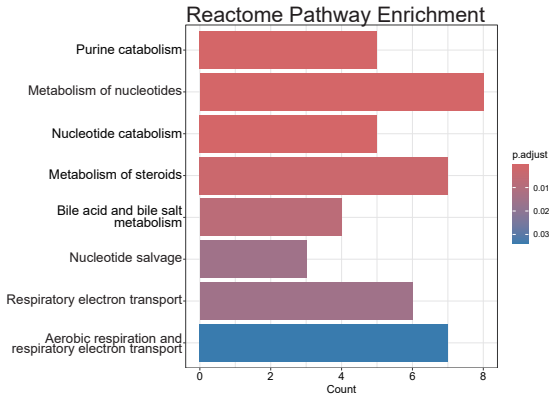


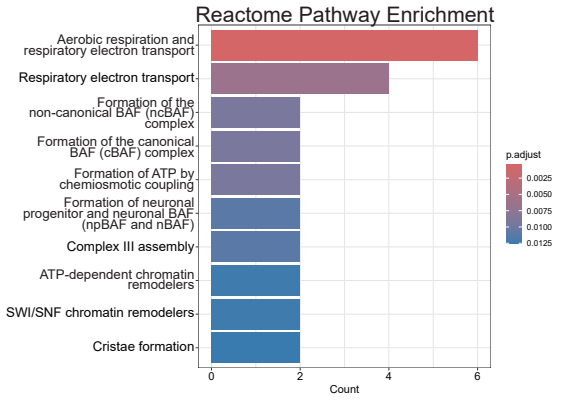
C





D





Nojiri, Kin *et al*. Figure S6