Genomic signatures of exceptional longevity and negligible aging in the long-lived red sea urchin

Jennifer M. Polinski[1], Kate R. Castellano[1], Katherine M. Buckley[2] and Andrea G. Bodnar[1]*

[1]Gloucester Marine Genomics Institute, Gloucester, MA 01930

[2]Department of Biological Sciences, Auburn University, Auburn, AL, 36849

*corresponding author

Keywords: sea urchin, longevity, aging, negligible senescence, comparative genomics, innate immunity, nervous system, genome stability

Summary

The red sea urchin (*Mesocentrotus franciscanus*) is one of the earth's longest living animals, reported to live more than 100 years with indeterminate growth, life-long reproduction, and no increase in mortality rate with age. To understand the genetic underpinnings of longevity and negligible aging, we constructed a chromosome-level assembly of the red sea urchin genome and compared it to that of short-lived sea urchin species. Genome wide syntenic alignments identified chromosome rearrangements that distinguish short- and long-lived species. Expanded gene families in long-lived species play a role in innate immunity, sensory nervous system, and genome stability. An integrated network of genes under positive selection in the red sea urchin were involved in genomic regulation, protein homeostasis, and mitochondrial function. Our results implicated known longevity genes in sea urchin longevity, but also revealed novel molecular signatures that promote long-term maintenance of tissue homeostasis, disease resistance, and negligible aging.

Introduction

There is a tremendous variety of life history strategies across the animal kingdom, with some animals achieving lifespans on the order of centuries without the physiological decline that typically accompanies aging. This phenomenon, referred to as negligible senescence, is characterized by a lack of increased mortality rate or decreased fecundity as an organism ages, in combination with maintenance of physiological function and disease resistance.[1] Animals with extraordinary longevity and negligible senescence have evolved mechanisms to promote long-term maintenance of tissue homeostasis and physiological function while avoiding degenerative and neoplastic disease. Understanding these mechanisms can reveal effective strategies to achieve longevity and healthy aging.

Comparative genomics between long- and short-lived species is a powerful approach to understand the evolution of longevity and enables unbiased discovery of genes and pathways that regulate lifespan. This approach has been used to identify molecular signatures related to longevity in mammals and other vertebrates, uncovering both shared and distinct strategies to modulate aging and disease resistance.[2-9] Investigating a wider range of long-lived animals, including those that exhibit negligible senescence, promises to reveal novel strategies to promote longevity and healthy aging and uncover pathways that could potentially be exploited for preventative or therapeutic avenues for age-related degenerative disease.

Sea urchins represent a promising group of organisms to advance our understanding of lifespan determination and healthy aging. There are about 1,000 extant sea urchin species with a wide range of lifespans, including species with exceptional longevity and negligible senescence. As echinoderms, sea urchins are non-chordate deuterostomes that share a close genetic relationship with vertebrates and have a long history as research models that have advanced our

understanding of cell cycle regulation, the role of chromosomes in inheritance, and gene regulatory networks of development.[10-12] They are also commercially fished which provides a wealth of data on their life histories and incidence of disease.[13] Life history data indicates that the red sea urchin, *Mesocentrotus franciscanus*, is one of the earth's longest living animals. It is reported to live more than 100 years and shows negligible senescence as defined by indeterminate growth, life-long reproduction, no age-associated increase in mortality rate or increased incidence of disease, including no known cases of cancer.[14-19] Negligible senescence has also been reported for other sea urchin species despite a wide range of lifespans.[17] This includes the variegated sea urchin, *Lytechinus variegatus*, which is reported to live for 3-4 years[20], the painted sea urchin, *Lytechinus pictus*, reported to live 5-7 years[21], and the purple sea urchin, *Strongylocentrotus purpuratus,* reported to live more than 50 years.[22,23]

Studies to date, conducted within the framework of known theories of aging, have demonstrated that both short- and long-lived sea urchin species lack many hallmarks of aging. Sea urchins maintain telomere length, antioxidant and proteasome enzyme activities, and regenerative capacity, and exhibit little accumulation of oxidative cellular damage with age.[24-26] Gene expression studies using tissues of long-lived species indicate that key cellular pathways involved in protein homeostasis, tissue regeneration, and neurological function are maintained with age.[27,28] Although genomes have been assembled for several sea urchin species, including *S. purpuratus*, *L. variegatus*, and *L. pictus*[11,29,30], to date, no assembled genome was available for the long-lived red sea urchin *M. franciscanus*. Here we report a chromosome-level assembly for the red sea urchin genome. Targeted analysis of this genome and comparisons between long- and short-lived species sheds light on the molecular, cellular, and systemic mechanisms that promote longevity and negligible senescence.

Results

Sequencing and assembly of the red sea urchin genome yielded 21 chromosome-length scaffolds

DNA isolated from testis of one male *M. franciscanus* was used for genome sequencing and assembly (Figure 1A). A total of 16 million Oxford Nanopore Technologies (ONT) long-reads, ranging in size from 1-195 kb and totaling 50.15 Gb, were used as input for the preliminary CANU assembly[31], and 543.3 million Illumina read pairs (157.6 Gb) were used for POLCA polishing.[32] Scaffolding with Hi-C data resulted in a 96.8% complete final assembly spanning 773.16 Mb across 21 chromosome-length scaffolds, containing 22,240 predicted gene models (Figure 1, Table S1, Table S2). The number of scaffolds was identical to the number of chromosomes reported for closely related species (*S. purpuratus* and *S. droebachiensis*) based on karyotype analysis (n=21).[33] Repeat annotations using a *de novo* RepeatModeler library with RepeatMasker[34,35] revealed that the *M. franciscanus* genome was comprised of 59.33% repetitive sequences, with 28.71% being unclassified. The overall percentage and composition of repeat elements in the *M. franciscanus* genome was similar to that of other sea urchin species including long-lived *S. purpuratus* and short-lived *L. variegatus* and *L. pictus*, with the largest group of classified repeats being long interspersed nuclear elements (LINE) (Figure S1, Table S3).

Syntenic alignments revealed chromosomal rearrangements between long- and short-lived sea urchin species

Although the genome sizes are similar between our selected short- and long-lived sea urchin species, synteny analysis revealed differences at the structural level. The predicted number of chromosomes in the long-lived red sea urchin *M. franciscanus* (n=21) was identical to

that of long-lived *S. purpuratus*, but differed from that of short-lived sea urchins *L. variegatus* and *L. pictus* (n=19).[29,30] Syntenic alignments revealed conserved interchromosomal organization between the long-lived species *M. franciscanus* and *S. purpuratus* (Figure 2A). However, comparing short- and long-lived species revealed that chromosome 2 in *L. pictus* and *L. variegatus* align with chromosomes 7 and 11 in *M. franciscanus* (Figure 2A, Figure S2). Further, chromosome 5 in *L. pictus* aligns with chromosomes 10 and 15 in *M. fransciscanus*, while chromosome 1 in *L. variegatus* aligns with chromosome 3 and 15 in *M. franciscanus* (Figure 2A, Figure S2). The two short-lived species share a high level of interchromosomal synteny except that part of chromosome 1 of *L. variegatus* aligns to chromosome 5 in *L. pictus*[30] (Figure 2A). Ancestral sea urchins (e.g., genera *Arbacia* and *Stomopneustes*) have n=22 chromosomes[36] suggesting that long-lived *M. franciscanus* and *S. purpuratus* have undergone one chromosomal fusion event, while the short-lived species undergone three fusion events during evolution.

Microsyntenic analysis revealed a high degree of intrachromosomal alignment between *M. franciscanus* and *S. purpuratus*, along with some large inversions, and revealed a high level of intrachromosomal rearrangements between *M. franciscanus* and short-lived species *L. pictus* and *L. variegatus* (Figure 2B-D). Consistent with previous reports[30], our comparisons of the *L. variegatus* and *L. pictus* genome sequences indicated that these two closely-related species are largely syntenic with some microsyntenic inversions and translocations (Figure 2E). Despite the large number of intrachromosomal rearrangements between long- and short-lived species, analysis of the highly conserved Hox gene cluster, indicated that the order and directionality of the Hox genes are conserved (Figure S2).

Genes under positive selection in the long-lived red sea urchin are associated with genomic regulation, stress response, nutrient sensing, peroxisome and mitochondria metabolism

Analysis of evolutionary selection, conducted on the coding regions of genes across four sea urchin species with different lifespans (*M. franciscanus, S. purpuratus, L. variegatus*, and *L. pictus*), revealed 446 orthogroups subject to positive selection in the long-lived red sea urchin *M. franciscanus* (p<0.001) (Table S4). Orthologs under positive selection were assigned to functional categories using a custom sea urchin gene ontology[28,37] with the largest proportion associated with Nucleic Acid Metabolism (100 genes), followed by Protein Homeostasis (70 genes), Metabolism (62 genes), Signaling (50 genes) and Mitochondria (46 genes) (Figure 3A, Table S4). Functional enrichment analysis on this gene set revealed significant enrichment of Gene Ontology (GO) terms related to nucleic acid metabolic processes, mitochondrial function and oxidative phosphorylation, protein transport and homeostasis, and cellular metabolism (Benjamini-Hochberg FDR corrected p-value <0.05) (Table S5). The top enriched KEGG (Kyoto Encyclopedia of Genes and Genomes) pathway was oxidative phosphorylation (Table S5).

Using the STRING database[38], we generated a protein-protein interaction (PPI) network from the genes subject to positive selection. This network had more interactions than expected for a random set of proteins of the same size and degree distribution drawn from the genome (p-value = 6.51×10^{-5}), indicating that the proteins encoded by the positively selected genes are at least partially biologically connected (Figure 3B, Figure S3). Gene expression patterns indicated that many of the positively selected genes were highly expressed during early development and showed distinct patterns of age-related expression in adult tissues (muscle, esophagus and radial nerve cord) (Figure S4, Table S4). The pattern of age-related gene expression was particularly striking in radial nerve cord which showed an upregulation with age of 187 of the 446 positively

selected genes (log2-fold change of >1.5) (Figure S4, Table S4) and an upregulation of 75 out of the 180 genes within the STRING network (log2-fold change of >1.5) (Figure 3C).

Fifteen of the positively selected genes have been previously associated with longevity and aging based on genetic perturbation experiments in short-lived model organisms represented in the GenAge database[39], and/or by comparative genomic analysis of long-lived animals[6,7] (Figure 3D, Table S6). These shared genes play a role in the mechanistic target of rapamycin (mTOR) pathway, insulin signaling, peroxisome biogenesis, mitochondrial function, and endoplasmic reticulum (ER) stress, among others (Table S6).

Within the Nucleic Acid Metabolism category, positively selected genes included 37 genes encoding transcription factors and transcriptional regulators, 24 genes encoding splicing factors and mRNA processing proteins**,** 13 genes encoding proteins involved in chromosome structure and chromatin remodeling, and 8 genes encoding proteins with a role in DNA damage response and DNA repair (Table S4). Of particular interest were genes encoding Activating Transcription Factor 4 (ATF4) and Forkhead Box Protein O1 (FOXO1). ATF4 is a transcription factor that acts as a regulator of metabolic and redox processes under normal cellular conditions and is a master regulator of the cellular stress response required for adaptation to ER stress, amino acid starvation, mitochondrial stress or oxidative stress.[40] FOXO transcription factors are essential regulators of cellular homeostasis and important determinants of aging and longevity downstream of insulin and insulin-like growth factor signaling.[41,42]

Signatures of positive selection were detected in several red sea urchin genes associated with lipid metabolism, including six that play a role in peroxisome biogenesis (Table S4). A total of 46 nuclear-encoded mitochondrial genes under positive selection carry out a wide range of mitochondrial functions, including respiration (13 genes), metabolism and fatty acid oxidation

(13 genes), mtDNA and protein synthesis (10 genes), mitochondrial organization and fission (6 genes), and the TCA cycle (2 genes). Positive selection was also detected in genes associated with protein homeostasis, including the ubiquitin proteasome pathway (18 genes) and the ER-associated degradation of misfolded proteins (ERAD) pathway (7 genes) (Table S4). This includes the gene encoding Hypoxia up-regulated protein 1 (HYOU1) which was also associated with longevity in rockfish.[6] There were 20 genes under positive selection encoding proteins that play a role in vesicle-mediated protein trafficking, including vacuolar protein sorting-associated protein 41 (VPS41) which is a pro-longevity gene in model organisms in the GenAge database[39] (Table S6).

Within the red sea urchin, positive selection was detected among several genes associated with cell signaling; the Ras-MAPK signal transduction pathway was the most highly represented (15 total genes; Table S4). Two genes under positive selection encoded proteins of the mTOR pathway: one subunit of mTOR1 (GATOR complex protein WDR59) and one subunit of mTOR2 (RICTOR). mTOR is a master regulator of cellular metabolism and well-known for its role in aging and lifespan determination across multiple taxa.[43] The insulin signaling pathway is also known to play a conserved role in aging across taxa[42,44] and, in addition to *FOXO*, three genes that play a role in insulin signaling were under positive selection in the red sea urchin genome: pappalysin1 (*PAPP1*), Suppressor of cytokine signaling 7 (*SOCS7*), and sirtuin-4 (*SIRT4*)[45-47] (Table S3).

Expanded gene families in long-lived sea urchins have a role in innate immunity, nervous system function, and nucleic acid metabolism

Ortholog analysis of eight echinoderm species including *M. franciscanus*, *S. purpuratus*, *L. variegatus*, and *L. pictus*, identified 36,621 gene family clusters. Using 1,245 single-copy orthologs, we generated a species-level phylogeny and identified 901 expanded and 1023 contracted gene families in the lineage giving rise to the long-lived sea urchins (*M. franciscanus* and *S. purpuratus*) (Figure 4A). Of the 901 expanded gene families, 73 were considered significant in CAFE analysis ($p<0.05$) and assigned to functional categories using a custom sea urchin gene ontology[28,37] with the top three functional categories being Immunity (18 orthogroups), Nervous System (18 orthogroups), and Nucleic Acid Metabolism (8 orthogroups) (Figure 4B, Table S7). Only four of the 1023 contracted gene families at the branch leading to long-lived sea urchin species were considered significant ($p<0.05$) and all four were annotated as uncharacterized.

In contrast, within the lineage leading to the short-lived sea urchins (*L. variegatus* and *L. pictus*), 41 gene families were significantly expanded, nine of which were assigned functional annotations (Table S7). One orthogroup of interest was iodothyronine deiodinase type 1 (*Dio1*) which regulates thyroid hormones that are essential for growth, development and basal metabolism and plays a role in sea urchin metamorphosis.[48,49] This expansion may be related to the accelerated maturation time and rapid growth of the short-lived sea urchin species.

Within the long-lived sea urchin species, expanded gene families associated with Immunity are predicted to play two distinct but complementary roles: 1) detecting and responding to pathogenic microbes; and 2) maintaining tissue homeostasis. Specifically, genes within expanded families encode pattern recognition receptors, including Toll-like receptors (TLRs) and Nod-like receptors, as well as proteins that contain Scavenger Receptor Cysteine Rich (SRCR) and C-type lectin domains. Notably, several expanded gene families have been

implicated in antiviral immune responses in other systems including five families of TRIM E3-ubiquitin ligases and Sterile alpha motif and HD-domain containing protein 1 (SAMHD1) (Table S7).

In the Nervous System category, several G protein-coupled receptor (GPCR) gene families were expanded in the long-lived sea urchin lineage including adhesion type (Class B2) and rhodopsin type (Class A) (Table S7). These expanded repertoires may confer an enhanced ability to detect and respond to a diverse array of environmental stimuli. In addition to GPCRs, an expansion was evident in the gene family encoding low-density lipoprotein receptor-related protein 4 (LRP4), which plays an important role in formation and maintenance of neuromuscular junctions, synaptogenesis and synaptic organization, regeneration of peripheral nerves, and clearance of amyloid proteins.[50-52]

In the Nucleic Acid Metabolism functional category, gene families encoding histone-lysine N-methyltransferases were expanded in the branch leading to long-lived sea urchins including N-lysine methyltransferase KMT5A and KMT5A-A, as well as Histone-lysine N-methyltransferases SMYD2 (including SMYD2-A, SMYD2-B) and SMYD3 (Table S7). Histone-lysine N-methyltransferases play an important role in regulating chromatin states and regulating the accessibility of DNA to trans-acting factors that mediate transcription, DNA repair, DNA replication, and transposition.[53] Several families of transposons were expanded in the long-lived red sea urchin (Table S7) including the retrovirus-related Pol polyprotein from transposon 17.6 and retrovirus-related Pol polyprotein from transposon 412. Orthologs of these transposons are highly upregulated in response to heat stress in the coral *Acropora hyancinthus* and may contribute to stress adaptation.[54]

Given the central role of immunity, nervous system function, and genome stability in promoting survival, disease resistance, and longevity, we conducted more comprehensive targeted analyses of predicted genes within each of these categories in the long-lived red sea urchin genome.

The *M. franciscanus* immune gene repertoire revealed expansion of pattern recognition receptors and regulators of anti-viral immunity

The immune system plays a critical role in promoting survival and longevity by mediating the complex interactions between a host and the microbial world and also maintaining tissue homeostasis by detecting and removing dying or aberrant cells. Sea urchins and other echinoderms lack adaptive immunity but possess a complex and well-characterized innate immune system.[55-57] A comprehensive survey of the immune gene repertoire in the red sea urchin genome revealed an expanded suite of genes encoding receptors, regulators, and effectors of the innate immune response (Figure 5, Table S8).

One of the hallmarks of sea urchin immunity is a large array of genes that encode pattern recognition receptors (PRRs) that are homologous to those that mediate vertebrate innate immunity (e.g., TLRs, NLRs, RLRs, and SRCRs). While these sequences exhibit substantial diversity at the amino acid level, they can be reliably identified based on broadly conserved domain architectures.[58] The red sea urchin genome contains an expanded receptor repertoire consisting of 167 genes encoding TLRs, 82 NACHT and leucine-rich repeat containing proteins (NLR), 17 genes encoding Retinoic Acid-Inducible Gene 1 or RIG-I-like receptors (RLR) and 329 scavenger receptor cysteine-rich (SRCR) proteins that contain a total of 709 SRCR domains (Figure 5, Table S8). In other systems, the NLRs, RLRs, and SRCRs are often highly divergent

and consist of many short exons, posing challenges for gene prediction algorithms. However, in echinoderms, TLRs are encoded in single exons, which allows for careful sequence inspection. Notably, of the 167 TLR genes, 81 are likely pseudogenes due to truncations, frameshifts and/or internal stop codons. This is consistent with the dynamic evolutionary history of invertebrate TLRs.[57] In addition to the RLRs that recognize different classes of nucleic acids, the red sea urchin genome contains orthologs encoding proteins that mediate cytosolic DNA sensing in mammals, including DDX41, RNA pol III and MRE11 (Figure 5, Table S8). Notably, however, clear orthologs of the DNA sensor cGMP-AMP synthase (cGAS) and the stimulator of interferon genes (STING) were absent from the red sea urchin genome.

Interactions between PRRs and their cognate ligands initiate an intracellular signaling cascade that leads to the upregulation of transcripts that encode cytokines, which amplify and shape the subsequent immune response, and effector molecules that kill microbes either directly or indirectly.[59] In *M. franciscanus,* we identified orthologs for components of signaling pathways downstream of TLRs that result in the activation of NF-κB and TBK-1/IRF as well as mitochondrial antiviral signaling protein (MAVS), the immediate RLR adaptor. Orthologs of many genes whose expression is induced by interferon in mammals were identified in the red sea urchin genome, but notably, we did not identify an ortholog encoding interferon. This is consistent with the hypothesis that this cytokine is restricted to jawed vertebrates.[60] However, it is possible that distant orthologs are present within sea urchins but not detected due to high rates of divergence. As in other echinoderms, the only identifiable interleukin ortholog in *M. franciscanus* was IL-17, which is present in a family of 23 paralogs (Table S8). Additional cytokine genes detected include four orthologs encoding tumor necrosis factor (TNF), which, in mammals, serves a variety of proinflammatory functions. The red sea urchin genome also

contains orthologs for each of the transcription factor families that regulate vertebrate hematopoiesis (Table S8).

Tripartite motif (TRIM) proteins are a versatile family of E3 ubiquitin ligases with a broad array of functions. Recent studies suggest that TRIM proteins are involved in innate immunity at multiple levels by regulating the activity of PRRs, adaptor molecules and transcription factors.[61] TRIM proteins have also been implicated in antiviral activities, including interfering with several stages of viral life cycles and inducing autophagy in virally-infected cells. [62,63] Based on the unique domain architecture of these proteins (an N-terminal RING-BBox-coiled-coil [RBCC] motif in addition to one or more C-terminal domains), we identified a total of 97 TRIM orthologs within the *M. franciscanus* genome. This includes significant duplication of two TRIM proteins that exist as single-copy orthologs in mammals: TRIM33 (29 homologs) and TRIM56 (39 homologs) (Figure 5, Table S8). TRIM33 has been implicated in several unique aspects of immunity. Upon interacting with the dsRNA sensor DHX33, TRIM33 activates the NOD-, LRR- and pyrin domain-containing protein 3 (NLRP3)-dependent inflammasome, resulting in a proinflammatory response.[61] TRIM33 has also been associated with HIV-infection by targeting viral proteins for degradation, a function that may extend to other retroviruses.[64] Additionally, TRIM33 has been shown to play roles in tumor suppression[65] and loss-of-function resulted in accelerated aging of hematopoietic stem cells.[66] TRIM56 was originally reported to have a role in regulating the response to intracellular dsDNA[67] but has also been shown to have several roles in antiviral immunity, including regulating signaling pathways downstream of TLR3.[68] Given the diverse roles of these proteins within mammalian systems, it is difficult to speculate on their functions in the red sea urchin. However, the multiplicity of

these gene families may suggest that individual sea urchin TRIM proteins are undergoing sub- or neofunctionalization.

Targeted analysis of the *M. franciscanus* genome revealed a complex repertoire of genes associated with the sensory nervous system

Gene family analysis revealed an expansion of several gene families associated with neuronal function in the long-lived sea urchin lineage. Targeted examination of the neural gene complement in the red sea urchin genome identified an extensive repertoire of more than 1800 genes **(**Table S9). The most remarkable features include an extensive repertoire of GPCRs and transient receptor protein (TRP) channels, an expanded number of genes involved in the cholinergic signaling system in addition to the expansion of genes encoding LRP4 (Figure 6A and 6B, Table S9, Table S10).

In total, 1132 GPCRs were identified (based on the presence of Pfam domains for 7 transmembrane receptors: PF00001, PF00002, PF00003), representing 5.1% of all predicted genes in the red sea urchin genome. The GPCR family includes 959 Class A (rhodopsin-type receptors), 143 Class B (secretin and adhesion receptors), 27 Class C, and 3 Class F (Frizzled) receptors (Figure 6C, Table S9). Class A, or rhodopsin-type receptors, are the largest group of GPCRs in many organisms. They are important for chemical sensing and photoreception and react to a range of stimuli, including small molecules, photons, peptides and proteins.[69] The large number of rhodopsin-type GPCRs suggests that these genes play critical roles in sea urchins. This may include chemosensory and photosensory functions for feeding, predator avoidance, habitat selection, mating and coordinating reproductive behaviors such as mass spawning events, larval settlement and metamorphosis.[70] One expanded rhodopsin-type GPCRs was GRL101,

which belongs to a family of leucine-rich repeat containing GPCRs that contain multiple low-density lipoprotein class A (LDLa) motifs (Figure 6D) that are thought to directly transduce signals carried by large extracellular lipoprotein complexes into neuronal events.[71] High levels of expression of GRL101 in the sensory organs of other echinoderms suggests a role in chemosensation[72], but the function of this gene has not been characterized.

In addition to the large repertoire of GPCRs, the red sea urchin genome had 68 homologs encoding TRP channels, categorized into seven subfamilies, more than twice the number found in other animals including humans (Figure 6B, Table S9). TRP channels can detect many types of environmental stimuli, including mechanosensation and thermosensation, mediating response to touch, pressure, vibrations, and temperature.[70] The red sea urchin genome contains 46 genes annotated as *TRPA-1*, a family that is important for thermoreception in many animals including sea urchins.[73] The extensive repertoire of GPCRs and TRP channels indicates well-developed sensory perception that may promote survival in the benthic marine environment.

The red sea urchin genome has a rich repertoire of 68 genes encoding nicotinic acetylcholine receptors (nAChR) along with 39 genes encoding acetylcholinesterase and 14 encoding cholinesterases that break down acetylcholine, in addition to 89 muscarinic acetylcholine receptors (mAChR - Class A GPCRs) (Figure 6B, Table S9). The cholinergic system is an ancient signaling pathway that regulates diverse physiological processes and contributes to development, immune, and environmental responses. In sea urchins, cholinergic signaling mediates fertilization and early embryonic development.[74,75] The transcription patterns for cholinergic genes in the red sea urchin show a wide variation among the early life stages and adult tissues indicating a large diversity of functions (Table S9). Expanded arrays of nAChRs are

found in some molluscan genomes (Figure 6B), which has been predicted to confer enhanced capacity to quickly respond to variable environmental stimuli.[76]

The *M. franciscanus* genome contains expansions of key tumor suppressor and DNA repair genes

Maintaining genome integrity is critical to promote longevity and to impede the formation of neoplastic disease. We therefore surveyed the *M. franciscanus* genome for genes known to play a role in genome stability. A targeted list of 469 genes associated with DNA repair, cell cycle regulation, tumor suppression, apoptosis, chromatin structure and silencing was surveyed, with 450 (96%) identified as having at least one homolog in the red sea urchin genome. Notably, multiple copies of some tumor suppressor and DNA repair genes were identified (Figure 7, Table S11).

Efficient DNA repair is essential to maintain genome integrity and cellular function, and homologs for 82 of the 83 repair genes surveyed were identified in the red sea urchin genome (Table S11). Of particular note were three homologs encoding Excision Repair Cross-Complementing group 6 protein (ERCC6) (Figure 7A). ERCC6 is a DNA-binding protein involved in transcription-coupled nucleotide excision repair that allows RNA polymerase II-blocking lesions to be rapidly removed from the transcribed strand of active genes.[77] Expression patterns indicate that the *ERCC6* homologs were highly expressed in sea urchin embryonic and larval stages (Figure 7B, Table S11) where they may play a role in repair of cyclobutane pyrimidine dimers (CPD) and 6-4 pyrimidine-pyrimidone photoproducts (6–4PP) induced by UV radiation in the surface seawater. Notably, all three *ERCC6* homologs were upregulated with age in the red sea urchin radial nerve cord (Figure 7B, Table S11). ERCC6 is known to play an

essential role in neurological function in humans and mutations in *ERCC6* are associated with Cockayne Syndrome, a progressive developmental and neurodegenerative disorder, with growth abnormalities, progeroid features, and sun sensitivity.[78] Two homologs encoding X-Ray Repair Cross-Complementing Protein 1 (XRCC1) were identified in the red sea urchin genome. XRCC1 is a scaffold protein recruited to the site of DNA damage through the activity of poly (ADP-ribose) polymerases (PARPs) and plays a key role in orchestrating base excision repair and single-strand break repair.[79,80] Impaired function in XRCC1 has been associated with increased cancer susceptibility in humans.[79]

The cell cycle is tightly regulated by cyclins, cyclin-dependent kinases (CDKs) and CDK inhibitors to prevent the propagation of mutations and tumorigenesis. The red sea urchin genome has a nearly complete complement of the bilaterian genes encoding cyclins and CDKs, with the exception of *CDK3*, *CDK15*, *CDK17* and *cyclin M* (Table S8). One homolog was identified for each of the cyclin-dependent kinase inhibitor families *INK4* (p16/p19) and *Cip1/Kip1* (p21/p27), which play important roles in cell cycle regulation and tumor suppression[81] (Table S11). Similarly, a single homolog of the tumor suppressor gene *TP53* was identified in the red sea urchin. *TP53* encodes a transcription factor that has a central role in safeguarding genome integrity through modulating expression of genes involved in cell cycle regulation, DNA repair, and apoptosis.[82,83] The red sea urchin homolog most closely resembled the ancestral family member *TP63* (Figure 7A, Table S11). In contrast, three homologs for the retinoblastoma tumor suppressor gene (*RB*) were identified in the red sea urchin genome (Figure 7A, Table S11). Phosphorylation of the RB protein prevents its binding to E2F transcription factors, enabling them to activate the expression of target genes that regulate cell cycle.[84] Homologs encoding both ataxia telangiectasia and Rad3-related protein (ATR) and ataxia-telangiectasia-mutated

protein (ATM) were present in the red sea urchin genome along with two homologs encoding serine/threonine-protein kinase CHK1 (Figure 7A) and one homolog encoding serine/threonine-protein kinase CHK2 (Table S8). The kinases ATR and ATM are recruited to the site of DNA damage and phosphorylate and activate the effector kinases CHK2 and CHK1, respectively, leading to cell cycle arrest.[83] There were two copies of each of the tumor suppressors phosphatase and tensin homolog deleted on chromosome 10 (*PTEN*) and von Hippel-Lindau disease tumor suppressor gene (*VHL*) (Figure 7A, Table S11). PTEN is a negative regulator of a major cell growth and survival signaling pathway, the phosphatidylinositol-3-kinase (PI3K)/AKT signaling pathway, and a decrease in PTEN activity results in cancer susceptibility and tumor progression in humans.[85] The VHL tumor suppressor protein plays a key role in cellular oxygen sensing by targeting hypoxia-inducible factors for ubiquitylation and proteasomal degradation.[86] Germline inactivation of the VHL gene in humans causes the von Hippel-Lindau hereditary cancer syndrome, and somatic mutations of this gene result in the development of hemangioblastomas and clear-cell renal carcinomas.[87] The tumor suppressor and DNA repair genes were highly expressed in red sea urchin developmental stages, and some were upregulated with age in adult tissues, particularly in radial nerve cord (Figure 7B). With the exception of VHL, the expansion of the tumor suppressor genes was not unique to the long-lived red sea urchin but occurred in at least one other sea urchin species regardless of lifespan (Figure 7C).

Discussion

Genetic manipulation of short-lived model organisms (e.g., *C. elegans*, *D. melanogaster*, and *M. musculus*) has identified many diverse genes that play a role in aging and lifespan

410 determination and revealed major pathways that modulate these processes across taxa.[88] Despite
411 these significant advances, alterations of single genes do not capture the complexity of the
412 multifactorial processes that promote longevity in natural populations, and lifespan extensions
413 achieved using this approach are far smaller than the variation of lifespans among species shaped
414 by natural selection.[6,7] Comparative genomic approaches using non-model organisms with
415 different lifespans are beginning to reveal the full scale of multigenic traits that promote
416 longevity in nature.[2-9] However, these studies are limited by the genomic resources available for
417 non-model organisms and are often confounded by large evolutionary distances, and phenotypic
418 differences when comparing animals with different lifestyles and habitats. In this study, we
419 generated a high-quality reference genome for the long-lived red sea urchin and conducted
420 comparisons among sea urchin species with negligible senescence but very different lifespans.
421 This study is unique in that it focuses on animals with negligible senescence that share a high
422 degree of phenotypic similarity and similar habitats. In fact, three of the four species used in our
423 comparisons (*S. purpuratus*, *M. franciscanus*, and *L. pictus*) share overlapping geographic ranges
424 in the Northeast Pacific Ocean, while *L. variegatus* is found in the western central Atlantic
425 Ocean. Furthermore, the high-quality, chromosome-level genome assemblies available for these
426 four target sea urchin species enabled comparisons at the level of chromosome structure and
427 organization, in addition to sequence-level evolution.

428 Chromosomal rearrangements can alter loci copy number, relationships between linked
429 loci, gene expression patterns and rates of recombination or transposition, and are important
430 contributors to genetic variation that drives phenotypic diversity and speciation.[89] Although the
431 genome sizes of our selected sea urchin species are similar, the number of chromosomes varies
432 between short- and long-lived species with *L. variegatus* and *L pictus* having 19 chromosomes

while *S. purpuratus* and *M. franciscanus* have 21 chromosomes.[29,30,33] Syntenic alignments revealed high inter- and intrachromosomal similarity between the long-lived sea urchins (*M. franciscanus* and *S. purpuratus*) and large- and small-scale chromosomal rearrangements between long- and short-lived species, suggesting dynamic evolutionary processes acting upon these genomes. It is possible that genes whose number or regulation are impacted by genomic rearrangements may affect lifespan; however, inferences about impacts on longevity require closer examination of genes and regulatory elements affected by the rearrangements. Additionally, the power of these comparisons will increase as reference genomes become available for more sea urchin species with different lifespans across a wide evolutionary distance. Evolutionarily, our two short-lived species *L. variegatus* and *L. pictus* diverged approximately 5-10 Mya[90], while the long-lived species *S. purpuratus* and *M. franciscanus* diverged approximately 15-20 Mya[91], and the last common ancestor shared by *L. pictus* and *S. purpuratus* was approximately 40 Mya.[92]

Positive selection analysis was conducted to identify genetic variants in protein coding regions that have undergone adaptive evolution in the long-lived red sea urchin. This analysis revealed more than 400 positively selected genes that are involved in a wide variety of cellular processes including nutrient sensing and metabolism, protein homeostasis, genome regulation and maintenance, mitochondria and peroxisome function. Several of these genes are known to modulate lifespan and aging in short-lived model organisms or promote longevity in long-lived animals (e.g., mTOR, insulin signaling, ER stress response), suggesting these pathways also play a role in sea urchin longevity. Genes involved in nucleic acid metabolic processes represented the largest category under positive selection, including many transcription factors/regulators, splicing factors, and chromatin modifiers, indicating that genome maintenance and

transcriptional regulation play key roles in sea urchin longevity. Network analysis revealed that the products of many genes under positive selection were biologically connected, suggesting a molecular network of interconnected pathways related to negligible senescence. Clusters of co-expressed genes can be used to model gene regulatory events underlying longevity and negligible senescence, and it is notable that many positively selected genes in the red sea urchin were upregulated with age in radial nerve cord. A comprehensive investigation of gene expression patterns across a wide range of tissues in the red sea urchin will help to understand the organizational principles governing gene regulation and provide a framework to experimentally validate putative regulatory relationships.

The expanded gene families related to innate immunity, nervous system function, and nucleic acid metabolism in long-lived sea urchins prompted a targeted assessment of genes within each of these categories in the red sea urchin genome. We identified a highly conserved repertoire of genes involved in genome regulation and maintenance, including multiple copies of some key DNA repair and tumor suppressor genes (e.g., *RB*, *PTEN*, *VHL*, *CHK1, ERCC6, XRCC1*). High levels of expression throughout development suggest an important role in protecting the genome during early life stages. Additionally, the age-related upregulation in some adult tissues, particularly in radial nerve, suggests a protective role in neuronal tissues into advanced years. Expansions of these genes were not unique to the long-lived sea urchins, suggesting that they are not strictly associated with longevity, but instead, may contribute to the absence of neoplasia reported across sea urchin species.[18,19]

The nervous system serves as an interface between an organism's external and internal environments, and neuronal signaling has a major influence on survival and longevity.[93,94] It has been previously reported that environmental cues and signals that affect chemosensory and

thermosensory neurons modulate lifespan in *C. elegans* and *D. melanogaster* [93,94]. It is therefore possible that the extensive repertoire of genes involved in sensory perception identified in the red sea urchin genome promotes longevity through well-developed chemosensation, thermosensation, mechanoreception, and photosensitivity. The results of our comparative genomic analysis together with our previous gene expression study on long-lived red sea urchins[28] suggest that nervous system maintenance may act as a central regulator of negligible senescence and contribute to maintaining the physiology of extraneural tissues. This concept is supported by work in model organisms where manipulating longevity genes in the central nervous system slows aging at a systemic level and can extend lifespan.[95-98] The expansion of the cholinergic system and the LRP4 family suggests efficient synaptogenesis, synaptic organization, and maintenance are important in long-lived sea urchins. The LRP4 family has widespread functions in the nervous system including the formation of the neuromuscular junction, synaptogenesis and synaptic organization in the central and peripheral nervous system, and regeneration of peripheral nerves.[50,51] LRP4 functions as a receptor for a variety of ligands, including Agrin and Wnt, and organizes AChR clustering at the synapse.[50] In addition, LRP4 has recently been shown to promote the clearance of amyloid protein by astrocytes in an Alzheimer's Disease mouse model.[52]

The immune system plays an important role in maintaining tissue homeostasis and protecting the host from pathogenic microbes; both of these functions contribute to aging and longevity. Typically, aging is accompanied by a decline in immune function (i.e., immunosenescence) that predisposes organisms to infections, as well as immune system overactivation that results in chronic inflammation, a major risk factor for a number of aging-related diseases.[99,100] Several lines of evidence support the idea that optimized age-dependent

immune homeostasis promotes longevity.[100] In this study, we identified a robust array of PRRs encoded in the red sea urchin genome, including 167 genes that encode TLRs. However, careful analysis of the gene sequences suggests that just under half (48.5%) of these sequences appear to be pseudogenes. Pseudogenization is one of the primary mechanisms of gene loss. It is therefore possible that this reduced number of functional TLRs may help maintain immune homeostasis by limiting inappropriate inflammatory signals. Activation of PRRs and their corresponding downstream signaling cascades typically induces protective inflammatory responses to remove harmful microbes or initiate tissue repair.[59] However, PRR ligands are not restricted to microbial sources; endogenous DNA can also trigger the expression of inflammatory genes, and DNA sensing pathways associated with immunity have been implicated in the antitumor response.[59,101] It is therefore possible that PRR repertoires are "fine-tuned" to optimize the ability to detect microbes as well as mitigate cancer formation in sea urchins.

Lifespan and senescence vary enormously within and among taxa. Our genomic analyses of the red sea urchin suggest that an integrated network of molecular and systemic processes promote longevity and negligible senescence. Some of the genes and pathways identified in our study are known to play a role in aging and lifespan determination in other animals, while others may have novel contributions to long-term maintenance of tissue homeostasis and disease resistance. Our results provide a foundation for further mechanistic investigation into the roles specific genes play in regulating healthy aging and exploring the potential to translate these strategies to promote life-long good health and disease resistance in humans.

Limitations of the study

Taking advantage of the variation in lifespan between different sea urchin species with limited phenotypic differences, we used a combination of comparative and targeted genomic analyses to identify candidate genes and pathways that contribute to longevity and negligible senescence. Although this approach enables the formulation of testable hypotheses, demonstrating a genotype to phenotype link requires additional validation. The development of new molecular and cell culture tools for sea urchins[102-105] will enable future functional genomics studies to test gene function at the molecular, cellular, and systemic levels. Our positive selection analyses focused on coding regions of the genome; however, changes to regulatory regions can also generate significant phenotypic variation.[106] Future studies measuring epigenetic changes and open chromatin regions will help elucidate the role of non-coding regions in regulatory networks of longevity and negligible senescence. Our comparative analyses were focused on four sea urchin species, but the results will be strengthened as more high-quality reference genomes become available for additional species with different lifespans. Further, including multiple individuals within each species would account for any intraspecific genetic variation that may be present.

Acknowledgments

The authors are sincerely grateful to the Dusky Fund of the Essex County Community Foundation for the funding that supported this work. The authors would like to thank Catalina Offshore Products (San Diego, CA) for supplying the red sea urchin used for sequencing, and Phase Genomics (Seattle, WA) for conducting the Hi-C experiments.

Author Contributions

J.M.P. conducted sequencing, assembly, and annotation of the red sea urchin genome, gene family expansion/contraction analysis and TPM analysis of the sea urchin transcriptomes. K.R.C. analyzed repeat sequences, conducted synteny, gene density, and positive selection analyses, and contributed to the gene family expansion/contraction analysis. K.M.B. analyzed the immune gene repertoire. A.G.B. secured funding for this project, coordinated activities of all participants, conducted targeted analysis of the neuronal and genome stability genes, contributed to data analysis and interpretation, and wrote the manuscript. All authors contributed to generating figures and editing the manuscript.

Data and code availability

Raw sequencing data and the genome assembly are deposited on the National Center of Biotechnology Information (NCBI) under the BioProject accession PRJNA914702. All code used to produce the assembly, analyses, results, and visualizations are available on GitHub: https://github.com/jmpolinski/Red-Sea-Urchin-Genome.

Declaration of Interests

The authors declare no competing interests.

Figure Legends

Figure 1. The red sea urchin genome. (A) The *M. franciscanus* individual used for genomic sequencing. (B) Hi-C contact map showing the 21 longest scaffolds with highest contact density along the diagonal, each outlined square representing a putative chromosome. (C) Genome-wide gene density distribution visualized using a 500 kb window with color scale showing regions of low gene density in blue and high gene density in red.

Figure 2. Syntenic alignments between long- and short-lived sea urchin species. (A) Pairwise syntenic comparisons of chromosomes from *M. franciscancus*, *S. purpuratus, L. pictus* and *L. variegatus*. The vertical lines connect orthologous gene blocks with gray lines representing conserved syntenic blocks and colored lines indicating chromosomal rearrangements between species. (B-E) Oxford plots visualizing alignments of orthologous genes across individual chromosomes showing pairwise comparisons for long-lived (*M. franciscanus*, *S. purpuratus*), long- versus short-lived (*M. franciscanus, L. pictus* and *M. franciscanus, L. variegatus*) and short-lived (*L. pictus*, *L. variegatus*) sea urchins.

Figure 3. Genes under positive selection in the red sea urchin genome. (A) Number of positively selected genes assigned to each functional category using a custom sea urchin gene ontology. (B) STRING protein-protein interaction (PPI) network for proteins encoded by positively selected genes, showing only the nodes with one or more connections and using gene identifiers for the closest human homolog. Nodes assigned to the GO terms; mitochondrion (GO:0005739), protein localization (GO:0008104), and nucleic acid metabolic process (GO:0090304), are shown by different colors. Results of STRING analysis using sea urchin gene identifiers are shown in Figure S3. (C) Expression patterns of the positively selected genes represented in the PPI network across three adult tissues (muscle, esophagus and radial nerve cord) from young and old red sea urchins. Dendrograms on the left side indicate hierarchical clustering of homologs. (D) Sankey diagram showing overlap of the positively selected genes in the red sea urchin genome with genes that are known to modulate aging and longevity in other animals. Black font represents pro-longevity and red font represents anti-longevity genes.

Figure 4. Echinoderm phylogenetic tree and gene family expansion/contraction. (A) Phylogenetic tree inferred from 1,245 single-copy orthologs among eight echinoderm species

592 and *B. floridae* as the outgroup with numbers of expanded gene families marked in blue and
593 numbers of contracted gene families marked in red at each branch. (B) The 73 significantly
594 expanded gene families in long-lived sea urchins assigned to functional categories using a
595 custom sea urchin gene ontology.

596 **Figure 5. The red sea urchin genome contains a complex repertoire of genes involved in**
597 **immune response.** An overview of the putative red sea urchin immune system is shown.
598 Protein localization is based on available data in other systems. The intracellular signaling
599 cascades have been excluded for simplicity. Numbers in parentheses indicate the total number of
600 paralogs identified. Domain architectures are representative of the paralog groups; some
601 variation among individual proteins may occur. Protein domains are as follows: BROMO,
602 bromodomain (PF00439); CARD, caspase recruitment domain (PF00619); CTD, RIG-I receptor
603 C-terminal domain (PF18119); DEATH, death domain (PF00531); Helicase, DEAD-like
604 helicases superfamily (SM000487); LRR, leucine-rich repeat (PF00560); NACHT (PF05729);
605 PHD, plant homeodomain (SM00249); TIR, Toll/IL-1R region (PF01582); TRIM, tripartite
606 motif that consists of a RING-B-box-coiled-coil domains; SEFIR, Sef/IL-17R region (PF08357).

607 **Figure 6. Neuronal genes encoded in the red sea urchin genome.** (A) Number of homologs
608 encoding ion channels, ligand-gated ion channels (LGICs), and GPCRs in the red sea urchin (*M.*
609 *franciscanus*) genome compared to other metazoan species. (B) Number of TRP channels and
610 nAChR encoded in the red sea urchin genome compared to other metazoan species. (C)
611 Phylogenic tree representing the different classes of GPCRs encoded in the red sea urchin
612 genome. (D) Schematic representation of GRL101 showing the extracellular leucine rich repeat
613 (LRR) and low-density lipoprotein class A (LDLa) motifs, created using Protter.[107]

Figure 7. Tumor suppressor and DNA repair genes encoded in the red sea urchin genome. (A) Gene structure of selected tumor suppressor and DNA repair genes in the red sea urchin, *M. franciscanus*, genome showing exons as colored blocks and introns as black lines. The scale bar applies to exons only. Intron sizes shown are standardized and not representative of sequence length. (B) Heatmap of normalized gene expression levels for developmental stages (egg, 8-16 cell stage, morula, blastula, gastrula, prism, pluteus) and adult tissues (Aristotle's lantern muscle (ALM), esophagus, and radial nerve cord from young and old red sea urchins) represented as a gradient of log2 TPM (transcripts per million). (C) Number of homologs for the selected genes in sea urchin species *M. franciscanus* (MFR), *S. purpuratus* (SPU), *L. variegatus* (LVA), and *L. pictus* (LPI).

STAR Methods

DNA extraction from red sea urchin tissue

Gonad tissue from one male red sea urchin, *M. franciscanus* (Catalina Offshore Products, San Diego, CA), was stored frozen at -80°C until DNA extraction using a modified cetyltrimethylammonium bromide (CTAB) extraction procedure.[108] Briefly, gonad tissue was homogenized with a sterile pestle in a solution of 2% CTAB, 0.1% β-mercaptoethanol, 100 mM Tris (pH 8.0), 1.4 M NaCl, and 20 mM EDTA. Proteinase K was added to a final concentration of 0.5 mg/ml and incubated at 65°C for 1.5 hours. The sample was cooled briefly and RNase A was added at a final concentration of 20 µg/ml and incubated at 37°C for 15 minutes. The sample was extracted with chloroform and gDNA precipitated with ethanol. The gDNA was resuspended in 10 mM Tris-HCl and quantified with the Qubit broad-range dsDNA assay (Thermo Fisher Scientific, Waltham, MA).

Illumina library preparation and sequencing

gDNA was sheared to approximately 450 base pair (bp) fragments on a Covaris M220. A cleanup was performed with PCRclean DX beads (Aline Biosciences, Woburn, MA) to remove fragments less than 200 bp. The sequencing library was prepped with KAPA Hyper Prep DNA reagents (Roche Sequencing Solutions, Pleasanton, CA) following the manufacturer's guidelines for PCR-free library preparation. A double-sided size selection with ratios of 0.45-0.7X was performed with PCRclean DX beads to select for fragments that were approximately 450 to 600 bp. Library size was confirmed on a Fragment Analyzer (Agilent, Santa Clara, CA) and concentrations of adapter-ligated library were quantified via qPCR with the NEBNext® Library Quant Kit for Illumina (New England Biolabs, Ipswich, MA) prior to 2x150 paired-end sequencing on an Illumina NextSeq 500.

Oxford Nanopore library preparation and sequencing

For Oxford Nanopore Technologies (ONT) long-read sequencing, gDNA was enriched for fragments >1.5 kilobase (kb) using SPRI beads in modified buffers (0.5 M $MgCl_2$ and 5% PEG8000[109] and 10 mM Tris-HCl, 1 mM EDTA pH 8.0, 1.6 M NaCl, 11% PEG8000) for selecting DNA >1.5 kb. Average fragment sizes were determined on a Fragment Analyzer and concentrations were measured with the Qubit Broad Range dsDNA assay (Thermo Fisher Scientific, Waltham, MA). 1 µg of HMW DNA was used as starting material for the Nanopore Ligation Sequencing Kit (Oxford Nanopore Technologies, Oxford, UK). ONT libraries were prepared following the manufacturer's guidelines and sequenced on a MinION sequencer with R9.4.1 flow cells.

Hi-C library preparation and sequencing

Gonad tissue from the same male red sea urchin (*M. franciscanus*) used for DNA extraction and sequencing was frozen at -80°C immediately after dissection and shipped on dry ice to Phase Genomics for Hi-C library preparation with the Proximo™ Hi-C Kit (Phase Genomics, Seattle, WA). Sequencing of the Hi-C library was done on an Illumina NovaSeq.

Red sea urchin genome assembly

Fast5 files were converted to fastq format with the ONT Guppy basecalling software v4.2.2 using the fast algorithm, and NanoFilt v2.6.0[110] was used to trim the first 50 bases and filter out sequences with a minimum average read quality score of 10 or shorter than 1000 bp. Illumina data were trimmed and filtered with Trimmomatic v0.38 (LEADING:3 TRAILING:3 SLIDINGWINDOW:4:15 MINLEN:50).

Multiple assembly pipelines were employed to generate preliminary assemblies and the most contiguous and complete version, based on N50/L50 values and Benchmarking Single Copy Ortholog analysis (BUSCO v5.1.2 with gene set metazoa_odb10), was chosen as the preliminary assembly. The most complete preliminary assembly was generated with CANU v2.1.1[31] using only long-read sequencing data. The genome polishing tool POLCA[32], distributed with MaSuRCa v4.0.3, used Illumina data to correct errors in the long-read assembly. BUSCO analysis indicated that the assembly was highly duplicated with a completeness score of 97.6% but only 22.9% single-copy. Duplicate regions of the assembly were removed following steps outlined in the purge_dups pipeline (https://github.com/dfguan/purge_dups), using read depth metrics and a self-alignment with Minimap2 v2.18 to identify and remove duplicates and haplotigs.

Sequences shorter than 15 kb, 161 contigs totaling 1.66 Mb, were discarded from the assembly prior to scaffolding with Hi-C data. Juicer v1.6 with aligner BWA v0.7.17 was used to

analyze Hi-C data. The "merged_nodups.txt" file containing a list of paired alignments from the Hi-C data to the draft assembly was used as input with the draft assembly in 3D-DNA v190716, and Juicebox Assembly Tools (JBAT) was used to visualize Hi-C contact data and manually correct scaffolds.[111]

Gene model prediction, annotation, and expression analysis

RNA-seq data from a variety of developmental stages of the red sea urchin (*M. franciscanus*) was downloaded from SRA (NCBI project accession PRJNA637102, PRJNA531463, and PRJNA272924). Previously published RNA-seq data from Aristotle's lantern, esophagus, and radial nerve cord from young (<5 cm test diameter) and old (>15 cm test diameter) red sea urchins was also included (PRJNA562282).[28] All RNA-seq datasets were mapped to the genome assembly with HISAT2.[112] All individual sample alignment data was combined into a single BAM file that was used as evidence input into the BRAKER2 gene model prediction program.[113] After a preliminary run, BRAKER was rerun using the "hintsfile.gff" and Augustus species from the first run, resulting in an improved gene set based on BUSCO analysis. gFACs[114] was used to extract unique gene models from the BRAKER GTF file, and genes of particular interest were manually corrected based on visualizing mapped RNA-seq data in Tablet genome viewer.[115] Gene models were annotated using EnTAP[116] v0.10.8 and BLAST v2.7.1 searches against the SwissProt and RefSeq protein databases. Genome wide gene density was visualized per 500 kb windows using BEDTools v2.29.0 and RIdeogram v0.2.2.[117,118]

Gene expression analysis was conducted using the final corrected gene feature file (GFF) and genome assembly. RNA-seq data from the same SRA project and tissue type listed above were combined and mapped to the genome with the STAR aligner v2.7.3a [119], and the resulting BAM

files were used in edgeR[120] to generate gene counts and transcripts per million (TPM) expression data.

Repeat analysis

Repeat analysis was conducted on the newly assembled *M. franciscancus* genome, *S. purpuratus* genome (version 5.0, https://www.echinobase.org), *L. variegatus* genome (version 3.0, https://www.echinobase.org)[29,121], and *L. pictus* genome (version 2.0, NCBI).[30] Repeats were annotated with RepeatMasker[34] v 4.1.0 using a *de novo* repeat library. The de novo library was made using a combination of RepeatModeler[35] v2.0.2, Transposon PSI v1.0.0 (https://transposonpsi.sourceforge.net/), and LTRharvest[122] v1.6.2 to better identify diverged and species-specific repeats. Each program was run individually before combining and clustering (Usearch v11.0.667)[123], with a minimum sequence match of 80%, to remove redundant sequences that may have been identified in multiple programs.

Gene family expansion and contraction analysis

Eight echinoderm species including the sea urchins *M. franciscanus*, *S. purpuratus*, *L. variegatus*, *L. pictus*, the sea cucumber *Apostichopus japonicus*, the sea stars *Acanthaster planci* and *Asterias rubens*, and the sea lily *Anneissia japonica*, were chosen for gene family expansion/contraction analysis based on genome and protein quality using BUSCO v5.3.1 (Table S12).[124] *Branchiostoma floridae* was chosen as the outgroup species. Protein files were retrieved from Echinobase[125] or NCBI. Protein files were filtered using gFACS[114] v3 to remove isoforms for orthology analysis (--unique-genes-only flag). OrthoFinder[126] v2.5.4 was used to identify orthologs and paralogs across the gene model sets of the nine selected species using default settings (Table S12). Ortholog group counts and the rooted species tree, converted to an

ultrametric tree using ete2[127], were used as input for CAFE[128] v4.2.1 gene family expansion and contraction analysis.

Synteny analysis

Pairwise synteny comparisons were inferred between *M. franciscancus* and *S. purpuratus, M. franciscancus* and *L. variegatus, M. franciscancus* and *L. pictus,* and *L. variegatus* and *L. pictus* using MCscan.[129] *S. purpuratus* version 5.0 gene annotations were obtained from Echinobase (https://www.echinobase.org), version 3.0 of the *L. variegatus* gene annotations obtained from Echinobase[29,121], and version 2.0 of the *L. pictus* gene annotations obtained from NCBI.[30] LAST v1445 was used for genome wide alignments of coding regions, and filtering of tandem duplications and weak hits. Linkage clustering into syntenic blocks and visualization was performed with the MCscan python workflow (https://github.com/tanghaibao/jcvi/wiki/MCscan-(Python-version). Microsynteny visualization of the Hox cluster was also completed through MCscan modules.

Positive selection analysis

Positive selection analysis was conducted on the genomic coding regions of four sea urchin species with different lifespans (long-lived *M. franciscanus* and *S. purpuratus*, and short-lived *L. variegatus* and *L. pictus*) with *B. floridae* as the outgroup. OrthoFinder[126] v2.5.4 was run with the "-M msa" flag to include multiple sequence alignments. Orthogroups were filtered to keep only those containing all species. Pal2nal[130] v14 was used to convert multiple sequence files to phyllip format. Phylogenetic analysis of codons was completed using Phylogenetic Analysis of Maximum Likelihood (PAML) codeml v4.9 to identify branch-specific selection, setting *M. franciscanus* as the foreground branch.[131,132] For the null hypothesis or no selection, omega is set to 1. For the alternative hypothesis or positive selection, omega is set to >1. To

identify genes under positive selection in *M. franciscanus*, the log ratio test (LRT) was used to compare the log likelihood ratio (InL) values between the null and alternate hypotheses and a chi-square distribution test was performed to identify significant genes using $p<0.001$ as the cutoff. GO and KEGG enrichment analysis and protein-protein interactions among the positively selected genes were investigated using STRING v11.5 (http://www.stringdb.org).[38] STRING was run with the high confidence setting (0.7) and all active interaction sources selected (textmining, experiments, databases, co-expression, neighborhood, gene fusion and co-occurrence). STRING was run using sea urchin gene identifiers for the closest *S. purpuratus* homolog as input and selecting *S. purpuratus* as the target organism and also using the closest human homologs as input and selecting *Homo sapiens* as the target organism.[38]

Supplemental Information

Table S1 – Statistics for the *M. franciscanus* genome assembly and gene predictions
Table S2 – *M. franciscanus* gene models with annotations and gene expression analysis
Table S3 – Repeat elements in the *M. franciscanus*, *S. purpuratus*, *L. pictus*, and *L. variegatus* genomes
Table S4 – *M. franciscanus* genes under positive selection
Table S5 – GO and KEGG pathway enrichment for genes under positive selection in the *M. franciscanus* genome
Table S6 – *M. franciscanus* genes under positive selection that are known longevity genes in other organisms
Table S7 – Gene family expansion and contraction analysis
Table S8 – Immune gene repertoire in the *M. franciscanus* genome
Table S9 – Nervous system gene repertoire in the *M. franciscanus* genome
Table S10 – Number of ion channels and GPCRs encoded in the genome of the red sea urchin (*M. franciscanus*) and other metazoans
Table S11 – Genome Stability gene repertoire in the *M. franciscanus* genome
Table S12 – Species used as input for gene family expansion and contraction analysis
Figure S1 – Percent of repetitive elements in the *M. franciscanus*, *S. purpuratus*, *L. pictus* and *L. variegatus* genomes
Figure S2 – Syntenic alignments between the *M. franciscanus* and *L. variegatus* genomes (A) and the Hox gene cluster (B).

Figure S3 – STRING protein-protein interaction network for genes under positive selection in the *M. franciscanus* genome using gene identifiers for the closest *S. purpuratus* homolog as input and selecting *S. purpuratus* as the target organism. High confidence setting (0.7) was used in STRING-DB and only nodes with one or more interactions are shown. Nodes assigned to the selected GO terms are colored as indicated in the legend.

Figure S4 – Gene expression patterns of the positively selected genes in the *M. franciscanus* genome. (A) Expression of positively selected genes throughout early development and in adult tissues [young and old muscle (ALM), esophagus (ES) and radial nerve cord (RN)]. Panel B: Expression of positively selected genes in young and old muscle (ALM), esophagus (ES) and radial nerve cord (RN).

References

1. Finch, C.E., and Austad, S.N. (2001). History and prospects: symposium on organisms with slow aging. Experimental Gerontology *36*, 593-597. https://doi.org/10.1016/S0531-5565(00)00228-X.
2. Gorbunova, V., Seluanov, A., Zhang, Z., Gladyshev, V.N., and Vijg, J. (2014). Comparative genetics of longevity and cancer: insights from long-lived rodents. Nat Rev Genet *15*, 531-540. 10.1038/nrg3728.
3. Keane, M., Semeiks, J., Webb, A.E., Li, Y.I., Quesada, V., Craig, T., Madsen, L.B., van Dam, S., Brawand, D., Marques, P.I., et al. (2015). Insights into the evolution of longevity from the bowhead whale genome. Cell reports *10*, 112-122. 10.1016/j.celrep.2014.12.008.
4. Zhou, X., Dou, Q., Fan, G., Zhang, Q., Sanderford, M., Kaya, A., Johnson, J., Karlsson, E.K., Tian, X., Mikhalchenko, A., et al. (2020). Beaver and Naked Mole Rat Genomes Reveal Common Paths to Longevity. Cell Rep *32*, 107949. 10.1016/j.celrep.2020.107949.
5. Kolora, S.R.R., Owens, G.L., Vazquez, J.M., Stubbs, A., Chatla, K., Jainese, C., Seeto, K., McCrea, M., Sandel, M.W., Vianna, J.A., et al. (2021). Origins and evolution of extreme life span in Pacific Ocean rockfishes. Science *374*, 842-847. doi:10.1126/science.abg5332.
6. Treaster, S., Deelen, J., Daane, J.M., Murabito, J., Karasik, D., and Harris, M.P. (2023). Convergent genomics of longevity in rockfishes highlights the genetics of human life span variation. Science advances *9*, eadd2743. 10.1126/sciadv.add2743.
7. Sahm, A., Bens, M., Szafranski, K., Holtze, S., Groth, M., Görlach, M., Calkhoven, C., Müller, C., Schwab, M., Kraus, J., et al. (2018). Long-lived rodents reveal signatures of positive selection in genes associated with lifespan. PLOS Genetics *14*, e1007272. 10.1371/journal.pgen.1007272.
8. Glaberman, S., Bulls, S.E., Vazquez, J.M., Chiari, Y., and Lynch, V.J. (2021). Concurrent Evolution of Antiaging Gene Duplications and Cellular Phenotypes in Long-Lived Turtles. Genome Biology and Evolution *13*. 10.1093/gbe/evab244.
9. Yu, Z., Seim, I., Yin, M., Tian, R., Sun, D., Ren, W., Yang, G., and Xu, S. (2021). Comparative analyses of aging-related genes in long-lived mammals provide insights into

natural longevity. Innovation (Cambridge (Mass.)) *2*, 100108. 10.1016/j.xinn.2021.100108.
10. McClay, D.R. (2011). Evolutionary crossroads in developmental biology: sea urchins. Development *138*, 2639-2648. 10.1242/dev.048967.
11. Sodergren, E., Weinstock, G.M., Davidson, E.H., Cameron, R.A., Gibbs, R.A., Angerer, R.C., Angerer, L.M., Arnone, M.I., Burgess, D.R., and Burke, R.D. (2006). The genome of the sea urchin Strongylocentrotus purpuratus. Science *314*, 941-952.
12. Ernst, S.G. (2011). Offerings from an Urchin. Developmental Biology *358*, 285-294. https://doi.org/10.1016/j.ydbio.2011.06.021.
13. Lawrence, J.M. (2020). Sea Urchins: Biology and Ecology, Fourth Edition (Elsevier B.V.).
14. Ebert, T.A. (2008). Longevity and lack of senescence in the red sea urchin Strongylocentrotus franciscanus. Exp Gerontol *43*, 734-738. 10.1016/j.exger.2008.04.015.
15. Bodnar, A.G. (2009). Marine invertebrates as models for aging research. Exp Gerontol *44*, 477-484. 10.1016/j.exger.2009.05.001.
16. Bodnar, A.G. (2015). Cellular and molecular mechanisms of negligible senescence: insight from the sea urchin. Invertebrate reproduction & development *59*, 23-27.
17. Ebert, T.A. (2019). Negative senescence in sea urchins. Experimental Gerontology *122*, 92-98. https://doi.org/10.1016/j.exger.2019.04.018.
18. Robert, J. (2010). Comparative study of tumorigenesis and tumor immunity in invertebrates and nonmammalian vertebrates. Dev Comp Immunol *34*, 915-925. 10.1016/j.dci.2010.05.011.
19. Jangoux, M. (1987). Diseases of Echinodermata. IV. Structural abnormalities and general considerations on biotic diseases. Diseases of aquatic organisms *3*, 221-229.
20. Beddingfield, S.D., and McClintock, J.B. (2000). Demographic characteristics of Lytechinus variegatus (Echinoidea: Echinodermata) from three habitats in a North Florida Bay, Gulf of Mexico. Marine Ecology *21*, 17-40.
21. Detwiler, P.M. (1996). Demography, Growth, Mortality and Resource Allocation in the White Sea Urchin Lytechinus pictus. MS Thesis MS (San Diego State University).
22. Ebert, T.A. (2007). Growth and survival of postsettlement sea urchins. In Developments in Aquaculture and Fisheries Science, J.M. Lawrence, ed. (Elsevier), pp. 95-134. https://doi.org/10.1016/S0167-9309(07)80070-6.
23. Ebert, T.A. (2010). Demographic patterns of the purple sea urchin Strongylocentrotus purpuratus along a latitudinal gradient, 1985–1987. Marine Ecology Progress Series *406*, 105-120. 10.3354/meps08547.
24. Francis, N., Gregg, T., Owen, R., Ebert, T., and Bodnar, A. (2006). Lack of age-associated telomere shortening in long-and short-lived species of sea urchins. FEBS letters *580*, 4713-4717.
25. Du, C., Anderson, A., Lortie, M., Parsons, R., and Bodnar, A. (2013). Oxidative damage and cellular defense mechanisms in sea urchin models of aging. Free Radical Biology and Medicine *63*, 254-263.
26. Bodnar, A.G., and Coffman, J.A. (2016). Maintenance of somatic tissue regeneration with age in short-and long-lived species of sea urchins. Aging cell *15*, 778-787.

27. Loram, J., and Bodnar, A. (2012). Age-related changes in gene expression in tissues of the sea urchin Strongylocentrotus purpuratus. Mechanisms of ageing and development *133*, 338-347.
28. Polinski, J.M., Kron, N., Smith, D.R., and Bodnar, A.G. (2020). Unique age-related transcriptional signature in the nervous system of the long-lived red sea urchin Mesocentrotus franciscanus. Scientific Reports *10*, 9182. 10.1038/s41598-020-66052-3.
29. Davidson, P.L., Guo, H., Wang, L., Berrio, A., Zhang, H., Chang, Y., Soborowski, A.L., McClay, D.R., Fan, G., and Wray, G.A. (2020). Chromosomal-Level Genome Assembly of the Sea Urchin Lytechinus variegatus Substantially Improves Functional Genomic Analyses. Genome Biology and Evolution *12*, 1080-1086. 10.1093/gbe/evaa101.
30. Warner, J.F., Lord, J.W., Schreiter, S.A., Nesbit, K.T., Hamdoun, A., and Lyons, D.C. (2021). Chromosomal-Level Genome Assembly of the Painted Sea Urchin Lytechinus pictus: A Genetically Enabled Model System for Cell Biology and Embryonic Development. Genome Biology and Evolution *13*. 10.1093/gbe/evab061.
31. Koren, S., Walenz, B.P., Berlin, K., Miller, J.R., Bergman, N.H., and Phillippy, A.M. (2017). Canu: scalable and accurate long-read assembly via adaptive k-mer weighting and repeat separation. Genome Res *27*, 722-736. 10.1101/gr.215087.116.
32. Zimin, A.V., and Salzberg, S.L. (2020). The genome polishing tool POLCA makes fast and accurate corrections in genome assemblies. PLOS Computational Biology *16*, e1007981. 10.1371/journal.pcbi.1007981.
33. Eno, C.C., Bottger, S.A., and Walker, C.W. (2009). Methods for karyotyping and for localization of developmentally relevant genes on the chromosomes of the purple sea urchin, Strongylocentrotus purpuratus. The Biological Bulletin *217*, 306-312.
34. Smit, A.F.A., Hubley, R., and Green, P. (2013-2015). RepeatMasker Open-4.0. http://www.repeatmasker.org.
35. Flynn, J.M., Hubley, R., Goubert, C., Rosen, J., Clark, A.G., Feschotte, C., and Smit, A.F. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. Proceedings of the National Academy of Sciences *117*, 9451-9457. 10.1073/pnas.1921046117.
36. Uehara, T., Shingaki, M., Taira, K., Arakaki, Y., and Nakatomi, H. (2020). Chromosome studies in eleven Okinawan species of sea urchins, with special reference to four species of the Indo-Pacific Echinometra. In Biology of echinodermata, (CRC Press), pp. 119-129. doi.org/10.1201/9781003077565.
37. Tu, Q., Cameron, R.A., Worley, K.C., Gibbs, R.A., and Davidson, E.H. (2012). Gene structure in the sea urchin Strongylocentrotus purpuratus based on transcriptome analysis. Genome Research. 10.1101/gr.139170.112.
38. Szklarczyk, D., Gable, A.L., Nastou, K.C., Lyon, D., Kirsch, R., Pyysalo, S., Doncheva, N.T., Legeay, M., Fang, T., Bork, P., et al. (2020). The STRING database in 2021: customizable protein–protein networks, and functional characterization of user-uploaded gene/measurement sets. Nucleic Acids Research *49*, D605-D612. 10.1093/nar/gkaa1074.
39. Tacutu, R., Thornton, D., Johnson, E., Budovsky, A., Barardo, D., Craig, T., Diana, E., Lehmann, G., Toren, D., Wang, J., et al. (2017). Human Ageing Genomic Resources: new and updated databases. Nucleic Acids Research *46*, D1083-D1090. 10.1093/nar/gkx1042.
40. Wortel, I.M.N., van der Meer, L.T., Kilberg, M.S., and van Leeuwen, F.N. (2017). Surviving Stress: Modulation of ATF4-Mediated Stress Responses in Normal and

Malignant Cells. Trends in endocrinology and metabolism: TEM *28*, 794-806. 10.1016/j.tem.2017.07.003.
41. Martins, R., Lithgow, G.J., and Link, W. (2016). Long live FOXO: unraveling the role of FOXO proteins in aging and longevity. Aging Cell *15*, 196-207. https://doi.org/10.1111/acel.12427.
42. Barbieri, M., Bonafè, M., Franceschi, C., and Paolisso, G. (2003). Insulin/IGF-I-signaling pathway: an evolutionarily conserved mechanism of longevity from yeast to humans. American Journal of Physiology-Endocrinology and Metabolism *285*, E1064-E1071. 10.1152/ajpendo.00296.2003.
43. Weichhart, T. (2017). mTOR as Regulator of Lifespan, Aging, and Cellular Senescence: A Mini-Review. Gerontology *64*, 127-134. 10.1159/000484629.
44. Kenyon, C.J. (2010). The genetics of ageing. Nature *464*, 504-512. 10.1038/nature08980.
45. Oxvig, C. (2015). The role of PAPP-A in the IGF system: location, location, location. Journal of cell communication and signaling *9*, 177-187. 10.1007/s12079-015-0259-9.
46. Banks, A.S., Li, J., McKeag, L., Hribal, M.L., Kashiwada, M., Accili, D., and Rothman, P.B. (2005). Deletion of SOCS7 leads to enhanced insulin action and enlarged islets of Langerhans. J Clin Invest *115*, 2462-2471. 10.1172/jci23853.
47. Betsinger, C.N., and Cristea, I.M. (2019). Mitochondrial Function, Metabolic Regulation, and Human Disease Viewed through the Prism of Sirtuin 4 (SIRT4) Functions. Journal of Proteome Research *18*, 1929-1938. 10.1021/acs.jproteome.9b00086.
48. Heyland, A., Price, D.A., Bodnarova-buganova, M., and Moroz, L.L. (2006). Thyroid hormone metabolism and peroxidase function in two non-chordate animals. Journal of Experimental Zoology Part B: Molecular and Developmental Evolution *306B*, 551-566. doi:10.1002/jez.b.21113.
49. Taylor, E., Wynen, H., and Heyland, A. (2023). Thyroid hormone membrane receptor binding and transcriptional regulation in the sea urchin Strongylocentrotus purpuratus. Frontiers in endocrinology *14*, 1195733. 10.3389/fendo.2023.1195733.
50. DePew, A.T., and Mosca, T.J. (2021). Conservation and Innovation: Versatile Roles for LRP4 in Nervous System Development. Journal of developmental biology *9*. 10.3390/jdb9010009.
51. Gribble, K.D., Walker, L.J., Saint-Amant, L., Kuwada, J.Y., and Granato, M. (2018). The synaptic receptor Lrp4 promotes peripheral nerve regeneration. Nature Communications *9*, 2389. 10.1038/s41467-018-04806-4.
52. Zhang, H., Chen, W., Tan, Z., Zhang, L., Dong, Z., Cui, W., Zhao, K., Wang, H., Jing, H., Cao, R., et al. (2020). A Role of Low-Density Lipoprotein Receptor-Related Protein 4 (LRP4) in Astrocytic Aβ Clearance. J Neurosci *40*, 5347-5361. 10.1523/jneurosci.0250-20.2020.
53. Husmann, D., and Gozani, O. (2019). Histone lysine methyltransferases in biology and disease. Nature Structural & Molecular Biology *26*, 880-889. 10.1038/s41594-019-0298-7.
54. Traylor-Knowles, N., Rose, N.H., Sheets, E.A., and Palumbi, S.R. (2017). Early Transcriptional Responses during Heat Stress in the Coral Acropora hyacinthus. Biol Bull *232*, 91-100. 10.1086/692717.
55. Hibino, T., Loza-Coll, M., Messier, C., Majeske, A.J., Cohen, A.H., Terwilliger, D.P., Buckley, K.M., Brockton, V., Nair, S.V., Berney, K., et al. (2006). The immune gene

repertoire encoded in the purple sea urchin genome. Dev Biol *300*, 349-365. 10.1016/j.ydbio.2006.08.065.
56. Buckley, K.M., and Rast, J.P. (2015). Diversity of animal immune receptors and the origins of recognition complexity in the deuterostomes. Developmental & Comparative Immunology *49*, 179-189. https://doi.org/10.1016/j.dci.2014.10.013.
57. Buckley, K.M., and Rast, J.P. (2012). Dynamic evolution of toll-like receptor multigene families in echinoderms. Front Immunol *3*, 136. 10.3389/fimmu.2012.00136.
58. Buckley, K.M., and Rast, J.P. (2011). Characterizing Immune Receptors from New Genome Sequences. In Immune Receptors: Methods and Protocols, J.P. Rast, and J.W.D. Booth, eds. (Humana Press), pp. 273-298. 10.1007/978-1-61779-139-0_19.
59. Okude, H., Ori, D., and Kawai, T. (2020). Signaling Through Nucleic Acid Sensors and Their Roles in Inflammatory Diseases. Front Immunol *11*, 625833. 10.3389/fimmu.2020.625833.
60. Secombes, C.J., and Zou, J. (2017). Evolution of Interferons and Interferon Receptors. Front Immunol *8*, 209. 10.3389/fimmu.2017.00209.
61. Yang, W., Gu, Z., Zhang, H., and Hu, H. (2020). To TRIM the Immunity: From Innate to Adaptive Immunity. Front Immunol *11*, 02157. 10.3389/fimmu.2020.02157.
62. van Gent, M., Sparrer, K.M.J., and Gack, M.U. (2018). TRIM Proteins and Their Roles in Antiviral Host Defenses. Annual review of virology *5*, 385-405. 10.1146/annurev-virology-092917-043323.
63. Koepke, L., Gack, M.U., and Sparrer, K.M. (2021). The antiviral activities of TRIM proteins. Current opinion in microbiology *59*, 50-57. 10.1016/j.mib.2020.07.005.
64. Ali, H., Mano, M., Braga, L., Naseem, A., Marini, B., Vu, D.M., Collesi, C., Meroni, G., Lusic, M., and Giacca, M. (2019). Cellular TRIM33 restrains HIV-1 infection by targeting viral integrase for proteasomal degradation. Nat Commun *10*, 926. 10.1038/s41467-019-08810-0.
65. Aucagne, R., Droin, N., Paggetti, J., Lagrange, B., Largeot, A., Hammann, A., Bataille, A., Martin, L., Yan, K.P., Fenaux, P., et al. (2011). Transcription intermediary factor 1γ is a tumor suppressor in mouse and human chronic myelomonocytic leukemia. J Clin Invest *121*, 2361-2370. 10.1172/jci45213.
66. Quéré, R., Saint-Paul, L., Carmignac, V., Martin, R.Z., Chrétien, M.L., Largeot, A., Hammann, A., Pais de Barros, J.P., Bastie, J.N., and Delva, L. (2014). Tif1γ regulates the TGF-β1 receptor and promotes physiological aging of hematopoietic stem cells. Proc Natl Acad Sci U S A *111*, 10592-10597. 10.1073/pnas.1405546111.
67. Tsuchida, T., Zou, J., Saitoh, T., Kumar, H., Abe, T., Matsuura, Y., Kawai, T., and Akira, S. (2010). The ubiquitin ligase TRIM56 regulates innate immune responses to intracellular double-stranded DNA. Immunity *33*, 765-776. 10.1016/j.immuni.2010.10.013.
68. Fu, L., Zhou, X., Jiao, Q., and Chen, X. (2023). The Functions of TRIM56 in Antiviral Innate Immunity and Tumorigenesis. Int J Mol Sci *24*, 5046.
69. Yang, D., Zhou, Q., Labroska, V., Qin, S., Darbalaei, S., Wu, Y., Yuliantie, E., Xie, L., Tao, H., Cheng, J., et al. (2021). G protein-coupled receptors: structure- and function-based drug discovery. Signal transduction and targeted therapy *6*, 7. 10.1038/s41392-020-00435-w.
70. Roberts, R.E., Motti, C.A., Baughman, K.W., Satoh, N., Hall, M.R., and Cummins, S.F. (2017). Identification of putative olfactory G-protein coupled receptors in Crown-of-

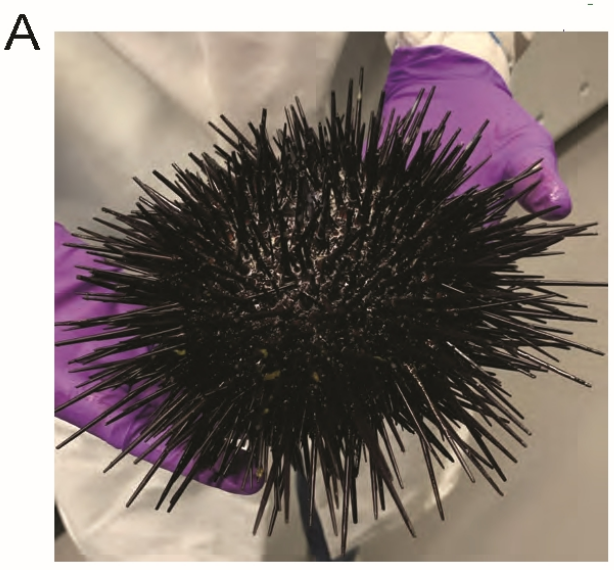
Thorns starfish, Acanthaster planci. BMC Genomics *18*, 400. 10.1186/s12864-017-3793-4.
71. Tensen, C.P., Van Kesteren, E.R., Planta, R.J., Cox, K.J., Burke, J.F., van Heerikhuizen, H., and Vreugdenhil, E. (1994). A G protein-coupled receptor with low density lipoprotein-binding motifs suggests a role for lipoproteins in G-linked signal transduction. Proceedings of the National Academy of Sciences *91*, 4816-4820. doi:10.1073/pnas.91.11.4816.
72. Roberts, R.E., Powell, D., Wang, T., Hall, M.H., Motti, C.A., and Cummins, S.F. (2018). Putative chemosensory receptors are differentially expressed in the sensory organs of male and female crown-of-thorns starfish, Acanthaster planci. BMC Genomics *19*, 853. 10.1186/s12864-018-5246-0.
73. Ding, J., Yu, Y., Yang, M., Shi, D., Li, Z., Chi, X., Chang, Y., Wang, Q., and Zhao, C. (2019). TRPA1 expression provides new insights into thermal perception by the sea urchin Strongylocentrotus intermedius. Journal of the Marine Biological Association of the United Kingdom *99*, 1825-1829. 10.1017/S0025315419000882.
74. Angelini, C., Baccetti, B., Piomboni, P., Trombino, S., Aluigi, M.G., Stringara, S., Gallus, L., and Falugi, C. (2004). Acetylcholine synthesis and possible functions during sea urchin development. European journal of histochemistry : EJH *48*, 235-243.
75. Jennings, N.A., Pezzementi, L., Lawrence, A.L., and Watts, S.A. (2008). Acetylcholinesterase in the sea urchin Lytechinus variegatus: characterization and developmental expression in larvae. Comp Biochem Physiol B Biochem Mol Biol *149*, 401-409. 10.1016/j.cbpb.2007.10.014.
76. Jiao, Y., Cao, Y., Zheng, Z., Liu, M., and Guo, X. (2019). Massive expansion and diversity of nicotinic acetylcholine receptors in lophotrochozoans. BMC Genomics *20*, 937. 10.1186/s12864-019-6278-9.
77. Spivak, G. (2015). Nucleotide excision repair in humans. DNA Repair (Amst) *36*, 13-18. 10.1016/j.dnarep.2015.09.003.
78. Brooks, P.J. (2013). Blinded by the UV light: how the focus on transcription-coupled NER has distracted from understanding the mechanisms of Cockayne syndrome neurologic disease. DNA Repair (Amst) *12*, 656-671. 10.1016/j.dnarep.2013.04.018.
79. McNeill, D.R., Lin, P.-C., Miller, M.G., Pistell, P.J., de Souza-Pinto, N.C., Fishbein, K.W., Spencer, R.G., Liu, Y., Pettan-Brewer, C., Ladiges, W.C., and Wilson, D.M., III (2011). XRCC1 haploinsufficiency in mice has little effect on aging, but adversely modifies exposure-dependent susceptibility. Nucleic Acids Research *39*, 7992-8004. 10.1093/nar/gkr280.
80. Hanzlikova, H., Gittens, W., Krejcikova, K., Zeng, Z., and Caldecott, K.W. (2016). Overlapping roles for PARP1 and PARP2 in the recruitment of endogenous XRCC1 and PNKP into oxidized chromatin. Nucleic Acids Research *45*, 2546-2557. 10.1093/nar/gkw1246.
81. Besson, A., Dowdy, S.F., and Roberts, J.M. (2008). CDK Inhibitors: Cell Cycle Regulators and Beyond. Developmental Cell *14*, 159-169. https://doi.org/10.1016/j.devcel.2008.01.013.
82. Aubrey, B.J., Strasser, A., and Kelly, G.L. (2016). Tumor-Suppressor Functions of the TP53 Pathway. Cold Spring Harbor perspectives in medicine *6*. 10.1101/cshperspect.a026062.

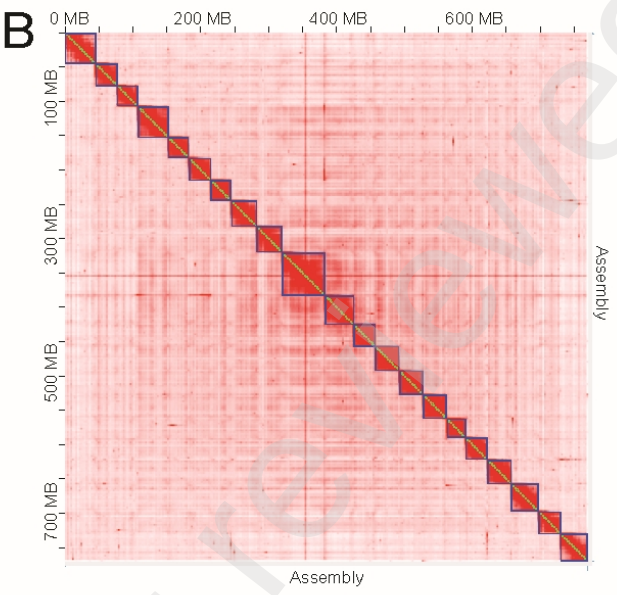
83. Barnum, K.J., and O'Connell, M.J. (2014). Cell cycle regulation by checkpoints. Methods Mol Biol *1170*, 29-40. 10.1007/978-1-4939-0888-2_2.
84. Vélez-Cruz, R., and Johnson, D.G. (2017). The Retinoblastoma (RB) Tumor Suppressor: Pushing Back against Genome Instability on Multiple Fronts. Int J Mol Sci *18*. 10.3390/ijms18081776.
85. Lee, Y.R., Chen, M., and Pandolfi, P.P. (2018). The functions and regulation of the PTEN tumour suppressor: new modes and prospects. Nature reviews. Molecular cell biology *19*, 547-562. 10.1038/s41580-018-0015-0.
86. Gossage, L., Eisen, T., and Maher, E.R. (2015). VHL, the story of a tumour suppressor gene. Nat Rev Cancer *15*, 55-64. 10.1038/nrc3844.
87. Kim, W.Y., and Kaelin, W.G. (2004). Role of VHL Gene Mutation in Human Cancer. Journal of Clinical Oncology *22*, 4991-5004. 10.1200/jco.2004.05.061.
88. Taormina, G., Ferrante, F., Vieni, S., Grassi, N., Russo, A., and Mirisola, M.G. (2019). Longevity: Lesson from Model Organisms. Genes *10*, 518.
89. Feulner, P.G.D., and De-Kayne, R. (2017). Genome evolution, structural rearrangements and speciation. Journal of Evolutionary Biology *30*, 1488-1490. https://doi.org/10.1111/jeb.13101.
90. Zigler, K.S., and Lessios, H.A. (2004). Speciation on the Coasts of The New World: Phylogeography and the Evolution of Bindin in the Sea Urchin Genus Lytechinus. Evolution *58*, 1225-1241. https://doi.org/10.1111/j.0014-3820.2004.tb01702.x.
91. Lee, Y.-H. (2003). Molecular Phylogenies and Divergence Times of Sea Urchin Species of Strongylocentrotidae, Echinoida. Molecular Biology and Evolution *20*, 1211-1221. 10.1093/molbev/msg125.
92. Nesbit, K.T., and Hamdoun, A. (2020). Embryo, larval, and juvenile staging of Lytechinus pictus from fertilization through sexual maturation. Dev Dyn *249*, 1334-1346. 10.1002/dvdy.223.
93. Jeong, D.E., Artan, M., Seo, K., and Lee, S.J. (2012). Regulation of lifespan by chemosensory and thermosensory systems: findings in invertebrates and their implications in mammalian aging. Front Genet *3*, 218. 10.3389/fgene.2012.00218.
94. Alcedo, J., Flatt, T., and Pasyukova, E.G. (2013). Neuronal inputs and outputs of aging and longevity. Front Genet *4*, 71. 10.3389/fgene.2013.00071.
95. Bishop, N.A., Lu, T., and Yankner, B.A. (2010). Neural mechanisms of ageing and cognitive decline. Nature *464*, 529-535. 10.1038/nature08983.
96. Bouchard, J., and Villeda, S.A. (2015). Aging and brain rejuvenation as systemic events. Journal of neurochemistry *132*, 5-19. 10.1111/jnc.12969.
97. Hansen, M., Rubinsztein, D.C., and Walker, D.W. (2018). Autophagy as a promoter of longevity: insights from model organisms. Nature Reviews Molecular Cell Biology *19*, 579-593. 10.1038/s41580-018-0033-y.
98. Simonsen, A., Cumming, R.C., Brech, A., Isakson, P., Schubert, D.R., and Finley, K.D. (2008). Promoting basal levels of autophagy in the nervous system enhances longevity and oxidant resistance in adult Drosophila. Autophagy *4*, 176-184. 10.4161/auto.5269.
99. Fulop, T., Dupuis, G., Baehl, S., Le Page, A., Bourgade, K., Frost, E., Witkowski, J.M., Pawelec, G., Larbi, A., and Cunnane, S. (2016). From inflamm-aging to immune-paralysis: a slippery slope during aging for immune-adaptation. Biogerontology *17*, 147-157. 10.1007/s10522-015-9615-7.

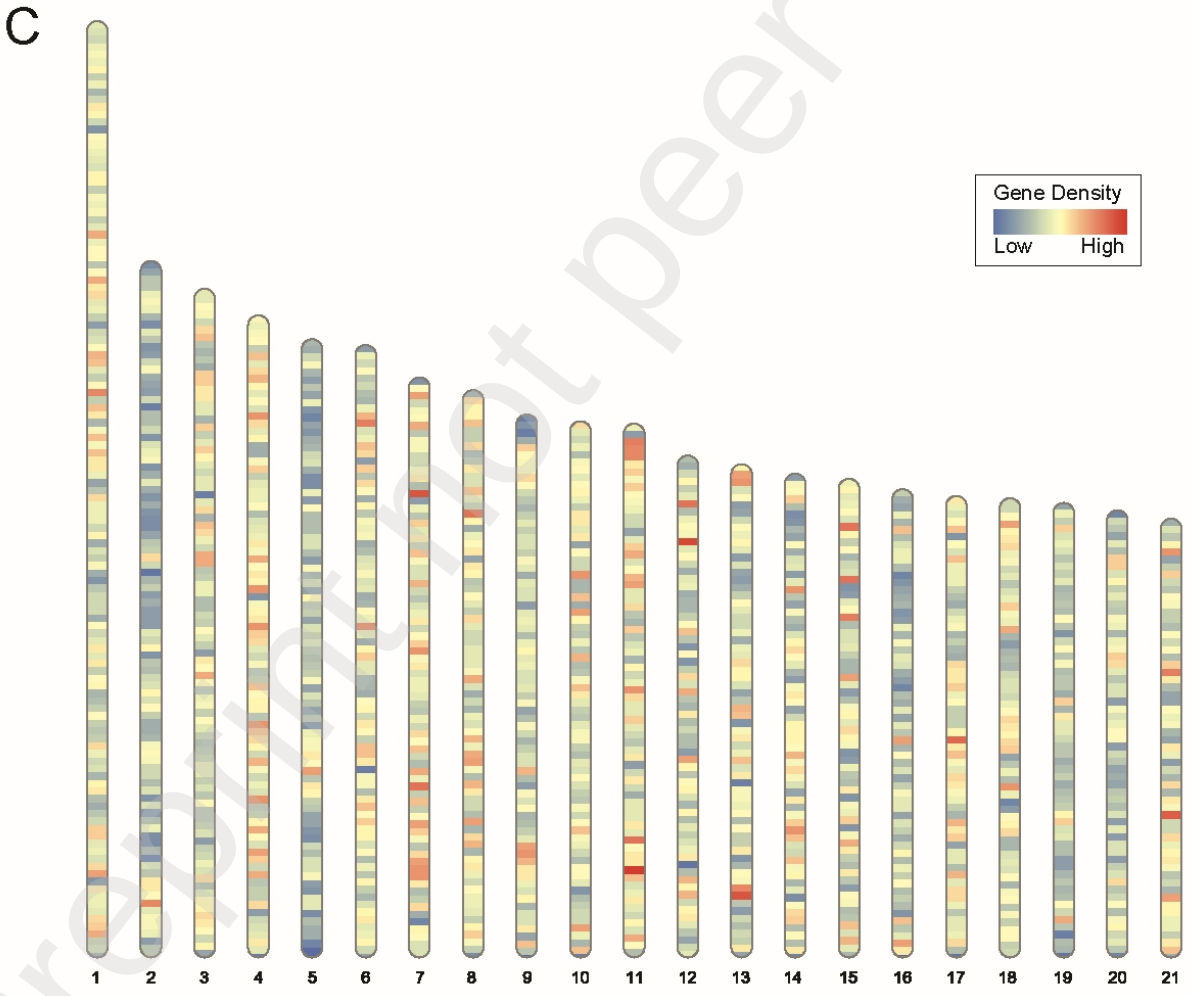
100. Garschall, K., and Flatt, T. (2018). The interplay between immunity and aging in Drosophila. F1000Research *7*, 160. 10.12688/f1000research.13117.1.
101. Yu, L., Wang, L., and Chen, S. (2010). Endogenous toll-like receptor ligands and their biological significance. J Cell Mol Med *14*, 2592-2603. 10.1111/j.1582-4934.2010.01127.x.
102. Lin, C.Y., Oulhen, N., Wessel, G., and Su, Y.H. (2019). CRISPR/Cas9-mediated genome editing in sea urchins. Methods Cell Biol *151*, 305-321. 10.1016/bs.mcb.2018.10.004.
103. Oulhen, N., Pieplow, C., Perillo, M., Gregory, P., and Wessel, G.M. (2022). Optimizing CRISPR/Cas9-based gene manipulation in echinoderms. Dev Biol *490*, 117-124. 10.1016/j.ydbio.2022.07.008.
104. Pieplow, A., Dastaw, M., Sakuma, T., Sakamoto, N., Yamamoto, T., Yajima, M., Oulhen, N., and Wessel, G.M. (2021). CRISPR-Cas9 editing of non-coding genomic loci as a means of controlling gene expression in the sea urchin. Dev Biol *472*, 85-97. 10.1016/j.ydbio.2021.01.003.
105. Oulhen, N., Morita, S., Warner, J.F., and Wessel, G. (2023). CRISPR/Cas9 knockin methodology for the sea urchin embryo. Mol Reprod Dev *90*, 69-72. 10.1002/mrd.23672.
106. Lewis, K.N., Soifer, I., Melamud, E., Roy, M., McIsaac, R.S., Hibbs, M., and Buffenstein, R. (2016). Unraveling the message: insights into comparative genomics of the naked mole-rat. Mammalian genome : official journal of the International Mammalian Genome Society *27*, 259-278. 10.1007/s00335-016-9648-5.
107. Omasits, U., Ahrens, C.H., Müller, S., and Wollscheid, B. (2013). Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics *30*, 884-886. 10.1093/bioinformatics/btt607.
108. Doyle, J.J., and Doyle, L.L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin *19*, 11-15.
109. Stortchevoi, A., Kamelamela, N., and Levine, S.S. (2020). SPRI Beads-based Size Selection in the Range of 2-10kb. Journal of biomolecular techniques : JBT *31*, 7-10. 10.7171/jbt.20-3101-002.
110. De Coster, W., D'Hert, S., Schultz, D.T., Cruts, M., and Van Broeckhoven, C. (2018). NanoPack: visualizing and processing long-read sequencing data. Bioinformatics *34*, 2666-2669. 10.1093/bioinformatics/bty149.
111. Durand, N.C., Shamim, M.S., Machol, I., Rao, S.S.P., Huntley, M.H., Lander, E.S., and Aiden, E.L. (2016). Juicer Provides a One-Click System for Analyzing Loop-Resolution Hi-C Experiments. Cell Systems *3*, 95-98. https://doi.org/10.1016/j.cels.2016.07.002.
112. Kim, D., Paggi, J.M., Park, C., Bennett, C., and Salzberg, S.L. (2019). Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol *37*, 907-915. 10.1038/s41587-019-0201-4.
113. Brůna, T., Hoff, K.J., Lomsadze, A., Stanke, M., and Borodovsky, M. (2021). BRAKER2: automatic eukaryotic genome annotation with GeneMark-EP+ and AUGUSTUS supported by a protein database. NAR Genomics and Bioinformatics *3*. 10.1093/nargab/lqaa108.
114. Caballero, M., and Wegrzyn, J. (2019). gFACs: Gene Filtering, Analysis, and Conversion to Unify Genome Annotations Across Alignment and Gene Prediction Frameworks. Genomics Proteomics Bioinformatics *17*, 305-310. 10.1016/j.gpb.2019.04.002.

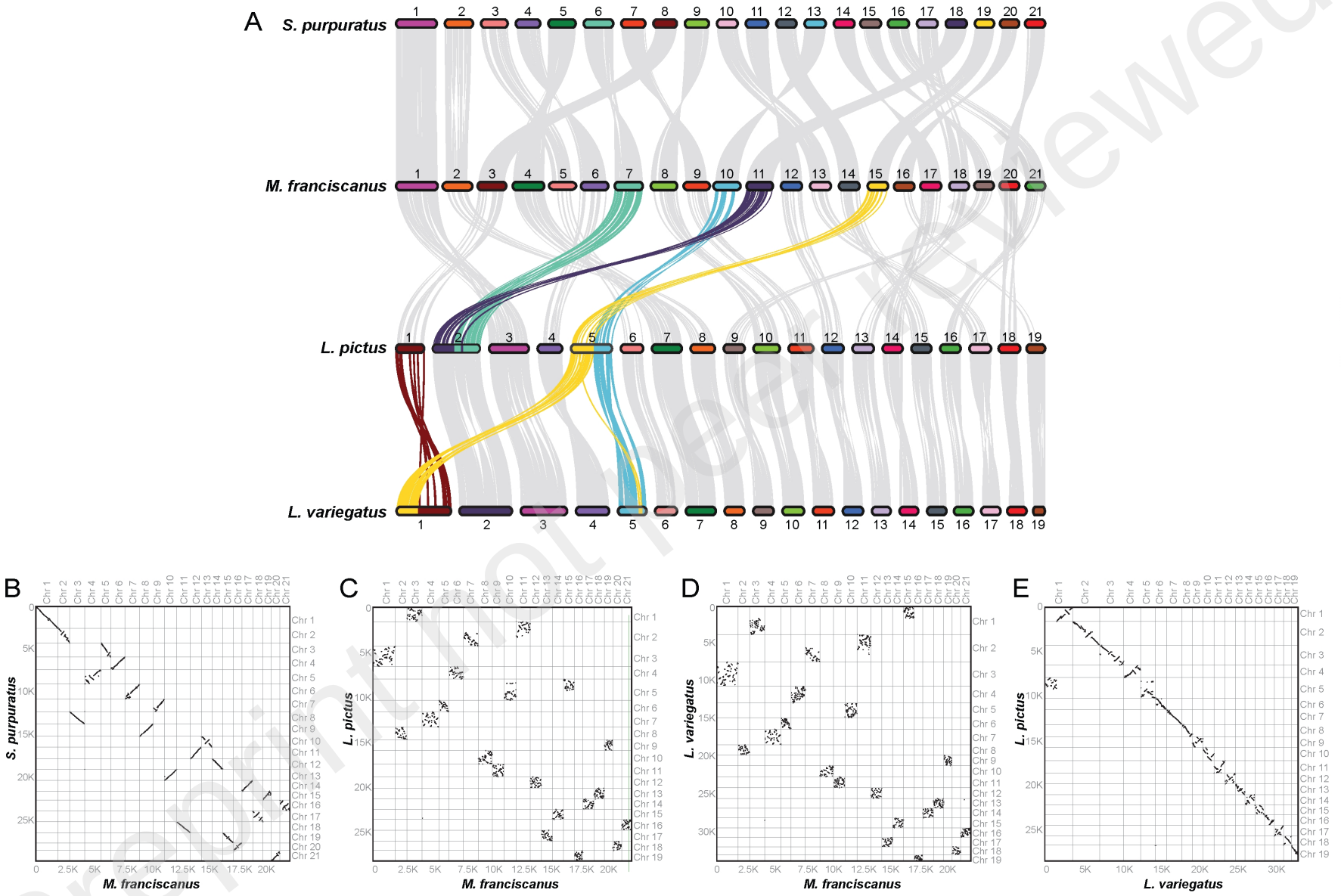
115. Milne, I., Stephen, G., Bayer, M., Cock, P.J., Pritchard, L., Cardle, L., Shaw, P.D., and Marshall, D. (2013). Using Tablet for visual exploration of second-generation sequencing data. Briefings in bioinformatics *14*, 193-202. 10.1093/bib/bbs012.
116. Hart, A.J., Ginzburg, S., Xu, M., Fisher, C.R., Rahmatpour, N., Mitton, J.B., Paul, R., and Wegrzyn, J.L. (2020). EnTAP: Bringing faster and smarter functional annotation to non-model eukaryotic transcriptomes. Molecular Ecology Resources *20*, 591-604. https://doi.org/10.1111/1755-0998.13106.
117. Quinlan, A.R., and Hall, I.M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics *26*, 841-842. 10.1093/bioinformatics/btq033.
118. Hao, Z., Lv, D., Ge, Y., Shi, J., Weijers, D., Yu, G., and Chen, J. (2020). RIdeogram: drawing SVG graphics to visualize and map genome-wide data on the idiograms. PeerJ. Computer science *6*, e251. 10.7717/peerj-cs.251.
119. Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., and Gingeras, T.R. (2012). STAR: ultrafast universal RNA-seq aligner. Bioinformatics *29*, 15-21. 10.1093/bioinformatics/bts635.
120. Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2009). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics *26*, 139-140. 10.1093/bioinformatics/btp616.
121. Arshinoff, B.I., Cary, G.A., Karimi, K., Foley, S., Agalakov, S., Delgado, F., Lotay, V.S., Ku, C.J., Pells, T.J., Beatman, T.R., et al. (2021). Echinobase: leveraging an extant model organism database to build a knowledgebase supporting research on the genomics and biology of echinoderms. Nucleic Acids Research *50*, D970-D979. 10.1093/nar/gkab1005.
122. Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. BMC Bioinformatics *9*, 18. 10.1186/1471-2105-9-18.
123. Edgar, R.C. (2010). Search and clustering orders of magnitude faster than BLAST. Bioinformatics *26*, 2460-2461. 10.1093/bioinformatics/btq461.
124. Manni, M., Berkeley, M.R., Seppey, M., Simão, F.A., and Zdobnov, E.M. (2021). BUSCO Update: Novel and Streamlined Workflows along with Broader and Deeper Phylogenetic Coverage for Scoring of Eukaryotic, Prokaryotic, and Viral Genomes. Molecular Biology and Evolution *38*, 4647-4654. 10.1093/molbev/msab199.
125. Arshinoff, B.I., Cary, G.A., Karimi, K., Foley, S., Agalakov, S., Delgado, F., Lotay, V.S., Ku, C.J., Pells, T.J., Beatman, T.R., et al. (2022). Echinobase: leveraging an extant model organism database to build a knowledgebase supporting research on the genomics and biology of echinoderms. Nucleic Acids Res *50*, D970-d979. 10.1093/nar/gkab1005.
126. Emms, D.M., and Kelly, S. (2019). OrthoFinder: phylogenetic orthology inference for comparative genomics. Genome Biol *20*, 238. 10.1186/s13059-019-1832-y.
127. Huerta-Cepas, J., Serra, F., and Bork, P. (2016). ETE 3: Reconstruction, Analysis, and Visualization of Phylogenomic Data. Mol Biol Evol *33*, 1635-1638. 10.1093/molbev/msw046.
128. Han, M.V., Thomas, G.W., Lugo-Martinez, J., and Hahn, M.W. (2013). Estimating gene gain and loss rates in the presence of error in genome assembly and annotation using CAFE 3. Mol Biol Evol *30*, 1987-1997. 10.1093/molbev/mst100.

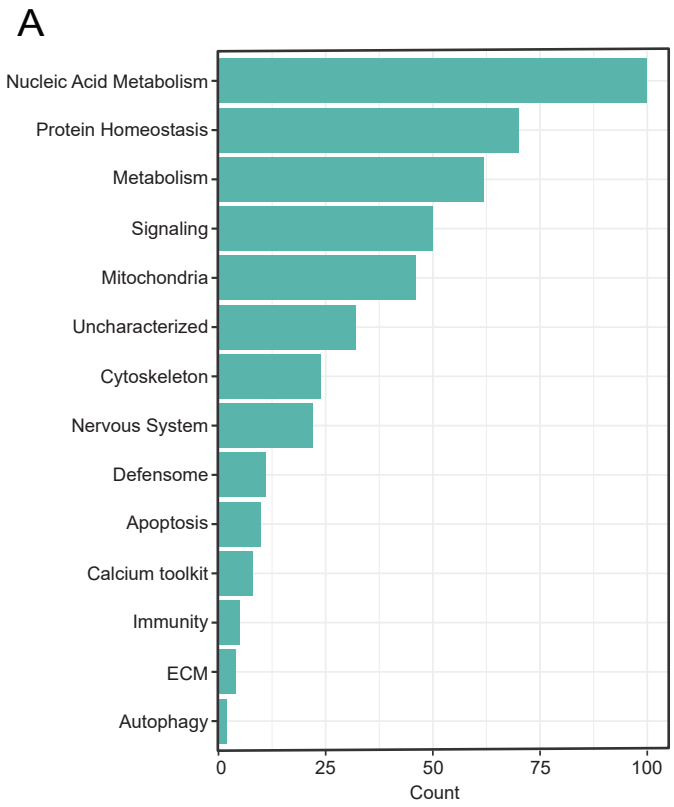
129. Tang, H., Wang, X., Bowers, J.E., Ming, R., Alam, M., and Paterson, A.H. (2008). Unraveling ancient hexaploidy through multiply-aligned angiosperm gene maps. Genome Res *18*, 1944-1954. 10.1101/gr.080978.108.
130. Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. Nucleic Acids Res *34*, W609-612. 10.1093/nar/gkl315.
131. Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. Computer applications in the biosciences : CABIOS *13*, 555-556. 10.1093/bioinformatics/13.5.555.
132. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol Biol Evol *24*, 1586-1591. 10.1093/molbev/msm088.

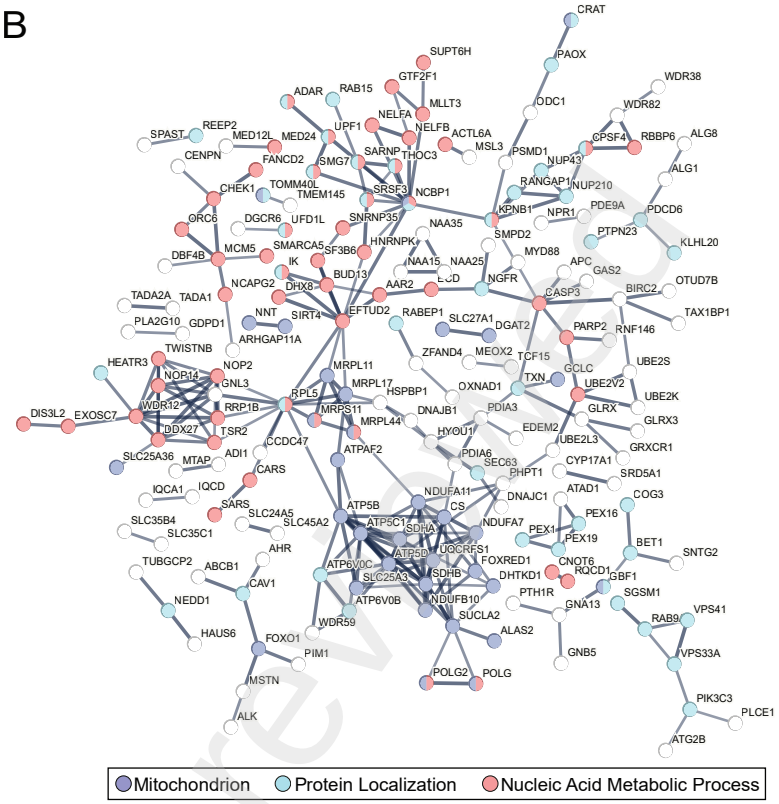


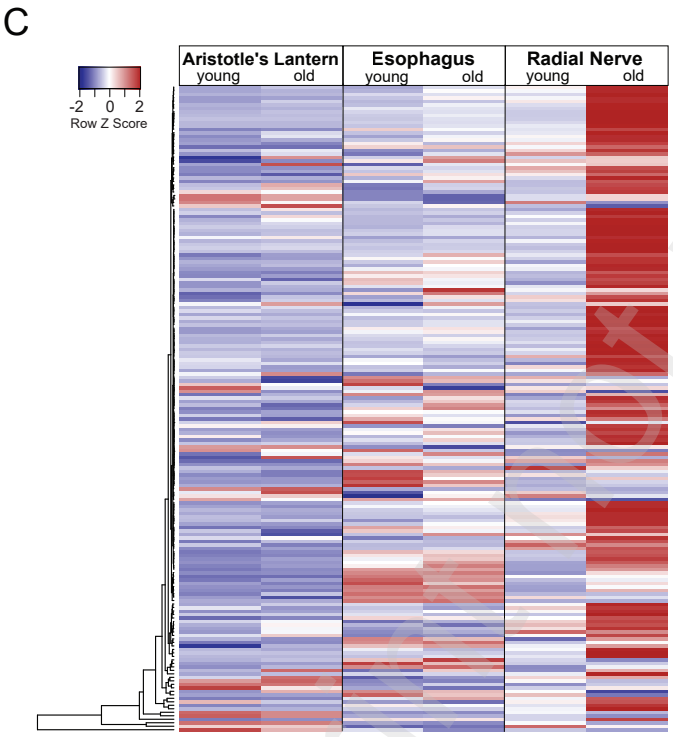


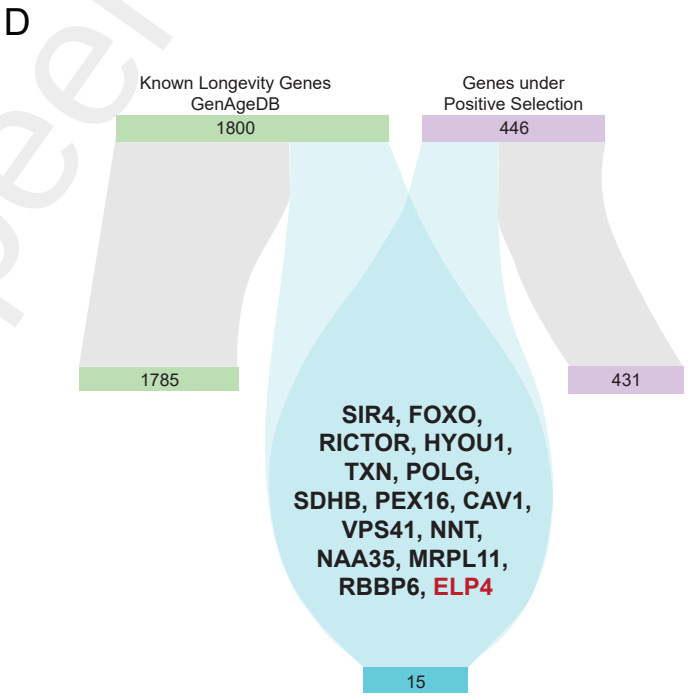




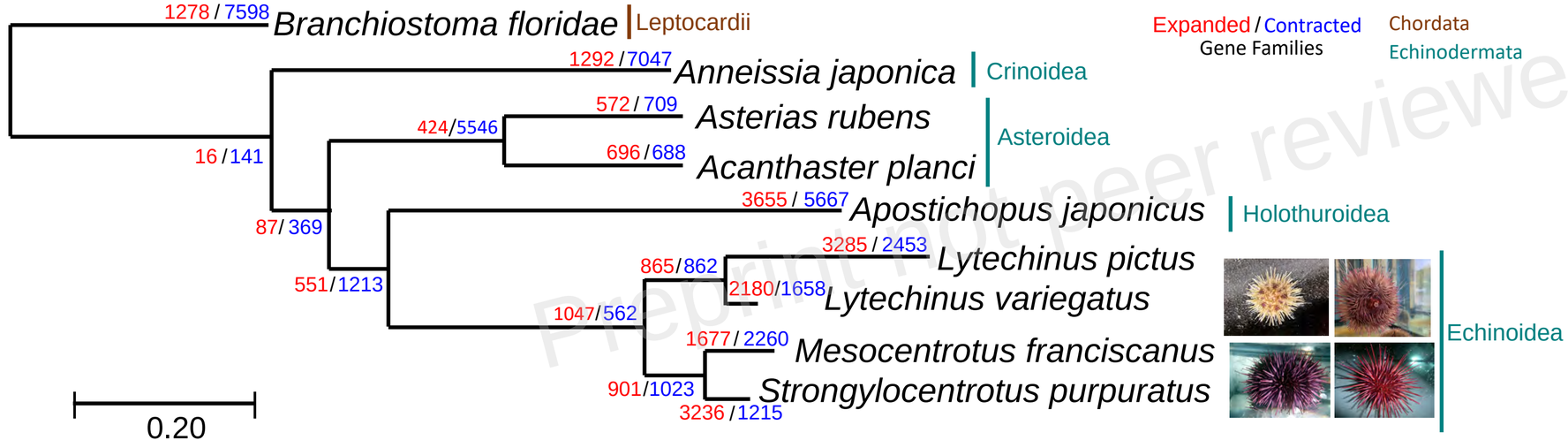




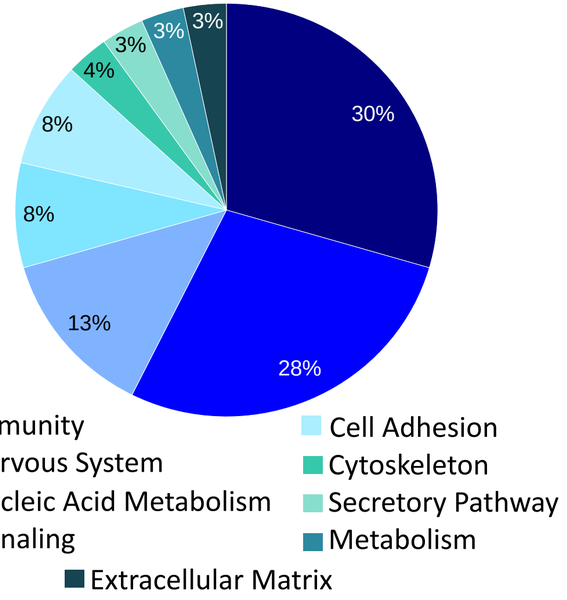


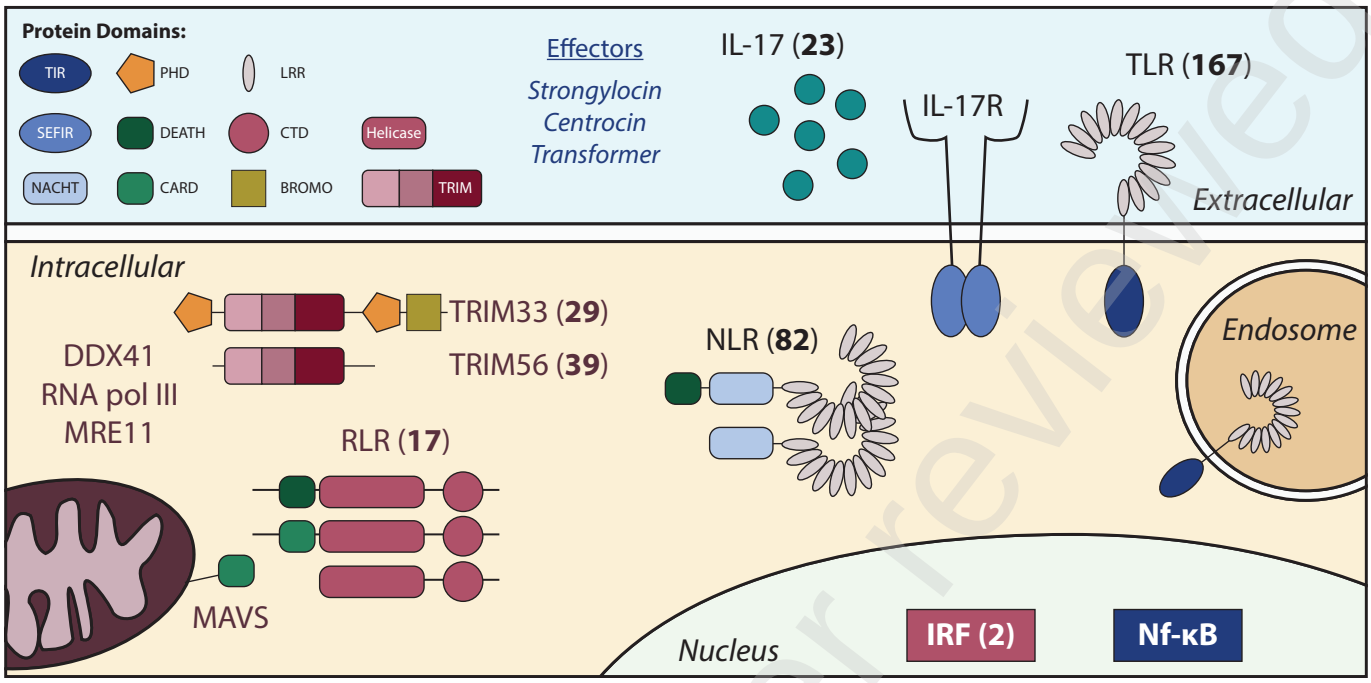


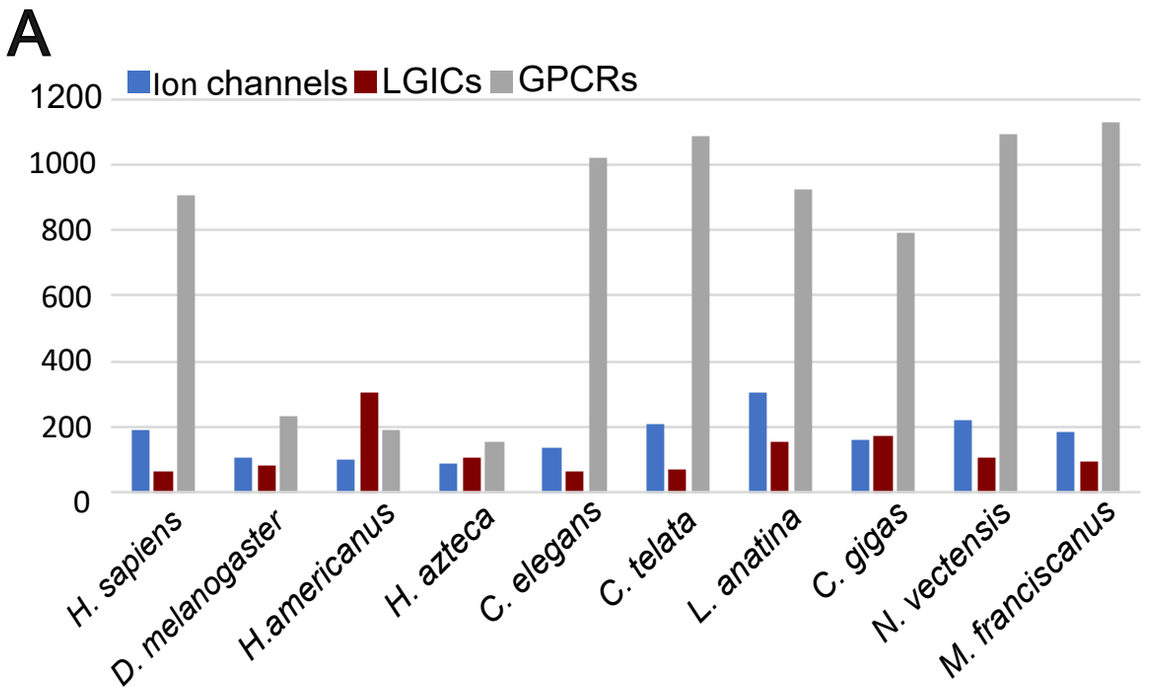
A

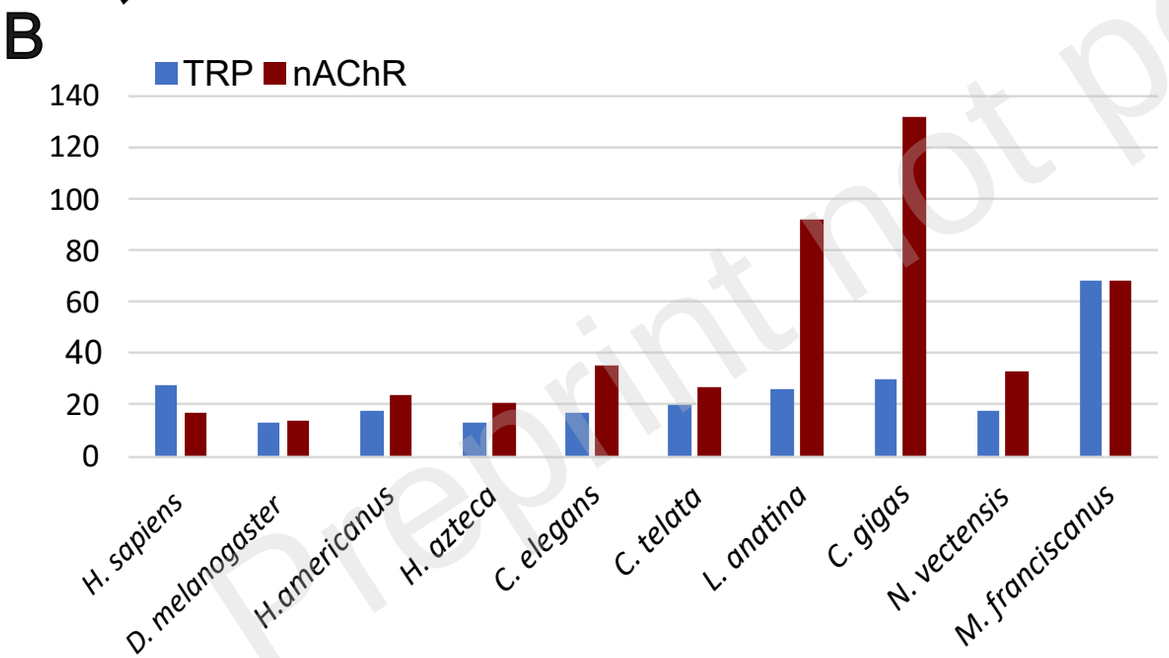


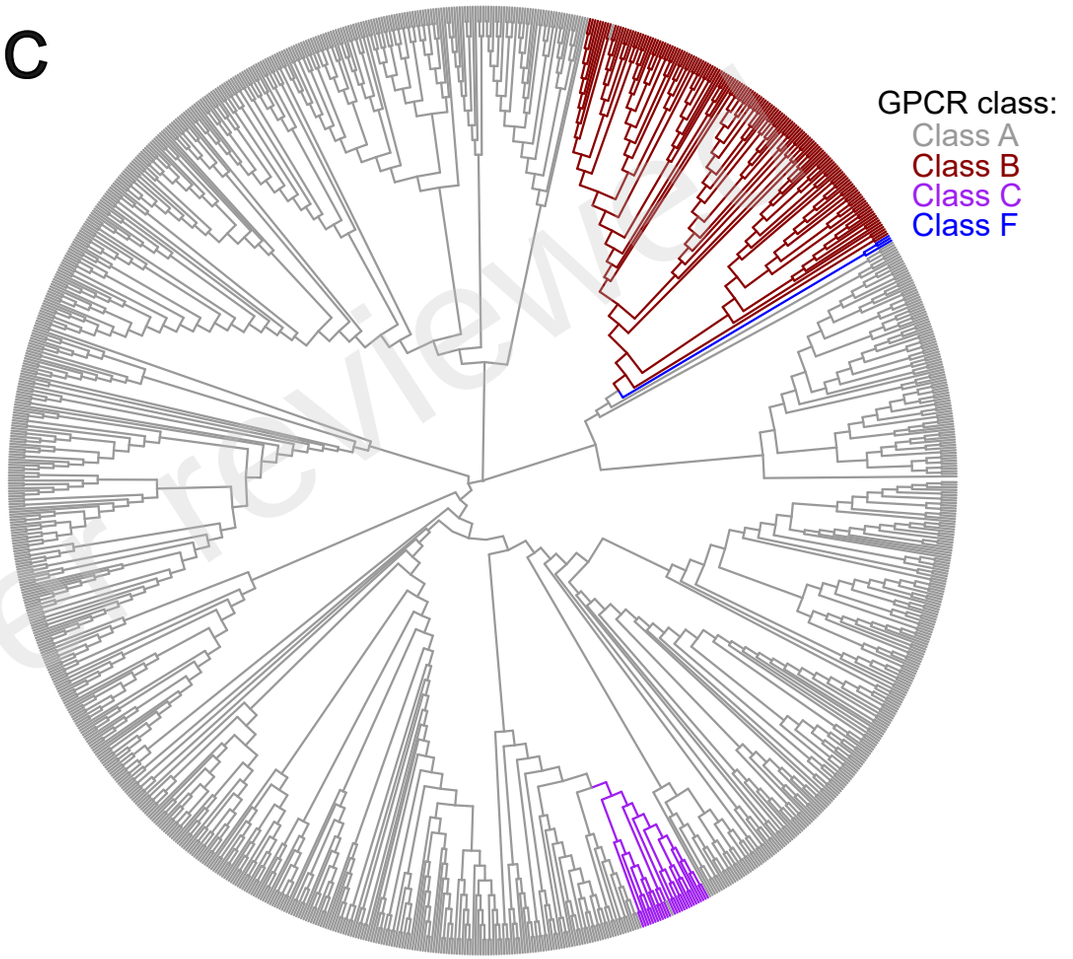
B

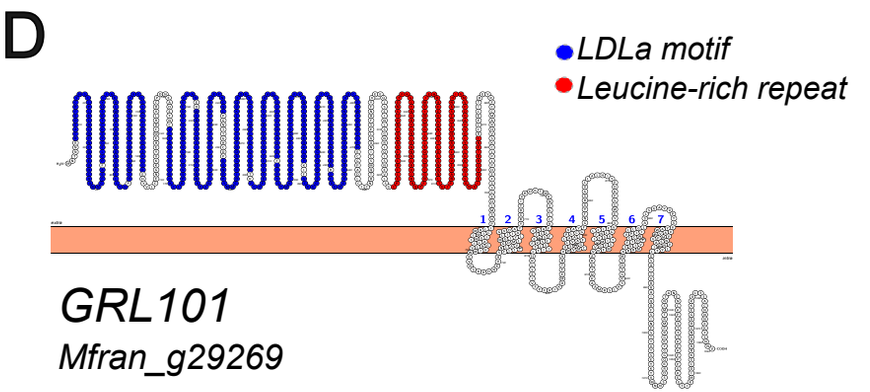


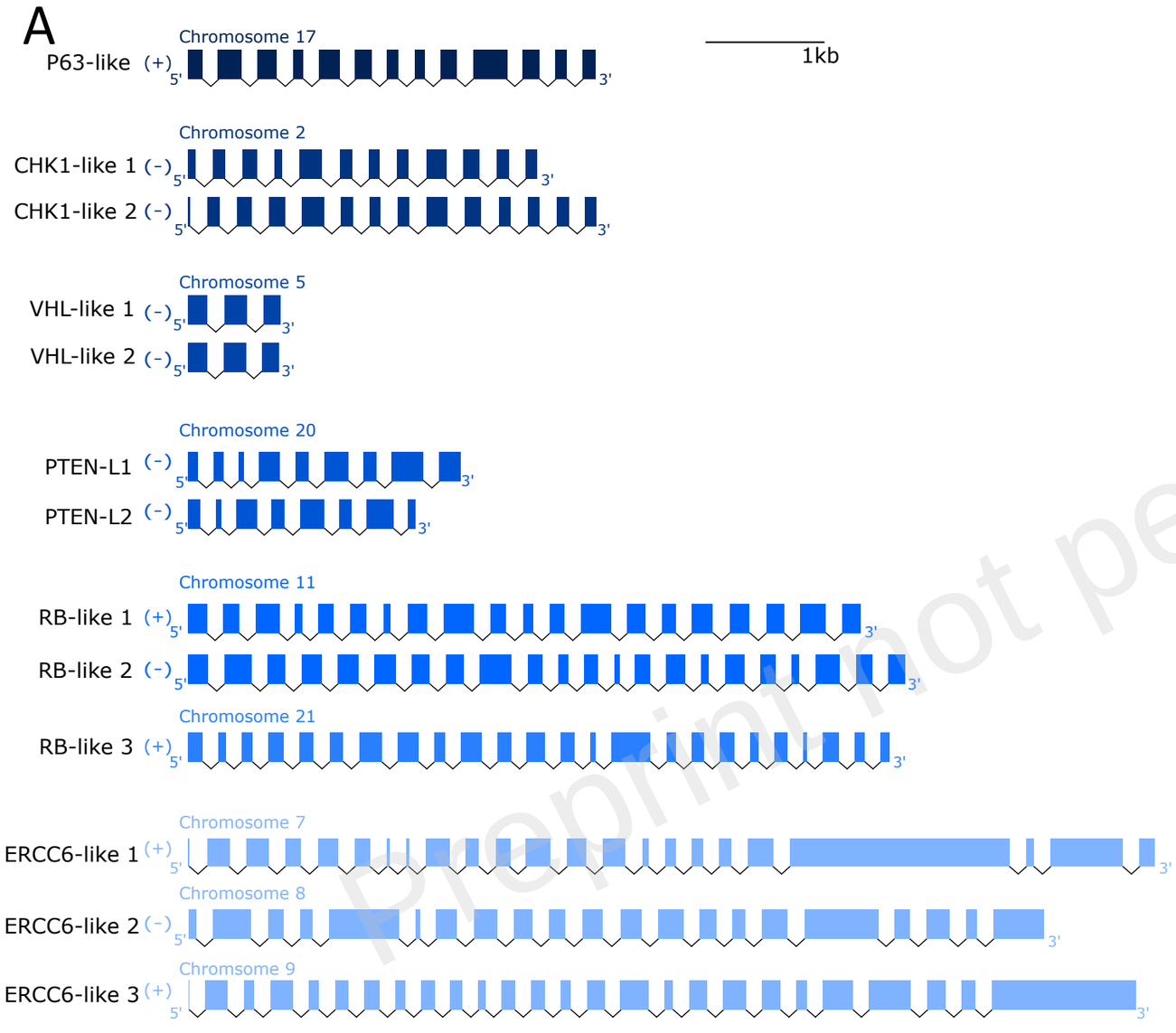


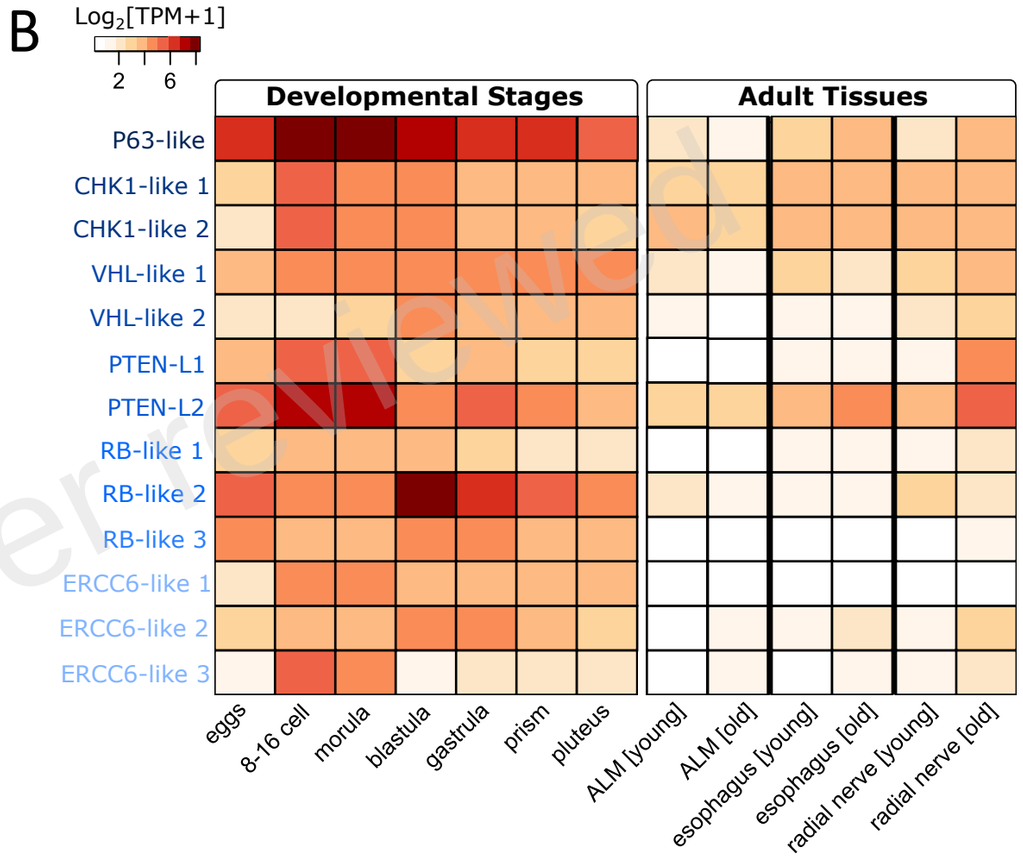












C

	P53/P63	CHK1	VHL	PTEN	RB	ERCC6
MFR	1	2	2	2	3	3
SPU	1	1	1	2	3	4
LVA	1	1	1	2	3	4
LPI	1	2	1	2	3	3

Genomic signatures of exceptional longevity and negligible aging in the long-lived red sea urchin

Jennifer M. Polinski[1], Kate R. Castellano[1], Katherine M. Buckley[2] and Andrea G. Bodnar[1]*

[1]Gloucester Marine Genomics Institute, Gloucester, MA 01930
[2]Department of Biological Sciences, Auburn University, Auburn, AL, 36849
*corresponding author

Supplemental Information

Table S1. Statistics for the *M. franciscanus* genome assembly and gene predictions.

Assembly	
Total length	773.16 Mb
Number of scaffolds	21
N50 (size/number scaffolds)	36.12 Mb / 9
N90 (size/number scaffolds)	30.24 Mb / 19
Number of contigs	4034
Contig N50	359.45 Kb / 656
Number of gaps	4,013
Percent of genome in gaps	0.25%
BUSCO (genome assembly)	metazoa_odb10
Complete, single BUSCO genes	910 (95.4%)
Duplicated BUSCO genes	13 (1.4%)
Fragmented BUSCO genes	23 (2.4%)
Missing BUSCO genes	8 (0.8%)
Gene predictions	
No. genes	22,210
Average gene length	13,423 bp
Average CDS length	1,558 bp
BUSCO (gene predictions)	metazoa_odb10
Complete, single BUSCO genes	887 (93.0%)
Duplicated BUSCO genes	27 (2.8%)
Fragmented BUSCO genes	31 (3.2%)
Missing BUSCO genes	9 (1.0%)

Table S3. Repeat elements in the *M. franciscanus*, *S. purpuratus*, *L. pictus*, and *L. variegatus* genomes.

	M. franciscanus Combined Repeat Library	S. purpuratus Combined Repeat Library	L. pictus Combined Repeat Library	*L. variegatus* Combined Repeat Library
Genome Version	1.0	5 DOI: 10.1126/science.1133609	2 DOI: 10.1093/gbe/evab061	3 DOI: 10.1093/gbe/evaa101
	Percentage of Sequence	Percentage of Sequence	Percentage of Sequence	Percentage of Sequence
Retroelements	25.27	29.48	27.52	26.73
SINEs:	4.67	8.24	5.84	7.86
Penelope	0.42	0.14	0.77	0.42
LINEs:	17.62	15.57	16.41	13.42
CRE/SLACS	0.00	0	0	0
L2/CR1/Rex	11.03	9.03	10.11	8.03
R1/LOA/Jockey	0.34	0.4	0.33	0.29
R2/R4/NeSL	0.10	0.08	0.21	0.17
RTE/Bov-B	1.90	3.05	2.73	2.61
L1/CIN4	0.54	1.46	2.1	2.17
LTR elements	2.98	5.67	5.28	5.44
BEL/Pao	0.33	0.51	0.37	0.36
Ty1/Copia	0.02	0.01	0	0
Gypsy/DIRS1	2.11	2.64	3.18	3.07
Retroviral	0.32	1.33	0.68	0.6
DNA transposons	5.35	6.58	9.01	7.96
hobo-Activator	0.67	1.49	2.81	2.47
Tc1-IS630-Pogo	2.21	0.99	0.74	0.77
En-Spm	0.00	0	0	0
MuDR-IS905	0.00	0	0	0
PiggyBac	0.01	0.01	0.03	0.02
Tourist/Harbinger	0.44	1.12	0.65	0.7
Other (Mirage, P-element, Transib):	0.17	0.05	0.07	0.15
Rolling-circles:	0.95	0.24	0.44	0.24
Unclassified:	28.71	21.79	26.72	24.85
Classified:	30.62	36.05	36.54	34.69
Total interspersed repeats:	59.33	57.84	63.26	59.54
Small RNA:	3.44	7.62	4.57	6.35
Satellites:	0.31	0.93	0.24	0.28
Simple repeats:	1.59	1.51	0.97	1.03
Low complexity:	0.22	0.21	0.18	0.2

Table S5. GO and KEGG pathway enrichment for genes under positive selection in the *M. franciscanus* genome.

Term	Description	Count in network	Strength	False discovery rate
GO Biological Process				
GO:0006099	Tricarboxylic acid cycle	6 of 34	0.94	0.0407
GO:0045333	Cellular respiration	14 of 158	0.64	0.0063
GO:0051168	Nuclear export	13 of 162	0.59	0.0204
GO:0006611	Protein export from nucleus	12 of 150	0.59	0.0343
GO:0015980	Energy derivation by oxidation of organic compounds	17 of 223	0.57	0.0052
GO:0006913	Nucleocytoplasmic transport	18 of 263	0.52	0.0089
GO:0006091	Generation of precursor metabolites and energy	27 of 405	0.51	0.00031
GO:0055114	Oxidation-reduction process	47 of 939	0.39	6.67E-05
GO:0006396	RNA processing	39 of 854	0.35	0.0038
GO:0016071	mRNA metabolic process	30 of 678	0.33	0.0341
GO:0006886	Intracellular protein transport	43 of 999	0.32	0.0045
GO:0045184	Establishment of protein localization	61 of 1564	0.28	0.0015
GO:0046907	Intracellular transport	59 of 1520	0.28	0.0024
GO:0042886	Amide transport	58 of 1553	0.26	0.0063
GO:0015031	Protein transport	56 of 1486	0.26	0.0065
GO:0034613	Cellular protein localization	57 of 1610	0.24	0.02
GO:0016070	RNA metabolic process	56 of 1584	0.24	0.0217
GO:0071705	Nitrogen compound transport	64 of 1823	0.23	0.0101
GO:0006139	Nucleobase-containing compound metabolic process	90 of 2659	0.22	0.0014
GO:0034641	Cellular nitrogen compound metabolic process	110 of 3282	0.21	0.00014
GO:0046483	Heterocycle metabolic process	95 of 2840	0.21	0.0013
GO:0006725	Cellular aromatic compound metabolic process	96 of 2882	0.21	0.0013
GO:0090304	Nucleic acid metabolic process	72 of 2178	0.21	0.0178
GO:0008104	Protein localization	71 of 2139	0.21	0.0178
GO:0010467	Gene expression	68 of 2056	0.21	0.0236
GO:0044249	Cellular biosynthetic process	85 of 2611	0.2	0.0069
GO:0071702	Organic substance transport	70 of 2173	0.2	0.036
GO:1901360	Organic cyclic compound metabolic process	99 of 3118	0.19	0.004
GO:1901576	Organic substance biosynthetic process	86 of 2734	0.19	0.0168
GO:0033036	Macromolecule localization	79 of 2473	0.19	0.02
GO:0009058	Biosynthetic process	87 of 2788	0.18	0.018
GO:0044237	Cellular metabolic process	229 of 7513	0.17	3.24E-10
GO:0008152	Metabolic process	247 of 8298	0.16	2.07E-10
GO:0006807	Nitrogen compound metabolic process	202 of 6852	0.16	1.04E-06
GO:0044238	Primary metabolic process	213 of 7332	0.15	9.49E-07
GO:0071704	Organic substance metabolic process	220 of 7755	0.14	2.01E-06
GO:0043170	Macromolecule metabolic process	170 of 6137	0.13	0.0024
GO:0044260	Cellular macromolecule metabolic process	139 of 4976	0.13	0.0138
GO:1901564	Organonitrogen compound metabolic process	143 of 5244	0.12	0.0239
GO:0009987	Cellular process	355 of 15024	0.06	4.40E-06
GO Molecular Function				
GO:0015036	Disulfide oxidoreductase activity	7 of 41	0.92	0.0488
GO:0016667	Oxidoreductase activity, acting on a sulfur group of donors	8 of 59	0.82	0.0488
GO:0016491	Oxidoreductase activity	33 of 726	0.35	0.0469
GO:0003824	Catalytic activity	170 of 5486	0.18	2.61E-06
GO Cellular Component				
GO:0000275	Mitochondrial proton-transporting ATP synthase complex, catalytic sector F(1)	3 of 5	1.47	0.0355
GO:0005746	Mitochondrial respirasome	9 of 84	0.72	0.0106
GO:0098800	Inner mitochondrial membrane protein complex	10 of 131	0.57	0.0441
GO:0098798	Mitochondrial protein complex	19 of 262	0.55	0.00085
GO:0005759	Mitochondrial matrix	25 of 479	0.41	0.0045
GO:0005743	Mitochondrial inner membrane	25 of 480	0.41	0.0045
GO:0019866	Organelle inner membrane	26 of 541	0.37	0.0094
GO:0005739	Mitochondrion	61 of 1611	0.27	0.00078
GO:1902494	Catalytic complex	46 of 1328	0.23	0.037
GO:0032991	Protein-containing complex	149 of 5073	0.16	0.00016
GO:0031090	Organelle membrane	104 of 3548	0.16	0.0091
GO:0005654	Nucleoplasm	113 of 3973	0.14	0.0114
GO:0070013	Intracellular organelle lumen	161 of 5857	0.13	0.0018
GO:0031981	Nuclear lumen	126 of 4733	0.11	0.0486
GO:0043231	Intracellular membrane-bounded organelle	279 of 10761	0.1	1.98E-06
GO:0043226	Organelle	338 of 13515	0.09	4.98E-09
GO:0043227	Membrane-bounded organelle	310 of 12427	0.09	1.98E-06
GO:0043229	Intracellular organelle	309 of 12528	0.08	5.91E-06
GO:0005737	Cytoplasm	283 of 11428	0.08	0.00011
GO:0005622	Intracellular	342 of 14276	0.07	1.98E-06
KEGG Pathways				
hsa00190	Oxidative phosphorylation	11 of 130	0.62	0.0448
hsa05014	Amyotrophic lateral sclerosis	19 of 352	0.42	0.0448

Table S6. *M. franciscanus* genes under positive selection that are known longevity genes in other organisms.

Gene	Pro- or Anti-Longevity	Description	Source
RICTOR	Pro	Rapamycin-insensitive companion of mTOR, subunit of mTORC2	GenAge
FOXO1	Pro	Forkhead transcription factor that regulates metabolic homeostasis and stress resistance	GenAge, Giannakou et al., 2004[1]
PEX16	Pro	Peroxisomal membrane protein peroxin 16 required for peroxisome biogenesis	GenAge
TXN	Pro	Thioredoxin, involved in many redox reactions and acts as an antioxidant	GenAge, Sahm et al., 2018[2]
VPS41	Pro	Vacuolar protein sorting-associated protein 41-vesicle-mediated protein trafficking	GenAge
CAV1	Pro	Caveolin, scaffold protein linking integrins to the Ras-ERK pathway promoting cell cycle progression	GenAge
SIRT4	Pro	Sirtuin-4, mitochondrial NAD-dependent protein lipoamidase, deacetylase and ADP-ribosyl transferase, regulates metabolism and lipid homeostasis	GenAge Parik et al., 2019[3]
POLG	Pro	Mitochondrial DNA polymerase subunit gamma-1	GenAge, Yu et al., 2022[4]
SDHB	Pro	Iron-sulfur protein subunit of the succinate dehydrogenase complex II of the mitochondrial respiratory chain	GenAge
NNT	Pro	Mitochondrial NAD(P) transhydrogenase	GenAge
MRPL11	Pro	Mitochondrial 39S ribosomal protein L11	GenAge, Phua et al., 2023[5]
NAA35	Pro	N-alpha-acetyltransferase 35, NatC auxiliary subunit involved in regulation of apoptosis	GenAge, Varland et al., 2023[6]
HYOU1	Pro	Hypoxia upregulated 1 – ER chaperone with important cytoprotective role under stress	Treaster et al., 2023[7]
RBBP6	Pro	E3 ubiquitin-protein ligase regulates DNA-replication, chromosome stability and apoptosis	Treaster et al. 2023[7]
ELP4	Anti	Elongator complex protein 4 of the histone acetyltransferase complex that associates directly with RNA polymerase II during transcriptional elongation	GenAge, Delaney et al., 2011[8]

Table S10. Number of ion channels and GPCRs encoded in the genome of the red sea urchin (*M. franciscanus*) and other metazoans.

	Potassium channel		Calcium channel		Sodium channel		Ion channels										Ligand-gated ion channels					
	K_v	K^+	Ca_v	Ca-2P	Na_v	NALCN	HCN	CNG	TRP	CatSper	H^+	ORAI	ENaC	CIC	Aquaporin	Anoctamin	P2XR	iGluR/IR	nAChR	GABA	GLR	GPCR
H. sapiens	40	38	10	2	10	1	7	6	28	4	1	3	8	9	13	10	7	18	17	19	5	907
D. melanogaster	17	9	3	0	2	1	1	4	13	0	0	1	26	3	7	21	0	61	14	4	3	234
H.americanus	13	27	3	2	1	1	1	3	18	1	2	1	16	3	4	6	1	264	24	5	11	188
H. azteca	11	24	3	1	2	1	1	4	13	0	0	1	15	2	6	6	0	54	21	11	20	155
C. elegans	16	53	3	0	0	2	0	6	17	0	0	1	25	6	8	2	0	10	35	10	7	1024
C. telata	14	33	8	2	2	1	1	4	20	0	2	1	83	8	22	8	1	24	27	8	12	1088
L. anatina	46	70	10	3	7	1	2	5	26	4	1	3	67	7	28	24	5	14	92	9	32	927
C. gigas	36	27	5	3	3	1	1	5	30	0	2	1	22	3	12	11	2	28	132	5	5	794
N. vectensis	50	55	7	3	5	1	2	4	18	7	2	1	35	6	8	17	3	15	33	42	16	1094
M. franciscanus	15	28	13	3	2	1	3	8	68	6	1	1	14	2	13	4	5	14	68	2	3	1132

Figure S1. Percent of repetitive elements in the *M. franciscanus*, *S. purpuratus*, *L. pictus* and *L. variegatus* genomes.

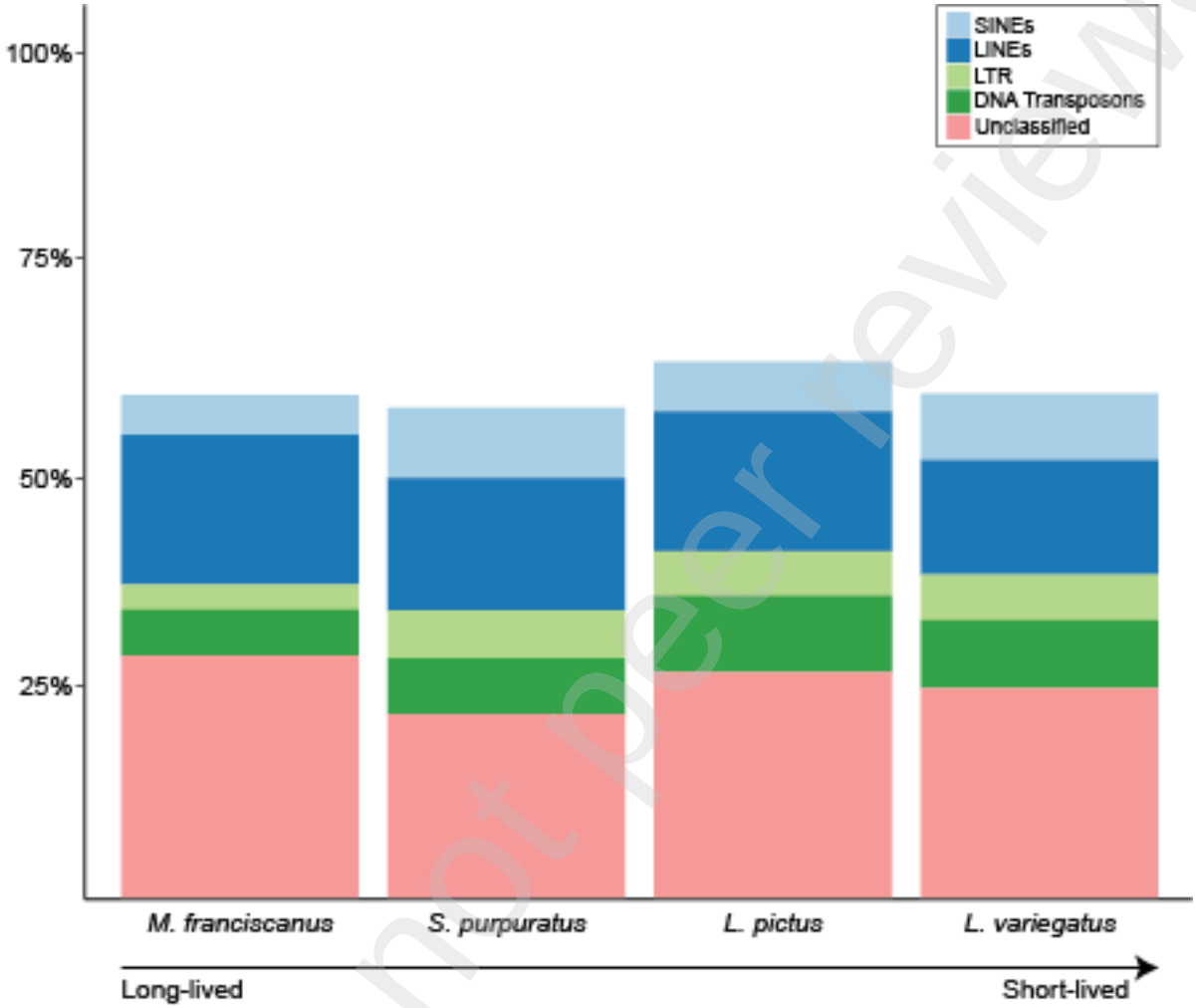


Figure S2. Syntenic alignments between the *M. franciscanus* and *L. variegatus* genomes (A) and the Hox gene cluster (B).

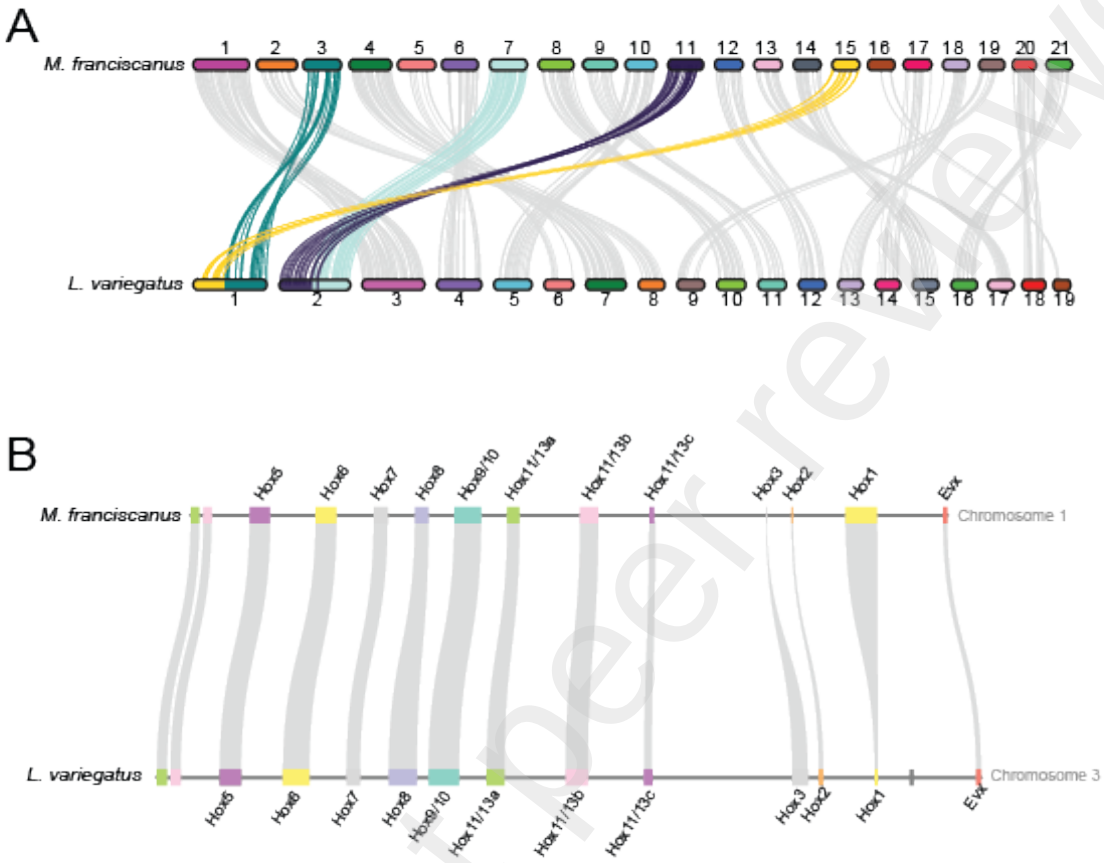


Figure S3. STRING protein-protein interaction network for genes under positive selection in the *M. franciscanus* genome using gene identifiers for the closest *S. purpuratus* homolog as input and selecting *S. purpuratus* as the target organism. High confidence setting (0.7) was used in STRING-DB and only nodes with one or more interactions are shown. Nodes assigned to the selected GO terms are colored as indicated in the legend.

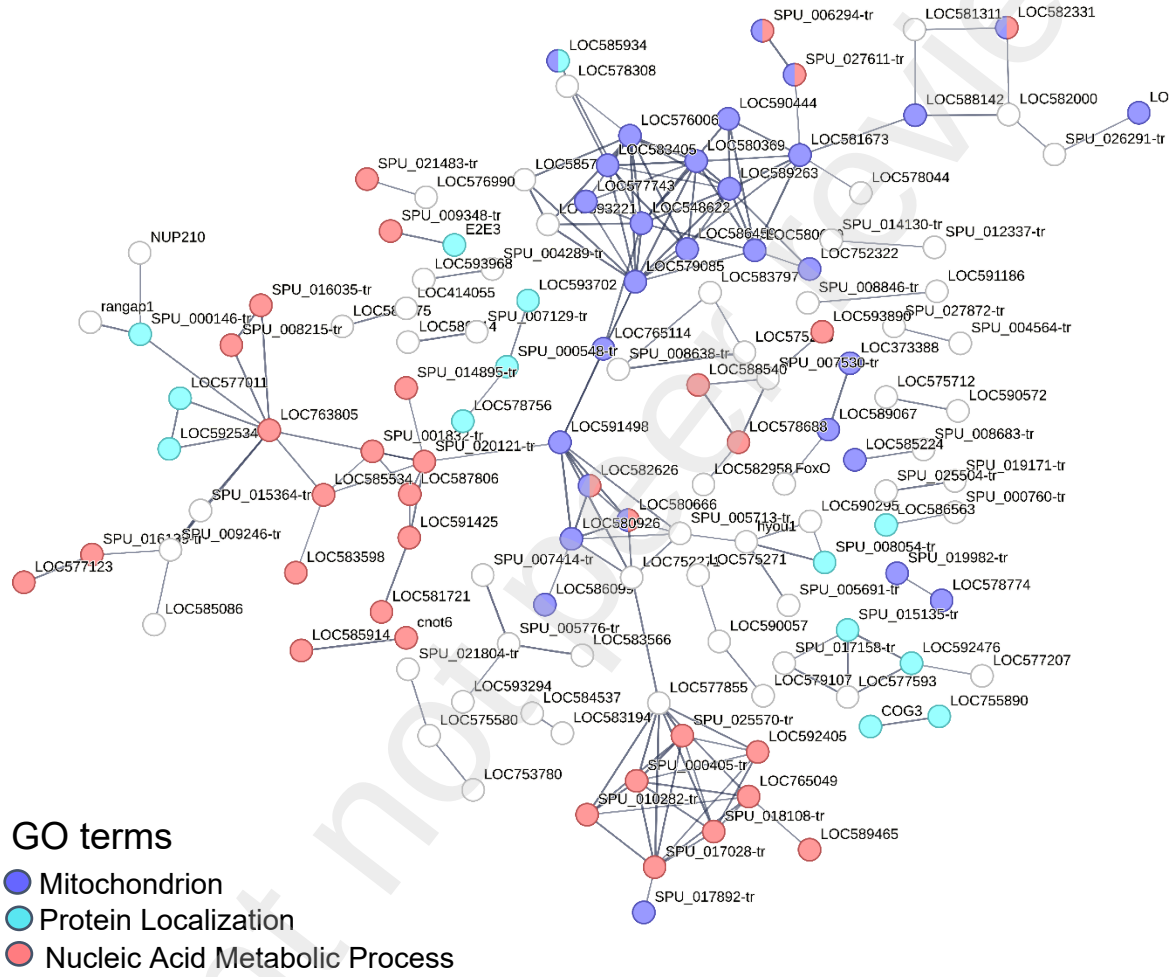
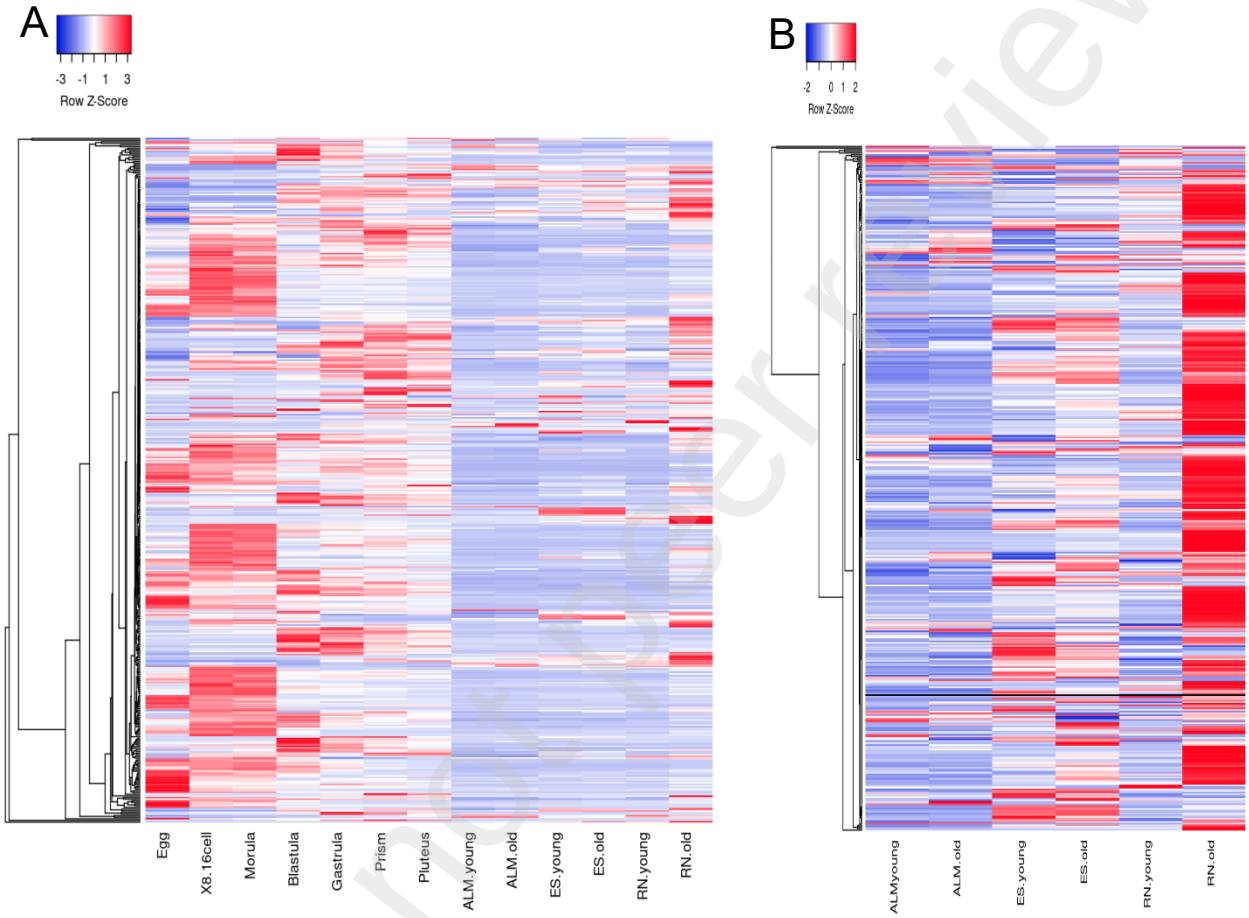


Figure S4. Gene expression patterns of the positively selected genes in the *M. franciscanus* genome. (A) Expression of positively selected genes throughout early development and in adult tissues [young and old muscle (ALM), esophagus (ES) and radial nerve cord (RN)]. Panel B: Expression of positively selected genes in young and old muscle (ALM), esophagus (ES) and radial nerve cord (RN).



Supplemental References

1. Giannakou, M.E., Goss, M., Jünger, M.A., Hafen, E., Leevers, S.J., and Partridge, L. (2004). Long-lived Drosophila with overexpressed dFOXO in adult fat body. Science *305*, 361. 10.1126/science.1098219.
2. Sahm, A., Bens, M., Szafranski, K., Holtze, S., Groth, M., Görlach, M., Calkhoven, C., Müller, C., Schwab, M., Kraus, J., et al. (2018). Long-lived rodents reveal signatures of positive selection in genes associated with lifespan. PLOS Genetics *14*, e1007272. 10.1371/journal.pgen.1007272.
3. Parik, S., Tewary, S., Ayyub, C., and Kolthur-Seetharam, U. (2018). Loss of mitochondrial SIRT4 shortens lifespan and leads to a decline in physical activity. Journal of Biosciences *43*, 243-247. 10.1007/s12038-018-9754-5.
4. Yu, T., Slone, J., Liu, W., Barnes, R., Opresko, P.L., Wark, L., Mai, S., Horvath, S., and Huang, T. (2022). Premature aging is associated with higher levels of 8-oxoguanine and increased DNA damage in the Polg mutator mouse. Aging Cell *21*, e13669. https://doi.org/10.1111/acel.13669.
5. Phua, C.Z.J., Zhao, X., Turcios-Hernandez, L., McKernan, M., Abyadeh, M., Ma, S., Promislow, D., Kaeberlein, M., and Kaya, A. (2023). Genetic perturbation of mitochondrial function reveals functional role for specific mitonuclear genes, metabolites, and pathways that regulate lifespan. GeroScience. 10.1007/s11357-023-00796-4.
6. Varland, S., Silva, R.D., Kjosås, I., Faustino, A., Bogaert, A., Billmann, M., Boukhatmi, H., Kellen, B., Costanzo, M., Drazic, A., et al. (2022). N-terminal acetylation shields proteins from degradation and promotes age-dependent motility and longevity. bioRxiv.
7. Treaster, S., Deelen, J., Daane, J.M., Murabito, J., Karasik, D., and Harris, M.P. (2023). Convergent genomics of longevity in rockfishes highlights the genetics of human life span variation. Science advances *9*, eadd2743. 10.1126/sciadv.add2743.
8. Delaney, J.R., Murakami, C.J., Olsen, B., Kennedy, B.K., and Kaeberlein, M. (2011). Quantitative evidence for early life fitness defects from 32 longevity-associated alleles in yeast. Cell Cycle *10*, 156-165. 10.4161/cc.10.1.14457.