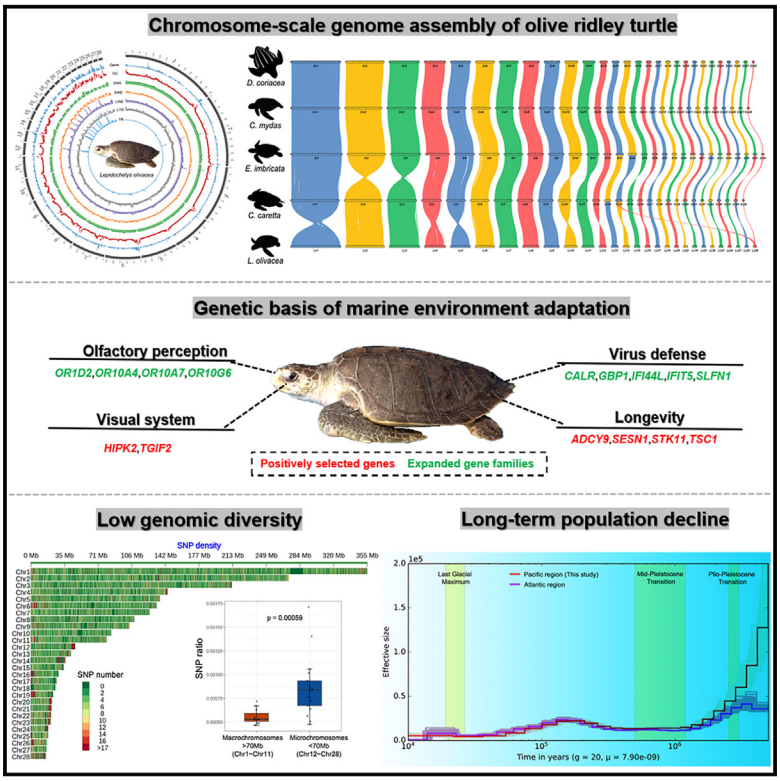
iScience

Article

Genomic insights into marine environment adaptation and conservation of the threatened olive ridley turtle (*Lepidochelys olivacea*)

Graphical abstract



Highlights

- Chromosome-scale genome assembly of olive ridley turtle
- Genetic footprints of marine environment adaptation
- Adaptive changes in genes linked to olfaction, vision, virus defense, and longevity
- Low genomic diversity and long-term population decline of olive ridley turtle

Authors

Lin Yang, Yiting Chen, Shaofeng Wang, ..., Xin Du, Wenliang Zhou, Fuwen Wei

Correspondence

zhouwl@gmlab.ac.cn (W.Z.), weifw@ioz.ac.cn (F.W.)

In brief

Evolutionary biology; Genomics; Nature conservation

Yang et al., 2025, iScience *28*, 111776
February 21, 2025
https://doi.org/10.1016/j.isci.2025.111776

CellPress

iScience

CellPress
OPEN ACCESS

Article

Genomic insights into marine environment adaptation and conservation of the threatened olive ridley turtle (*Lepidochelys olivacea*)

Lin Yang,[1,2,7] Yiting Chen,[1,3,7] Shaofeng Wang,[4] Chenglong Zhang,[5] Xin Huang,[6] Xin Du,[6] Wenliang Zhou,[1,8,*] and Fuwen Wei[1,2,6,*]

[1]Center for Evolution and Conservation Biology, Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China
[2]Jiangxi Key Laboratory of Conservation Biology, College of Forestry, Jiangxi Agricultural University, Nanchang, Jiangxi 330045, China
[3]Department of Ocean Science, The Hong Kong University of Science and Technology, Hong Kong, China
[4]Guangdong Huidong Sea Turtle National Nature Reserve, Huizhou, China
[5]Marine Protected Area Administration of Sansha City, Hainan, China
[6]CAS Key Laboratory of Animal Ecology and Conservation Biology, Institute of Zoology, Chinese Academy of Sciences, Beijing 100101, China
[7]These authors contributed equally
[8]Lead contact
*Correspondence: zhouwl@gmlab.ac.cn (W.Z.), weifw@ioz.ac.cn (F.W.)
https://doi.org/10.1016/j.isci.2025.111776

SUMMARY

Sea turtles are marine flagship species and most of them are currently in a threatened state. Long-term surviving in the ocean has driven significant morphological and physiological changes for this group, which makes them an ideal model for studying adaptive evolution of marine environments. Herein, we present a chromosome-scale genome of *Lepidochelys olivacea* with a genome size of 2.22 Gb and a contig N50 of 97.3 Mb. Comparative genomic analyses uncovered a suite of adaptive changes in genes related to olfaction, vision, virus defense, and longevity, which may help explain the genetic underpinnings of its marine environment adaptation. We also observed that the genome-wide heterozygosity of *L. olivacea* was low (6.45e-4), consistent with its prolonged population decline. Overall, our study provides valuable genetic resources for understanding evolutionary adaptations to aquatic environment and for the conservation of this threatened species.

INTRODUCTION

Sea turtles are the flagship species of marine ecosystems for their crucial roles in coral reef/seagrass protection, energy cycling, and maintaining ecological stability.[1] They also were referred to as 'living fossils of the ocean' with a long evolutionary history beyond 100 million years.[2] The ecological shift from land to the ocean drives a series of adaptive changes in sea turtles. For example, in terms of morphology, their carapace has evolved into a streamlined shape, and their limbs have become paddle-shaped similar to fish,[3] which is beneficial for their swimming. In addition, due to spending almost all their lives in the ocean, except for female adults coming onshore to lay eggs, their senses such as smell, vision, hearing, and salinity perception have perfectly adapted to the marine environment.[4,5] Therefore, sea turtles are a good model for studying adaptive evolution to marine environments.

The olive ridley turtle (*Lepidochelys olivacea*) is one of seven sea turtles globally that belongs to Cheloniidae family, known for its olive-colored carapace during adulthood.[6–8] It was mainly inhabited in warm and tropical waters and primarily distributed in the Pacific, Indian and Atlantic oceans.[8] Compared to other sea turtles, *L. olivacea* has a smaller body with adult weight ranging from 25 to 46 kg, and displays variable and asymmetrical scute counts, ranging from five to nine plates on each side, with six to eight being most commonly observed.[9] *L. olivacea* is chiefly carnivorous and its prey items mainly include starfish, sea urchins, sea squirts, bivalves, shrimp, and crabs.[10] Remarkably, *L. olivacea* also exhibits a striking reproductive behavior known as ''arribadas'', whereby thousands of turtles converge on specific beaches (e.g., Ostional Beach of Costa Rica; Playa Escobilla of Mexico; and Odisha of India) to lay millions of eggs,[11–13] simultaneously. However, the population of *L. olivacea* was greatly reduced from historical estimates due to egg harvesting, fisheries bycatch, and pollution.[8,14] *L. olivacea* was listed as 'Vulnerable' by the IUCN early in 2008.[8] Until now, there has been no accurate estimation of the population size for this threatened species globally, except for some famous rookeries such as Costa Rica, Brazil, Mexico, and Sri Lanka.[15–17]

For decades, genomic data has played an essential role in uncovering threatened species' evolutionary adaptations, population histories, and endangered mechanisms, and further

Table 1. Genome assembly, annotation and genomic diversity statistics of *L. olivacea*

Statistic	Value
Genome assembly	
Genome size (Gb)	2.22
Contig N50 (Mb)	97.3
Scaffolds N50 (Mb)	136
Number of Contigs	190
Number of scaffolds	71
Hi-C anchor percentage (%)	98.63
Chromosome number	28
Genome annotation	
Repeat content (%)	56.59
Number of predicted protein-coding gene	18,273
Functional annotated gene (%)	97.38
BUSCO results for protein-coding gene	C:98.6% [S:97.6%,D:1.0%], F:0.0%,M:1.4%,n:7480
GC content (%)	44.17
Genomic variation	
Number of whole-genome heterozygotes (SNP)	1,432,715
Genome-wide heterozygosity	0.000644683
Repeat-region heterozygosity	0.000792482
Non-exon-region heterozygosity	0.000648327
Exon-region heterozygosity	0.000400618

applying to species conservation.[18–20] In this study, we reported the first high-quality chromosome-scale genome of *L. olivacea*. Based on this genome assembly, we investigated its potential genetic footprint of marine environment adaptation, genomic diversity, and demographic history. Altogether, this study provides valuable genomic resources for studying the genomic evolution and conservation of this threatened species.

RESULTS

Chromosome-scale genome assembly and annotation

A total of 106.49 Gb HiFi reads and 373 Gb Hi-C sequencing data were generated by PacBio Sequel II and Illumina NovaSeq platform, respectively (Table S1). Based on these sequencing data, we obtained a chromosome-level genome assembly of *L. olivacea*, with the genome size of 2.22 Gb and contig N50 length of 97.32 Mb (Table 1). The genome size was similar to that of other published sea turtles[21–23] (*Chelonia mydas,* GCF_015237465.2; *Dermochelys coriacea*, GCF_009764565.3; *Eretmochelys imbricata*, GCA_030012505.1; *Caretta caretta*, GCF_023653815.1). The genome assembly contains 71 scaffolds that comprise 28 pseudo-chromosomes (covering 98.63% of the whole genome) and 43 unplaced scaffolds (Figures 1A and 1B; Tables 1 and S2).

We identified a total of ~56.59% (1.26Gb) repeat elements in the assembled genome, with the top three categories being DNA transposons (21.42%), long interspersed nuclear elements (LINE, 20.70%) and long terminal repeats (LTR, 14.11%), respectively (Figure 1B; Tables 1 and S3). Total of 18,273 protein-coding genes were predicted in the genome assembly and 97.38% of them were functionally annotated (Table 1). The BUSCO assessment of protein-coding genes based on Sauropsida_odb10 showed that the genome of *L. olivacea* harbored 98.6% of high conserved orthologues (Table 1), of which 97.6% were classified as "complete and singlecopy", 1% as "complete and duplicated" and 1.4% as "missing".

To further evaluate the completeness of the *L. olivacea* genome assembly, whole genome colinearity analyses were performed among five sea turtle species: *D. coriacea*, *C. mydas*, *E. imbricata*, *C. caretta*, and *L. olivacea*. The results revealed a high degree of sequence consistency across these species, consistent with findings from a previous study.[21] However, one notable exception was observed in a chromosome fusion/fission event between *C. caretta* and *L. olivacea*, where a portion of chromosome 13 and chromosome 28 of *C. caretta* corresponded to the chromosome 28 of *L. olivacea* (Figure 1C). In addition, using HiFi reads, we successfully assembled the complete mitochondrial genome of *L. olivacea*, with a total length of 16,687 bp (Figure S1). Haplotypic analysis of the D loop region showed that the reference genome individual clustered with haplotypes from individuals in Indonesia and Sri Lanka (Figure S2).

Phylogeny and evolution of protein-coding genes

Comparative genome analyses of 17 species (*L. olivacea* + 15 turtles +1 outgroup) (Table S4) revealed a total of 19,209 orthogroups, 455 species-specific orthogroups, 333,067 genes, and 7,148 single-copy orthologous genes. Based on these 7,148 single-copy genes, a maximum likelihood divergence-time tree was constructed. As shown in the phylogenetic tree (Figure 2A), *L. olivacea* had a close evolutionary relationship with *C. caretta* and the estimated divergence time of both species was approximately 11.1 Mya (95% highest posterior density [HPD] = 6.07–16.9). *D. coriacea* diverged from the ancestor of *C. mydas*, *E. imbricata*, *C. caretta* and *L. olivacea* approximately 58.3 Mya (95% highest posterior density [HPD] = 47.25–72.76). Overall, the inferred phylogenetic relationship of turtles was consistent with previous studies.[24–26]

Gene families analyses revealed that *L. olivacea* has 76 significantly expanded, 573 significantly contracted, and 56 rapidly evolving gene families (Figure 2A). Functional enrichment of 76 expanded gene families showed the significant GO/KEGG pathways mainly involved to "sensory perception of smell", "response to virus", "regulation of canonical NF-κB signal transduction", and "inflammatory response" (Figure 2B). Notably, olfactory receptor (OR) genes were the largest orthogroups in *L. olivacea* genome, accounting for 36.6% (329/900) of the entire expanded gene families. The number of OR class I genes (hydrophilic type) was ~3.4 times (254/75) to that of OR class II genes (hydrophobic type) (Figure 2C). We also observed that in *L. olivacea* genome, OR class I genes were only distributed on chromosome 1, while OR class II genes were distributed on chromosomes 6, 13, 14, and 28 (Figure 2D), which were similar with the genomes of *D. coriacea* and *C. mydas*.[21] In-depth examination of expanded gene families also revealed that 11 genes (e.g., *CALR*, *GBP1*, *GBP2*, *IFIT4*, *IFIT5*, *LGALS9*, *RNASE2*, *SLFN13*, and *TRIM41*) are assigned to virus defense (Figure 3A).

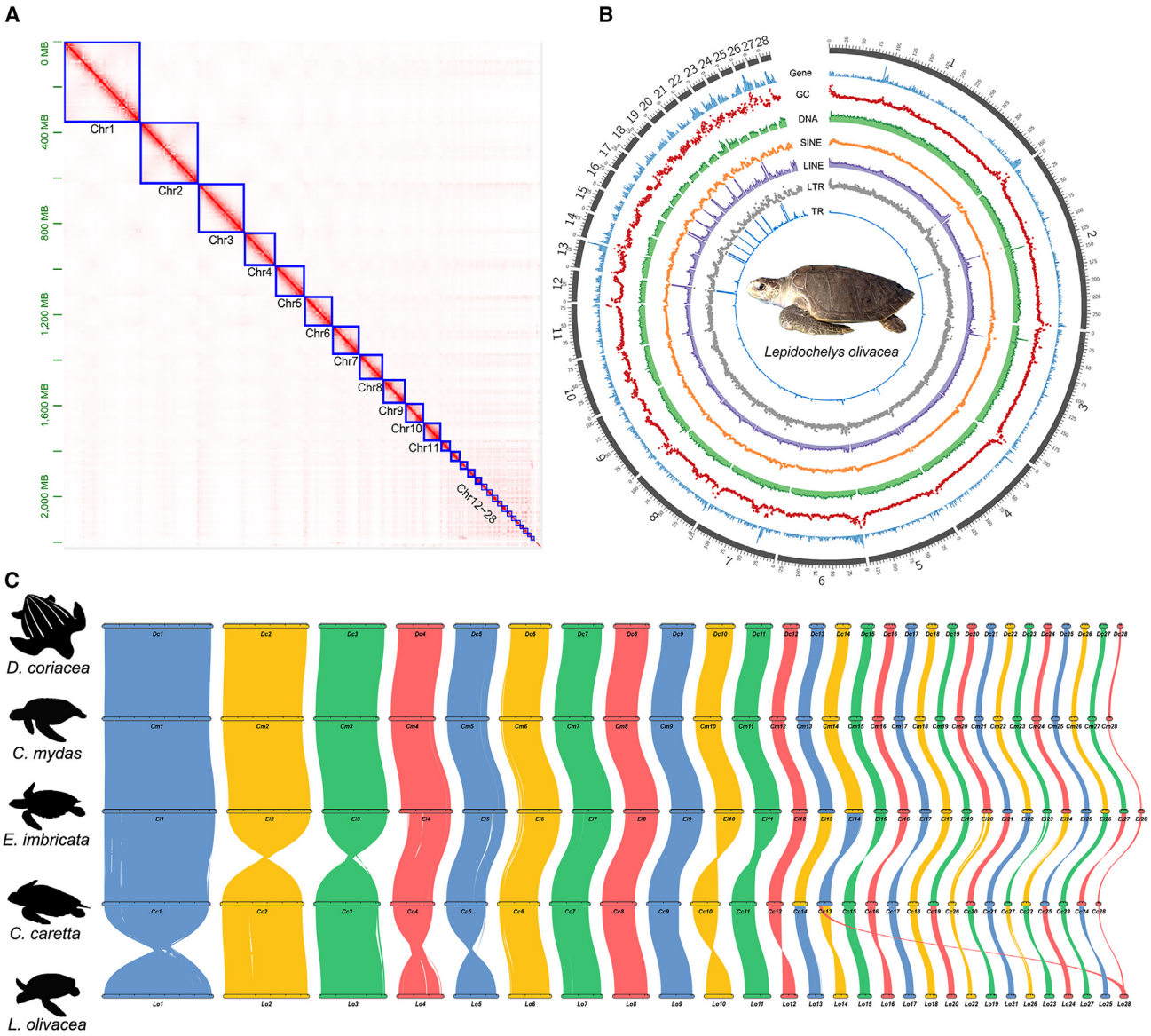


Figure 1. Genome architecture and synteny

(A) Hi-C map of *L. olivacea.* A total of 28 chromosomes were anchored in the genome of *L. olivacea*.

(B) Genome features of *L. olivacea*. Total of seven features including gene density, GC content, and five repeat element types (DNA, SINE, LINE, LTR, and TR) were plotted. Overall, the content or percentage of macro-chromosomes (Chr1∼Chr11, >70Mb) for most of features was lower than that of micro-chromosomes (Chr12∼Chr28, <70Mb).

(C) Genome collinearity among five sea turtles: *D. coriacea* (Leatherback turtle), *C. mydas* (Green turtle), *E. imbricata* (Hawksbill turtle), *C. caretta* (Loggerhead turtle) and *L. olivacea* (Olive ridley turtle). One chromosome fusion/fission event occurred between *C. caretta* and *L. olivacea*. Specifically, a portion of chromosome 13 and chromosome 28 of *C. caretta* corresponded to the chromosome 28 of *L. olivacea*.

Additionally, the contracted gene families were mainly enriched in the "innate immune response" and "response to bacterium" (Table S5).

To assess the evolutionary selective pressure of *L. olivacea*, we calculated the d_N/d_S ratio based on 7,148 single-copy genes. As a result, we detected 267 positively selected genes (PSGs) and 110 rapidly evolving genes (REGs) in the *L. olivacea* lineage (Tables S6 and S7). For PSGs, GO enrichment analysis revealed that they were mainly related to the "cell proliferation", "head development", "monoatomic cation transport", "response to xenobiotic stimulus", and "eye development" (Figures 3B and S3). KEGG enrichment analysis showed that they were mainly assigned to the "cell cycle", "DNA replication", and "Longevity regulating pathway" (Figures 3B and S3). Interestingly, we found that 4 PSGs (*ADCY9*, *STK11*, *TSC1*, and *SESN1*) are involved in longevity regulation (Figure 3A). For REGs, GO/KEGG enrichment analyses indicated that significant terms were mainly involved with "visual system development", "response to

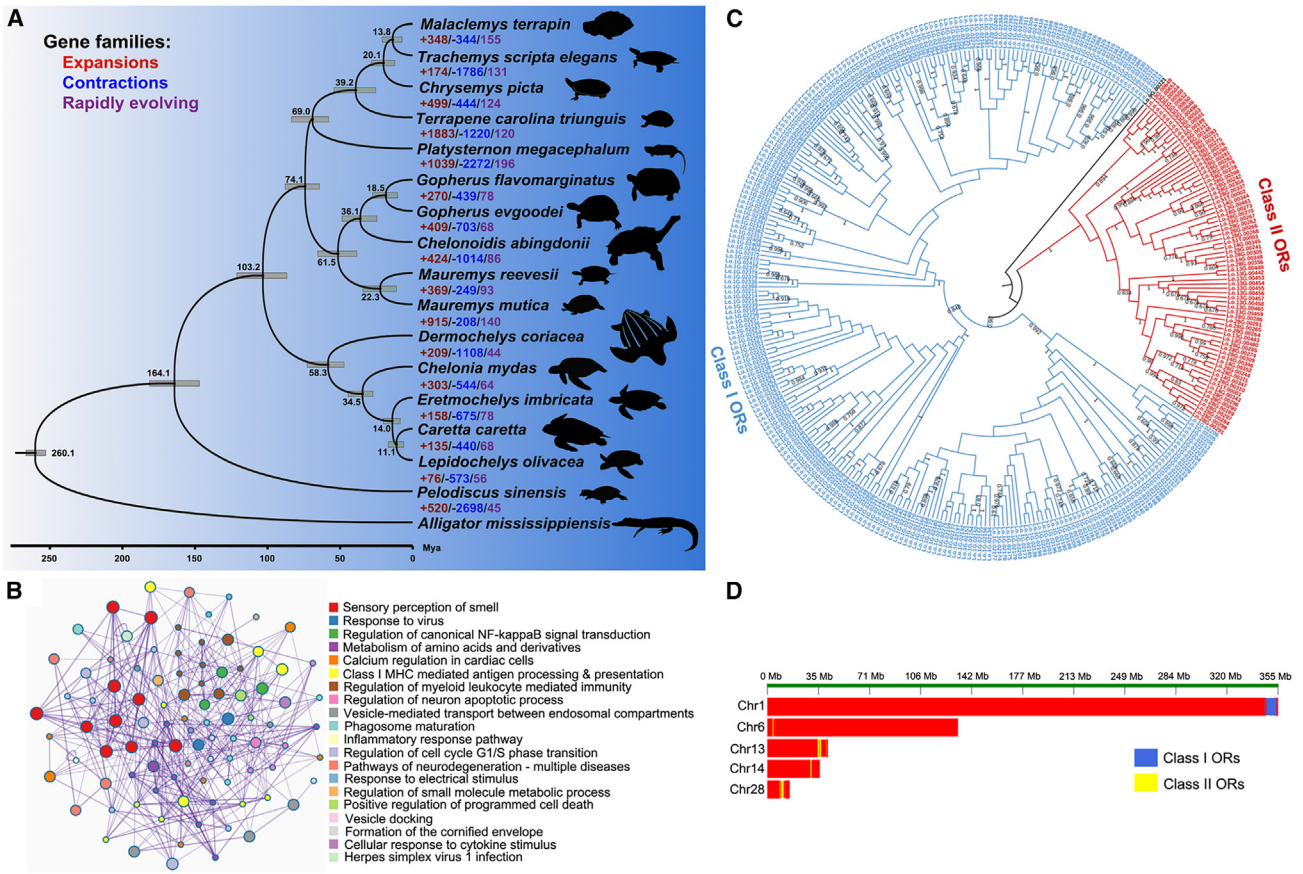


Figure 2. Phylogeny, expanded gene families, and expanded olfactory receptor genes of *L. olivacea*

(A) Phylogenetic relationship and divergence time estimation of 17 species including 16 turtles and 1 outgroup that used for comparative genomics analyses. The numbers of expanded (red), constructed (blue), and rapidly evolving (purple) gene families were labeled on each lineage/branch.

(B) Functional enrichment of 156 significantly expanded gene families in *L. olivacea*. The genes of sensory perception of smell (marked in red points) were the largest gene families in the *L. olivacea* genome, accounting for 36.6% of the entire expanded gene families.

(C) Phylogenetic tree of 329 expanded olfactory receptor genes (ORs) that composed of 254 hydrophilic class I ORs and 75 hydrophobic class II ORs in the genome of *L. olivacea*.

(D) The chromosome distribution of expanded olfactory receptor genes of *L. olivacea*. Class I ORs were only distributed on chromosome 1, while Class II ORs were distributed on chromosomes 6, 13, 14, and 28.

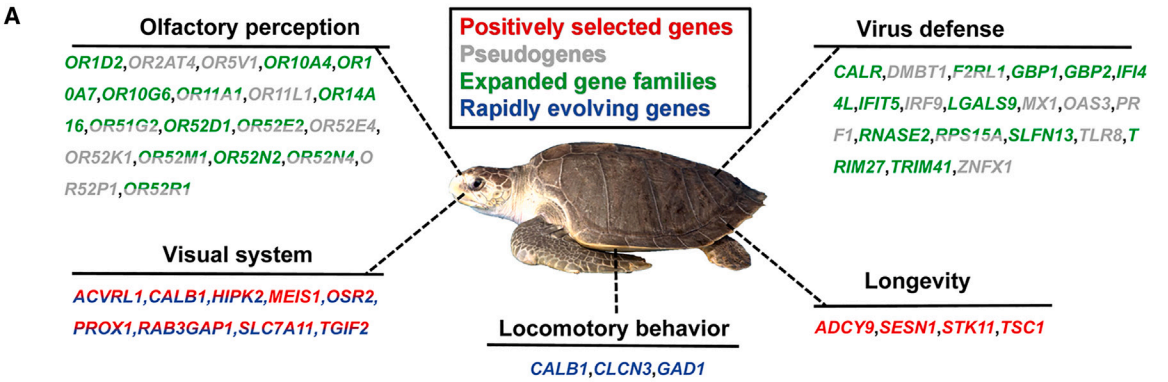
growth factor", "hemopoiesis", "regulation of autophagy", and "T cell receptor signaling pathway" (Figure S4). Further inspection of REGs observed that 8 genes (e.g., *CALB1*, *HIPK2*, *RAB3-GAP1*, and *TGIF2*) are related to visual system and 3 genes (*CALB1*, *CLCN3*, and *GAD1*) are associated with locomotory behavior (Figure 3A).

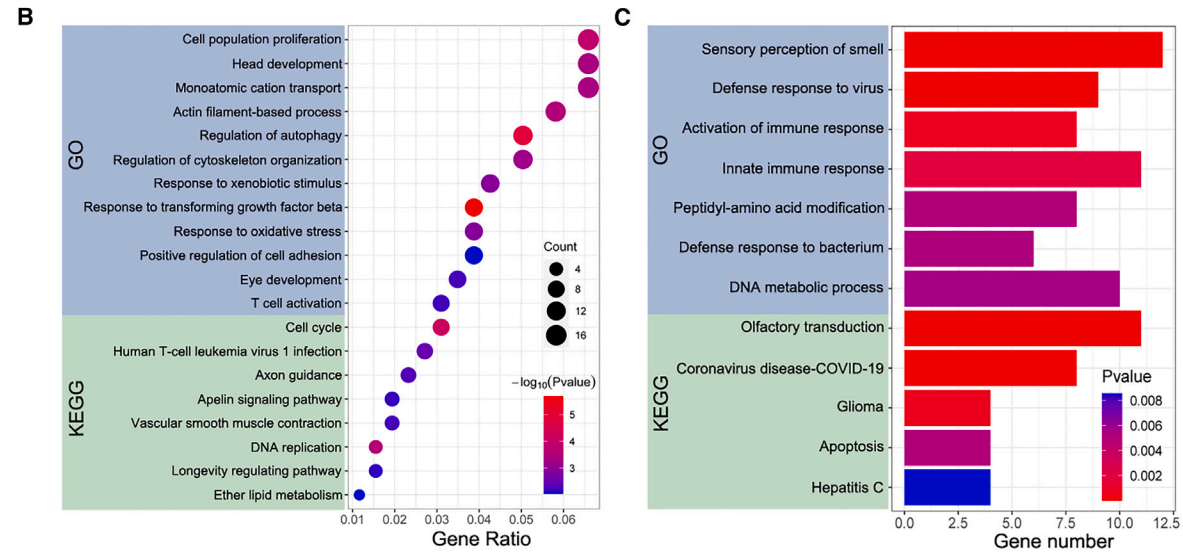
Pseudogenes play important biological roles for species' adaptive evolution.[27] Using the TOGA pipeline, a total of 156 pseudogenes were identified in the *L. olivacea* lineage (Table S8). Enrichment analyses revealed that these genes were mainly associated with terms of "sensory perception of smell", "olfactory transduction", "defense response to virus/bacterium", "innate immune response", and "DNA metabolic process" (Figures 3C and S5). In-depth analysis of pseudogenes revealed 12 genes (e.g., *OR2AT4*, *OR5V1*, *OR11L1*, *OR52E4*, *OR52K1*, and *OR52P1*) are related to olfactory perception and 11 genes (e.g., *DMBT1*, *IRF9*, *MX1*, *PRF1*, *ZNFX1*, and *ZNFX1*) are assigned to virus defense (Figure 3A).

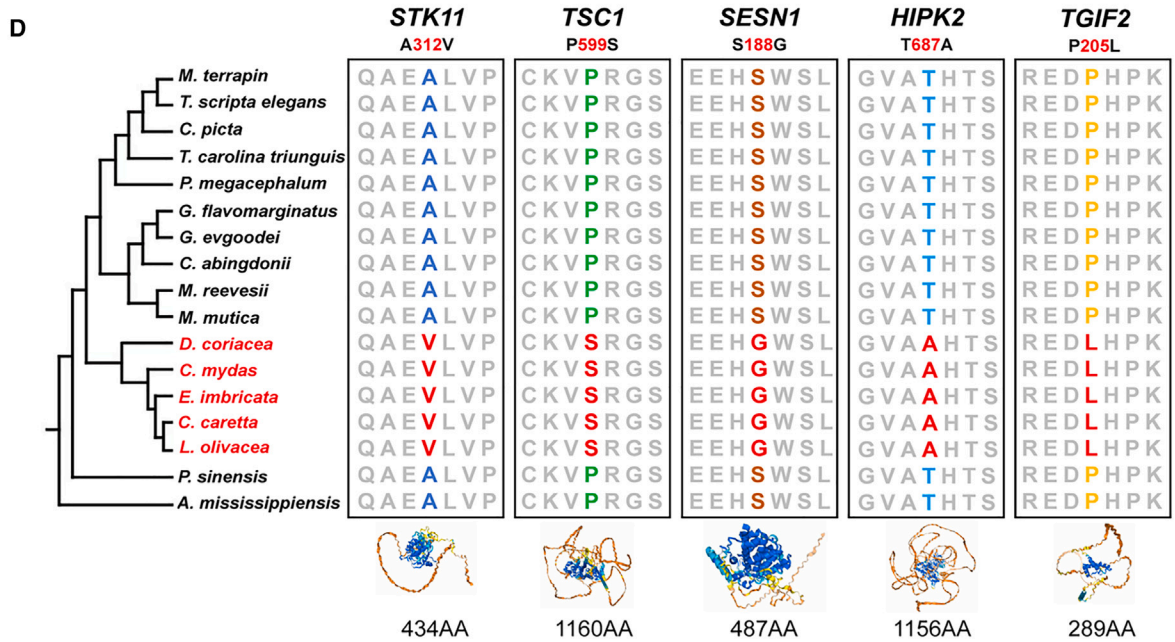
Moreover, in the detection of specific amino acid substitutions of the sea turtle group, we identified a total of 164 protein-coding genes comprising 549 substitutions overlapped with PSGs of *L. olivacea.* In-depth inspection observed that 3 genes (*STK11, TSC1* and *SESN1*) were associated with "longevity regulation" and that 2 genes (*HIPK2* and *TGIF2*) were related to "visual system" (Figures 3A and 3D). Notably, most of the amino acid substitutions of the above 5 genes were located in protein-domain regions (Figure 3D), suggesting that these genes may play an important role in the longevity and vision of sea turtles.

Genomic diversity, genetic load, and demographic history

Genetic variation is essential for the adaptability of a species/population.[28] In this study, a total of 1,432,715 SNP heterozygous sites and 623,810 INDEL heterozygous sites were identified in the genome of *L. olivacea*, with corresponding genome-wide heterozygosity of 6.45e-4 and 2.81e-4, respectively (Table S9).







(legend on next page)

We observed that the density pattern of SNP and INDEL for each chromosome was similar and the heterozygote density of macro-chromosomes (Chr1-Chr11, >70Mb) for both variant types was significantly lower than that of micro-chromosomes (Chr12-Chr28, <70Mb) (Figures 4A and 4B). To detect the neutral (non-coding region) and adaptive genetic variation (coding region) of *L. olivacea*, we further calculated the genomic heterozygosity on different genomic regions and found that for the SNP variant type, the order of genomic heterozygosity from high to low was repeat region (7.92e-4), non-exon region (6.48e-4), whole-genome region (6.45e-4), and exon region (4.01e-4), respectively (Figure 4C; Table S9). For the INDEL type, the heterozygosity pattern of different region was the same as the SNP variants, with order of repeat region (3.09e-4) > non-exon region (2.84e-4) > whole-genome region (2.81e-4) > exon region (8.45e-5) (Figure 4C; Table S9).

To further assess the genetic load of *L. olivacea*, we quantified the degree of impact of the mutations for both variant types and defined them as four levels: high-impact, moderate-impact, low-impact, and modifier-impact. We found that for SNP type, the order of variant-count ratio percentage from high to low was modifier-impact (42.3%, 606,138 sites), low-impact (0.51%, 7,310 sites), moderate-impact (0.45%, 6,517 sites), and high-impact (0.00789%, 113 sites), respectively (Figure 4D; Table S10). For INDEL type, the order of variant-count ratio percentage from high to low was modifier-impact (50.7%, 316,127 sites), high-impact (0.39%, 2,433 sites), low-impact (0.17%, 1,068 sites), and moderate-impact (0.07%, 439 sites), respectively (Figure 4D; Table S10). Furthermore, we checked the high-impact mutations of both SNP and INDEL and obtained 63 corresponding overlapping genes. The GO enrichment analysis revealed that they were mainly related to "sensory perception of smell", "response to wounding", and "inorganic ion transmembrane transport" (Figure 4F). These GO terms were similar with the results of expanded gene families and pseudogenes.

To reconstruct the demographic history of *L. olivacea*, pairwise sequential Markovian coalescent (PSMC) analysis was performed. As showed in Figure 4E, the demographic history of *L. olivacea* can be traced back to ~5 million years ago (Mya) to 10 Kya and the effective population sizes (*Ne*) have been decreasing historically (Figure 4E). The *Ne* fluctuation of Pacific region olive ridleys were similar to that of Atlantic region (Figure 4E). Notably, the *Ne* of *L. olivacea* (both Pacific and Atlantic regions) declined drastically in Plio-Pleistocene Transition and experienced two severe population bottlenecks in Mid-Pleistocene Transition and Last Glacial Maximum (Figure 4E). Consulting historical climate data found that the common features of these three ancient periods were the significant increase of global ice cover and the decrease of global temperature,[29–32] suggesting that the ancient glacial periods may have a significant impact on the *Ne* of *L. olivacea*.

DISCUSSION

Genetic basis of marine environment adaptation in *L. olivacea*

Sea turtles transitioned from the land to the ocean approximately 100 million years ago.[26,33] This ecological shift drives a range of genetic adaptive changes of sensory systems (e.g., olfaction and vision) and immune system (e.g., defense of virus and bacterium) for sea turtles in response to the long-term aquatic environmental stimuli. In this study, genomic analyses of the olive ridley turtle (*L. olivacea*) revealed the potential genetic basis underlying the physiological adaptations to the marine environment and longevity of this species.

Olfactory perception

The sense of smell plays an important role in foraging, avoiding predators, and reproduction for most marine species.[34] For sea turtles, their olfactory organs are composed of a lower chamber epithelium (LCE) and an upper chamber epithelium (UCE) with associated glands.[35] The LCE distributed with Class I ORs tend to be hydrophilic and mainly used to detect water odors, while the UCE distributed with Class II ORs tend to be hydrophobic and mainly used to detect air odors.[35] Notably, we observed that OR genes accounting for 36.6% (329/900) of the entire expanded gene families were the largest gene families in the *L. olivacea* genome, consistent with two previous studies of other sea turtle species (e.g., *C. mydas*[36] and *D. coriacea*[21]). The expanded 329 OR genes were composed of 254 hydrophilic class I ORs and 75 hydrophobic class II ORs (Figure 2C), suggesting that *L. olivacea* not only can detect the water-soluble odorants, but also can catch the airborne odorants and indeed, this ability may play a role in its navigation or foraging under natural conditions. Furthermore, the number of hydrophilic class I ORs was significantly higher (~3.4 times) than that of hydrophobic class II ORs (Figure 2C), implying that these expanded class I ORs may reflect the reliance and adaptation to the long-term aquatic environment for *L. olivacea*.

Additionally, we also found that some OR genes have underwent pseudogenization (Figures 3A and 3C). Similar results were also found in a previous *C. mydas* paper.[36] Meanwhile, multiple OR genes mutations were predicted to have a high impact on gene function (Figure 4F). These results suggested that natural selection may play multiple roles in the marine environment adaptation process for *L. olivacea*, with some olfactory genes being advantageous for aquatic adaptation and then expanding, while others are disadvantageous and underwent purifying selection or pseudogenization.

Visual system

Vision is an essential sensory process for both terrestrial and marine species.[37,38] For sea turtles, visual cues can help for their navigation, nest selection, and determining seaward direction.[39,40]

Figure 3. Genetic basis of marine environment adaptation of *L. olivacea*

(A) Genes with signatures of adaptive evolution for olfactory perception, visual system, virus defense, locomotory behavior, and longevity regulation in *L. olivacea*.

(B) GO/KEGG functional enrichment of positive selected genes in *L. olivacea*.

(C) GO/KEGG functional enrichment of pseudogenes in *L. olivacea*.

(D) The specific amino acid substitutions in genes *STK11*, *TSC1*, *SESN1*, *HIPK2*, and *TGIF2* of sea turtle group (marked in red). Each amino acid mutation was highlighted in a distinct color. The protein 3D structures of 5 genes were plotted using AlphaFold2 and most of the specific amino acid substitutions were located in domain region.

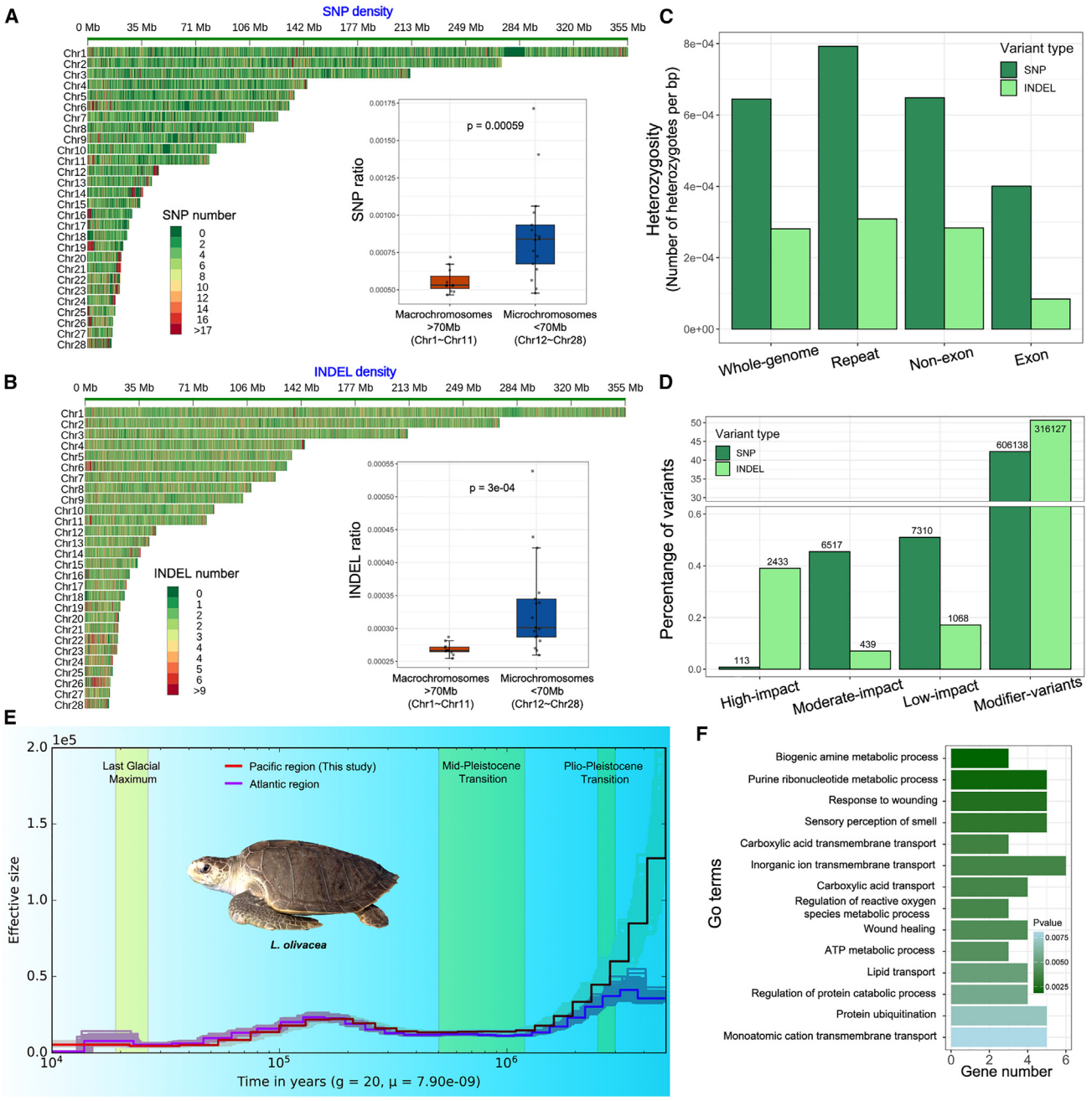


Figure 4. Genetic variation, genetic load and demographic history of *L. olivacea*

(A and B) SNP and INDEL density of 28 chromosomes in the genome of *L. olivacea*. The variant ratios of macro-chromosomes (Chr1∼Chr11, >70Mb) for both SNP and INDEL were significantly lower than that of micro-chromosomes (Chr12∼Chr28, <70Mb). Data are represented as mean ± SEM.

(C) Heterozygosity comparison of both SNP and INDEL in different regions including whole-genome, repeat region, non-exon region, and exon region.

(D) The percentage comparison of four annotated variant types including high-impact, moderate-impact, low-impact, and modifier variants in the genome of *L. olivacea*.

(E) Demographic history of Pacific (this study) and Atlantic (SRR15328381) olive ridleys simulated by PSMC with 100 bootstraps.

(F) Go enrichment of genes overlapping with the high-impact variants for both SNP and INDEL.

Comparing to land and/or freshwater turtles, sea turtles have a more spherical and rigid lens with a high refractive index, which compensates for the lack of corneal refraction.[38,41] The vision of sea turtles are primarily in the shorter wave-lengths (450–620 nm), which can intensively discern relatively small objects within the marine environment.[38] Significantly, positively selected/rapidly evolving genes (e.g., *CALB1*, *HIPK2*, *RAB3GAP1*, and *TGIF2*) of *L. olivacea* were enriched in the eye development

pathway (Figures 3A and 3B). In addition, two visual genes (*HIPK2* and *TGIF2*) have specific amino acid mutation sites in sea turtle group compared with other turtles (Figure 3D). *HIPK2*, an essential protein that promotes Notch signal transduction, plays numerous roles in regulation of eyeball size, lens formation and retinal morphogenesis during late embryogenesis.[42,43] Loss-of-function mutations of *HIPK2* in *Drosophila* displayed small, rough eyes.[42] Knockout of *HIPK2* in mice leads to a greater frequency of small eyes with a deficient lens and abnormally laminated and thickened retinas.[43] A previous study showed that *TGIF2* was highly expressed in the mouse retina at early stages of development and the knockdown of *TGIF2* in retinal explants resulted in a decreased number of rod photoreceptors but an increased number of cone photoreceptors.[44] In short, we speculated that mutations of these vision-related genes may contribute to the differences in the visual organs of sea turtles compared to land/freshwater turtles, although further functional experiments are needed for verification.

Virus defense

Viruses are one of the most abundant "lifeforms" in the ocean.[45] Viral infections affect the survival, health, and community composition of numerous marine organisms.[46] For sea turtles, they have a unique strong immune system that evolved over millions of years in a persistent process of host-pathogen arms race.[47] In our results, we found that some expanded family genes (e.g., *CALR*, *GBP1*, *IFIT5*, *LGALS9*, *RNASE2*, *SLFN13*, and *TRIM41*) were enriched in the pathway of virus defense and inflammatory response in the genome of *L. olivacea* (Figures 2B and 3A). Similarly, multiple pseudogenes (e.g., *DMBT1*, *IRF9*, *MX1*, *OSA3*, *PRF1*, *TLR8*, and *ZNXF4*) also were enriched in the pathway of virus or bacteria defense and innate immune response (Figures 3A and 3C). The presence of these innovative genes may help *L. olivacea* resist numerous marine viruses during its long-term evolutionary process.

Longevity regulation

Turtles are good models for studying longevity due to their slow and negligible senescence.[48–51] The estimated lifespan range of sea turtles is around 50–90 years,[52] significantly higher than that of most of other turtles, except for a few cases such as the giant tortoises.[48] Remarkably, four positively selected genes (*TSC1*, *SESN1*, *STK11,* and *ADCY9*) related to longevity were identified in the *L. olivacea* lineage (Figures 3A and 3B). Three (*TSC1*, *SESN1*, and *STK11*) of them have specific amino acid variants in sea turtle group relative to other turtles (Figure 3D). Moreover, two mutations (P599S in *TSC1*, S188G in *SESN1*) of these gene will lead to changes in the physicochemical properties of amino acids. A previous study proved that *TSC1* transgenic mice exhibited changes in multiple signaling pathways and biological processes related to the mTOR-*TSC1/2* axis, resulting in improved overall health status, particularly in female mice with increased survival rates.[53] Over-expression of *TSC1* extended lifespan in *Drosophila*[54]. It is also proved that over-expression of *SESN1* could prolong lifespan in *Caenorhabditis elegans* by defense against heat, hydrogen peroxide and the heavy metal copper.[55] *STK11* encodes a serine/threonine kinase and its over-expression increased lifespan by down-regulation of *S6K* in *Drosophila*[56]. To sum up, the specific mutations of these genes may play some roles in lifespan regulation for sea turtles, although functional experiments are required in the future to verify the efficacy.

Low genomic diversity and long-term population decline of *L. olivacea*

Genetic variation could to some extent reflect the ability to adapt to environmental changes for a species/population.[28] In our results, comparing to other sea turtle species except for *D. coriacea*,[21] *L. olivacea* exhibited a lower level of genome-wide heterozygosity (6.45e-4) (Table S9), which was generally consistent with a previous estimate based on olive ridley resequencing data mapped to the *C. mydas* reference genome.[25] The genetic diversity of neutral region such as the repeat and non-coding region was higher than that of coding region in genome of *L. olivacea*. Similar patterns were also found in other four sea turtles (*D. coriacea*, *C. mydas*, *E. imbricata*, and *C. caretta*)[21–23,25] (Table S9). In addition, we observed that for both SNP and INDEL variant types, the genomic diversity of macrochromosomes was significantly lower than that of microchromosomes (Figures 4A and 4B). This pattern was also seen in other genetic indicators, such as gene density, LINE, and TR repeat elements (Figure 1B). In fact, this is a known pattern observed in non-avian/avian and reptiles, where microchromosomes recombine more often and are subject to more rearrangements than macrochromosomes.[21,57] The underlying biological significance of this phenomenon deserves further investigation.

The demographic history showed that the population sizes of *L. olivacea* have been in a declining state from ~5 million years ago (Mya) to 10 Kya (Figure 4E), which may explain why *L. olivacea* currently harbored a low genomic diversity. Three ancient glacial events[29–32] (Plio-Pleistocene Transition, Mid-Pleistocene Transition and Last Glacial Maximum) with the occurrence of increase of global ice cover and decrease of global temperature further accelerated the population size reduction for *L. olivacea*. Besides, for decades, the population of *L. olivacea* has been substantially declining by the effect of human disturbance such as habitat degradation, fisheries bycatch, and pollution.[8,14] Therefore, both ancient climate changes and human disturbance may cause the current low genomic diversity of *L. olivacea* and more attention and protection should be focused in the future for this threatened species.

Conclusions

In this study, we generated a high-quality chromosome-scale genome assembly of *L. olivacea*, with contig N50 of 97.3Mb and protein-coding gene BUSCO of 98.6%. Our findings revealed the potential genetic footprints of marine environment evolutionary adaptation for *L. olivacea* including olfactory perception, visual system, virus/or bacteria defense, and longevity regulation. Furthermore, we also found a low genomic diversity in the genome of *L. olivacea*, which was consistent with its long-term historically population size decline. In summary, this study provides valuable genetic resources for phenotypic/physiological and evolutionary biology, and conservation studies of *L. olivacea*.

Limitations of the study

Sea turtles are a distinctive group of large reptiles that have undergone secondary adaptation to the marine environment,

representing a significant ecological transition from land to sea. Our study presents a high-quality genome of the olive ridley turtle and conducts comparative genomic analyses across five known sea turtle species. However, achieving a comprehensive understanding of the evolutionary history of the entire sea turtle lineage will require further in-depth analyses, particularly with the inclusion of the remaining two sea turtle species. Moreover, the candidate genes identified in this study—associated with olfaction, vision, and longevity—warrant further functional investigations and experimental validation to elucidate their roles in marine adaptation. Given the global distribution and migration of sea turtles, international collaboration and genomic resources sharing are critical for the development of effective conservation and management strategies for this endangered group.

RESOURCE AVAILABILITY

Lead contact

Further information and requests can be directed to Prof. Wenliang Zhou (zhouwl@gmlab.ac.cn) and Prof. Fuwen Wei (weifw@ioz.ac.cn).

Materials availability statements

The study did not generate new unique reagents.

Data and code availability

- The genomic data in this work has been released at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation (https://ngdc.cncb.ac.cn/bioproject/browse/PRJCA030559) (BioProject: PRJCA030559). The raw sequencing data of PacBio HiFi, Hi-C used for *L. olivacea* genome assembly has been deposited in the Genome Sequence Archive at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation (https://ngdc.cncb.ac.cn/gsa/browse/CRA020668) (GSA: CRA020668). The genome assembly of *L. olivacea* reported in this study has been deposited in the Genome Warehouse at the National Genomics Data Center, Beijing Institute of Genomics, Chinese Academy of Sciences/China National Center for Bioinformation (https://ngdc.cncb.ac.cn/gwh/Assembly/86766/show) (GWH: GWHFHHW00000000.1).
- This paper does not report original code.
- Any additional information required to reanalyze the data reported in this work is available from the lead contact upon request.

ACKNOWLEDGMENTS

This study was funded by National Natural Science Foundation of China (32300419 and 31821001), China Postdoctoral Science Foundation (2023M740814), Science and Technology Department, Guangdong Province (2021QN02H103, 2023A1111110001), the Ministry of Science and Technology (China) of China (2022YFF1301601), Guangdong Forestry Administration (SLYJ2023B4004), Science and Technology Department of JiangXi Province (20233ACB209001), and the PI Project of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) (GML2020GD0804, GML2022GD0804).

AUTHOR CONTRIBUTIONS

F.W. and W.L. conceived and supervised the project. L.Y., Y.T., S.F., C.L., W.L., and F.W. collected the samples of *L. olivacea.* L.Y., X.H., and X.D. performed the bioinformatics data analyses. F.W., W.L., L.Y., and Y.T. discussed the analyses results. L.Y. wrote and revised the manuscript.

DECLARATION OF INTERESTS

The authors declare no competing interests.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
  - Genomic DNA extraction
  - Library preparation and genomic DNA sequencing
  - Hi-C library preparation and sequencing
  - Genome assembly
  - Repeat annotation
  - Gene annotation
  - Genome assessment and synteny
  - Mitochondrial genome assembly and haplotype identification
  - Gene family identification and phylogeny tree construction
  - Expansion and contraction of gene families
  - Positive selected genes and rapidly evolving genes
  - Pseudogene
  - Identification of specific amino acid substitutions
  - Heterozygosity and genetic load
  - Demographic history simulation
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.isci.2025.111776.

Received: September 28, 2024
Revised: December 2, 2024
Accepted: January 7, 2025
Published: January 8, 2025

REFERENCES

1. Heithaus, M.R. (2013). Predators, prey, and the ecological roles of sea turtles. Biol. Sea Turtles *3*, 249.
2. Hirayama, R. (1998). Oldest known sea turtle. Nature *392*, 705–708.
3. Thewissen, J.G.M., Taylor, M.A., and Hall, B.K. (2007). Aquatic adaptations in the limbs of amniotes. In Fins into limbs: Evolution, development, and transformation, pp. 310–322.
4. Schmidt-Nielsen, K., and Fange, R. (1958). Salt glands in marine reptiles. Nature *182*, 783–785.
5. Nagelkerken, I., Doney, S.C., and Munday, P.L. (2019). Consequences of anthropogenic changes in the sensory landscape of marine animals. Oceanogr. Mar. Biol. *57*, 229–264.
6. Kalman, B. (2004). Endangered Sea Turtles (Crabtree Publishing Company).
7. Marcovaldi, M.A. (1999). Status and distribution of the olive ridley turtle, *Lepidochelys olivacea*, in the Western Atlantic Ocean. In In Proceedings of the Regional Meeting, Marine Turtle Conservation in the Wider Caribbean Region: A Dialogue for Effective Regional Management (Santo Domingo), pp. 16–18.
8. Abreu-Grobois, A., and Plotkin, P. (2008). Lepidochelys olivacea. In The IUCN Red List of Threatened Species, pp. 6–8.
9. Pacific Sea Turtle Recovery Team (U.S.) (1997). Recovery plan for US Pacific populations of the olive ridley turtle. Lepidochelys Olivacea.

10. Behera, S.K., Sivakumar, K., Choudhury, B., and John, S. (2014). Diet preference and prey of olive ridley turtles (*Lepidocheyls olivacea*) along East Coast of India, Odisha. Open J. Ocean Coast Sci. *1*, 73–82.

11. Ocana, M., Harfush-Melendez, M., and Heppell, S.S. (2012). Mass nesting of olive ridley sea turtles *Lepidochelys olivacea* at La Escobilla, Mexico: linking nest density and rates of destruction. Endanger. Species Res. *16*, 45–54.

12. Plotkin, P.T. (2007). Biology and Conservation of Ridley Sea Turtles (JHU Press).

13. Cáceres-Farias, L., Reséndiz, E., Espinoza, J., Fernández-Sanz, H., and Alfaro-Núñez, A. (2022). Threats and vulnerabilities for the globally distributed Olive Ridley (*Lepidochelys olivacea*) sea turtle: A historical and current status evaluation. Animals *12*, 1837.

14. Shanker, K., Ramadevi, J., Choudhury, B.C., Singh, L., and Aggarwal, R.K. (2004). Phylogeography of olive ridley turtles (*Lepidochelys olivacea*) on the east coast of India: implications for conservation theory. Mol. Ecol. *13*, 1899–1909.

15. Shanker, K., Abreu-Grobois, A., Bezy, V., Briseno, R., Colman, L., Girard, A., Girondot, M., Jensen, M.P., Manoharakrishnan, M., Rguez-Baron, J.M., and Valverde, R. (2021). Olive Ridleys: the quirky turtles that conquered the world. In In The State of the World's Sea Turtles SWOT: Report 16 (Oceanic Society), pp. 24–33.

16. Vilaça, S.T., Hahn, A.T., Naro-Maciel, E., Abreu-Grobois, F.A., Bowen, B.W., Castilhos, J.C., Ciofi, C., FitzSimmons, N.N., Jensen, M.P., Formia, A., et al. (2022). Global phylogeography of ridley sea turtles (*Lepidochelys spp.*): evolution, demography, connectivity, and conservation. Conserv. Genet. *23*, 995–1010.

17. Sosa-Cornejo, I., Peinado-Guevara, L.I., Contreras-Aguilar, H.R., Enciso-Saracho, F., Sandoval-Bautista, M., Enciso-Padilla, I., and Campista-León, S. (2022). Evaluation of nest management phases for *Lepidochelys olivacea* at two beaches in Northwest Mexico. Environ. Monit. Assess. *194*, 130.

18. Hu, Y., Wu, Q.I., Ma, S., Ma, T., Shan, L., Wang, X., Nie, Y., Ning, Z., Yan, L.I., Xiu, Y., and Wei, F. (2017). Comparative genomics reveals convergent evolution between the bamboo-eating giant and red pandas. Proc. Natl. Acad. Sci. USA *114*, 1081–1086.

19. Yang, L., Wei, F., Zhan, X., Fan, H., Zhao, P., Huang, G., Chang, J., Lei, Y., and Hu, Y. (2022). Evolutionary conservation genomics reveals recent speciation and local adaptation in threatened takins. Mol. Biol. Evol. *39*, msac111.

20. Hu, Y., Wang, X., Xu, Y., Yang, H., Tong, Z., Tian, R., Xu, S., Yu, L., Guo, Y., Shi, P., et al. (2023). Molecular mechanisms of adaptive evolution in wild animals and plants. Sci. China Life Sci. *66*, 453–495.

21. Bentley, B.P., Carrasco-Valenzuela, T., Ramos, E.K.S., Pawar, H., Souza Arantes, L., Alexander, A., Banerjee, S.M., Masterson, P., Kuhlwilm, M., Pippel, M., et al. (2023). Divergent sensory and immune gene evolution in sea turtles with contrasting demographic and life histories. Proc. Natl. Acad. Sci. USA *120*, e2201076120.

22. Guo, Y., Tang, J., Zhuo, Z., Huang, J., Fu, Z., Song, J., Liu, M., Dong, Z., and Wang, Z. (2023). The first high-quality chromosome-level genome of *Eretmochelys imbricata* using HiFi and Hi-C data. Sci. Data *10*, 604.

23. Chang, G., Jones, S., Leelakumari, S., Ashkani, J., Culibrk, L., O'Neill, K., Tse, K., Cheng, D., Chuah, E., McDonald, H., and Kirk, H. (2023). The genome sequence of the Loggerhead sea turtle, Caretta caretta Linnaeus 1758. F1000Research *12*, 336.

24. Phillips, K.F., Martin, K.R., Stahelin, G.D., Savage, A.E., and Mansfield, K.L. (2022). Genetic variation among sea turtle life stages and species suggests connectivity among ocean basins. Ecol. Evol. *12*, e9426.

25. Vilaça, S.T., Piccinno, R., Rota-Stabelli, O., Gabrielli, M., Benazzo, A., Matschiner, M., Soares, L.S., Bolten, A.B., Bjorndal, K.A., and Bertorelle, G. (2021). Divergence and hybridization in sea turtles: Inferences from genome data show evidence of ancient gene flow between species. Mol. Ecol. *30*, 6178–6192.

26. Thomson, R.C., Spinks, P.Q., and Shaffer, H.B. (2021). A global phylogeny of turtles reveals a burst of climate-associated diversification on continental margins. Proc. Natl. Acad. Sci. USA *118*, e2012215118.

27. Cheetham, S.W., Faulkner, G.J., and Dinger, M.E. (2020). Overcoming challenges and dogmas to understand the functions of pseudogenes. Nat. Rev. Genet. *21*, 191–201.

28. Kardos, M., Armstrong, E.E., Fitzpatrick, S.W., Hauser, S., Hedrick, P.W., Miller, J.M., Tallmon, D.A., and Funk, W.C. (2021). The crucial role of genome-wide genetic variation in conservation. Proc. Natl. Acad. Sci. USA *118*, e2104642118.

29. Anee, S. (2021). Evidence for ice covered West Antarctica during late Pliocene/early Pleistocene transition. Doctoral dissertation (Wesleyan University).

30. Herbert, T.D. (2023). The mid-Pleistocene climate transition. Annu. Rev. Earth. Planet. Sci. *51*, 389–418.

31. An, Z., Zhou, W., Zhang, Z., Zhang, X., Liu, Z., Sun, Y., Clemens, S.C., Wu, L., Zhao, J., Shi, Z., et al. (2024). Mid-Pleistocene climate transition triggered by Antarctic Ice Sheet growth. Science *385*, 560–565.

32. Osman, M.B., Tierney, J.E., Zhu, J., Tardif, R., Hakim, G.J., King, J., and Poulsen, C.J. (2021). Globally resolved surface temperatures since the Last Glacial Maximum. Nature *599*, 239–244.

33. Chiarenza, A.A., Waterson, A.M., Schmidt, D.N., Valdes, P.J., Yesson, C., Holroyd, P.A., Collinson, M.E., Farnsworth, A., Nicholson, D.B., Varela, S., and Barrett, P.M. (2023). 100 million years of turtle paleoniche dynamics enable the prediction of latitudinal range shifts in a warming world. Curr. Biol. *33*, 109–121.e3.

34. Kishida, T. (2021). Olfaction of aquatic amniotes. Cell Tissue Res. *383*, 353–365.

35. Nakamuta, S., Mori, M., Ito, M., Kurita, M., Miyazaki, M., Yamamoto, Y., and Nakamuta, N. (2023). In situ hybridization analysis of olfactory receptor expression in the sea turtle olfactory organ. Cell Tissue Res. *393*, 253–264.

36. Wang, Z., Pascual-Anaya, J., Zadissa, A., Li, W., Niimura, Y., Huang, Z., Li, C., White, S., Xiong, Z., Fang, D., et al. (2013). The draft genomes of soft-shell turtle and green sea turtle yield insights into the development and evolution of the turtle-specific body plan. Nat. Genet. *45*, 701–706.

37. Osorio, D., and Vorobyev, M. (2005). Photoreceptor sectral sensitivities in terrestrial animals: adaptations for luminance and colour vision. Proc. Biol. Sci. *272*, 1745–1752.

38. Bartol, S.M., Musick, J.A., Lutz, P.L., and Wyneken, J. (2003). Sensory biology of sea turtles. Biol. Sea Turtles *II*, 79–102.

39. Lohmann, K.J., Witherington, B.E., Lohmann, C.M., and Salmon, M. (2017). Orientation, navigation, and natal beach homing in sea turtles. In The Biology of Sea Turtles, *Volume I* (CRC Press), pp. 108–135.

40. Ehrenfeld, D.W., and Carr, A. (1967). The role of vision in the sea-finding orientation of the green turtle (*Chelonia mydas*). Anim. Behav. *15*, 25–36.

41. Northmore, D.P., and Granda, A.M. (1991). Ocular dimensions and schematic eyes of freshwater and sea turtles. Vis. Neurosci. *7*, 627–635.

42. Lee, W., Andrews, B.C., Faust, M., Walldorf, U., and Verheyen, E.M. (2009). *HIPK* is an essential protein that promotes Notch signal transduction in the *Drosophila* eye by inhibition of the global co-repressor Groucho. Dev. Biol. *325*, 263–272.

43. Inoue, T., Kagawa, T., Inoue-Mochita, M., Isono, K., Ohtsu, N., Nobuhisa, I., Fukushima, M., Tanihara, H., and Taga, T. (2010). Involvement of the *HIPK* family in regulation of eyeball size, lens formation and retinal morphogenesis. FEBS Lett. *584*, 3233–3238.

44. Kuribayashi, H., Tsuhako, A., Kikuchi, M., Yoshida, N., Koso, H., and Watanabe, S. (2016). Role of transcription factor *TGIF2* in photoreceptor differentiation in the mouse retina. Exp. Eye Res. *152*, 34–42.

45. Suttle, C.A. (2007). Marine viruses—major players in the global ecosystem. Nat. Rev. Microbiol. *5*, 801–812.

46. Wommack, K.E., and Colwell, R.R. (2000). Virioplankton: viruses in aquatic ecosystems. Microbiol. Mol. Biol. Rev. *64*, 69–114.

47. Nash, A., and Ryan, E.J. (2023). Immunity in sea turtles: review of a host-pathogen arms race millions of years in the running. Animals *13*, 556.

48. Quesada, V., Freitas-Rodríguez, S., Miller, J., Pérez-Silva, J.G., Jiang, Z.F., Tapia, W., Santiago-Fernández, O., Campos-Iglesias, D., Kuderna, L.F.K., Quinzin, M., et al. (2019). Giant tortoise genomes provide insights into longevity and age-related disease. Nat. Ecol. Evol. *3*, 87–95.

49. Bradley Shaffer, H., Minx, P., Warren, D.E., Shedlock, A.M., Thomson, R.C., Valenzuela, N., Abramyan, J., Amemiya, C.T., Badenhorst, D., Biggar, K.K., et al. (2013). The western painted turtle genome, a model for the evolution of extreme physiological adaptations in a slowly evolving lineage. Genome Biol. *14*, R28.

50. Reinke, B.A., Cayuela, H., Janzen, F.J., Lemaitre, J.F., Gaillard, J.M., Lawing, A.M., Iverson, J.B., Christiansen, D.G., Martínez-Solano, I., Sánchez-Montes, G., et al. (2022). Diverse aging rates in ectothermic tetrapods provide insights for the evolution of aging and longevity. Science *376*, 1459–1466.

51. da Silva, R., Conde, D.A., Baudisch, A., and Colchero, F. (2022). Slow and negligible senescence among testudines challenges evolutionary theories of senescence. Science *376*, 1466–1470.

52. Mayne, B., Tucker, A.D., Berry, O., and Jarman, S. (2020). Lifespan estimation in marine turtles using genomic promoter CpG density. PLoS One *15*, e0236888.

53. Zhang, H.M., Diaz, V., Walsh, M.E., and Zhang, Y. (2017). Moderate lifelong overexpression of tuberous sclerosis complex 1 (*TSC1)* improves health and survival in mice. Sci. Rep. *7*, 834.

54. Kapahi, P., Zid, B.M., Harper, T., Koslover, D., Sapin, V., and Benzer, S. (2004). Regulation of lifespan in *Drosophila* by modulation of genes in the TOR signaling pathway. Curr. Biol. *14*, 885–890.

55. Yang, Y.L., Loh, K.S., Liou, B.Y., Chu, I.H., Kuo, C.J., Chen, H.D., and Chen, C.S. (2013). *SESN-1* is a positive regulator of lifespan in *Caenorhabditis elegans*. Exp. Gerontol. *48*, 371–379.

56. Funakoshi, M., Tsuda, M., Muramatsu, K., Hatsuda, H., Morishita, S., and Aigaki, T. (2011). A gain-of-function screen identifies *WDB* and *LKB1* as lifespan-extending genes in *Drosophila*. Biochem. Biophys. Res. Commun. *405*, 667–672.

57. Srikulnath, K., Ahmad, S.F., Singchat, W., and Panthum, T. (2021). Why do some vertebrates have microchromosomes? Cell *10*, 2182.

58. De Coster, W., and Rademakers, R. (2023). NanoPack2: population-scale evaluation of long-read sequencing data. Bioinformatics *39*, btad311.

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

60. Guan, D., McCarthy, S.A., Wood, J., Howe, K., Wang, Y., and Durbin, R. (2020). Identifying and removing haplotypic duplication in primary genome assemblies. Bioinformatics *36*, 2896–2898.

61. Zhou, C., McCarthy, S.A., and Durbin, R. (2023). YaHS: yet another Hi-C scaffolding tool. Bioinformatics *39*, btac808.

62. Durand, N.C., Robinson, J.T., Shamim, M.S., Machol, I., Mesirov, J.P., Lander, E.S., and Aiden, E.L. (2016). Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. Cell Syst. *3*, 99–101.

63. 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. Proc. Natl. Acad. Sci. USA *117*, 9451–9457.

64. Tarailo-Graovac, M., and Chen, N. (2009). Using RepeatMasker to identify repetitive elements in genomic sequences. Current *25*, 4–10.

65. Jurka, J., Kapitonov, V.V., Pavlicek, A., Klonowski, P., Kohany, O., and Walichiewicz, J. (2005). Repbase Update, a database of eukaryotic repetitive elements. Cytogenet. Genome Res. *110*, 462–467.

66. Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. Nucleic Acids Res. *27*, 573–580.

67. Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. Bioinformatics *24*, 637–644.

68. Burge, C., and Karlin, S. (1997). Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. *268*, 78–94.

69. Majoros, W.H., Pertea, M., and Salzberg, S.L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. Bioinformatics *20*, 2878–2879.

70. Keilwagen, J., Hartung, F., and Grau, J. (2019). GeMoMa: homology-based gene prediction utilizing intron position conservation and RNA-seq data. Methods Mol. Biol. *1962*, 161–177.

71. Bairoch, A., and Apweiler, R. (2000). The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. Nucleic Acids Res. *28*, 45–48.

72. Pruitt, K.D., Tatusova, T., and Maglott, D.R. (2005). NCBI Reference Sequence (RefSeq): a curated non-redundant sequence database of genomes, transcripts and proteins. Nucleic Acids Res. *33*, D501–D504.

73. Buchfink, B., Xie, C., and Huson, D.H. (2015). Fast and sensitive protein alignment using DIAMOND. Nat. Methods *12*, 59–60.

74. Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., and Zdobnov, E.M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics *31*, 3210–3212.

75. Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: an information aesthetic for comparative genomics. Genome Res. *19*, 1639–1645.

76. Wang, Y., Tang, H., DeBarry, J.D., Tan, X., Li, J., Wang, X., Lee, T.H., Jin, H., Marler, B., Guo, H., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. *40*, e49.

77. Uliano-Silva, M., Ferreira, J.G.R.N., Krasheninnikova, K., Darwin Tree of Life Consortium; Formenti, G., Abueg, L., Torrance, J., Myers, E.W., Durbin, R., Blaxter, M., and McCarthy, S.A. (2023). MitoHiFi: a python pipeline for mitochondrial genome assembly from PacBio high fidelity reads. BMC Bioinf. *24*, 288.

78. Allio, R., Schomaker-Bastos, A., Romiguier, J., Prosdocimi, F., Nabholz, B., and Delsuc, F. (2020). MitoFinder: Efficient automated large-scale extraction of mitogenomic data in target enrichment phylogenomics. Mol. Ecol. Resour. *20*, 892–905.

79. Laslett, D., and Canbäck, B. (2008). ARWEN: a program to detect tRNA genes in metazoan mitochondrial nucleotide sequences. Bioinformatics *24*, 172–175.

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

81. Kumar, S., Stecher, G., Li, M., Knyaz, C., and Tamura, K. (2018). MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. Mol. Biol. Evol. *35*, 1547–1549.

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

83. Ranwez, V., Harispe, S., Delsuc, F., and Douzery, E.J.P. (2011). MACSE: Multiple Alignment of Coding Sequences accounting for frameshifts and stop codons. PLoS One *6*, e22594.

84. Castresana, J. (2000). Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. *17*, 540–552.

85. Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics *30*, 1312–1313.

86. Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. Mol. Biol. Evol. *24*, 1586–1591.

87. Kumar, S., Stecher, G., Suleski, M., and Hedges, S.B. (2017). TimeTree: a resource for timelines, timetrees, and divergence times. Mol. Biol. Evol. *34*, 1812–1819.

88. De Bie, T., Cristianini, N., Demuth, J.P., and Hahn, M.W. (2006). CAFE: a computational tool for the study of gene family evolution. Bioinformatics *22*, 1269–1271.

89. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., and Chanda, S.K. (2019). Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nat. Commun. *10*, 1523.

90. Kirilenko, B.M., Munegowda, C., Osipova, E., Jebb, D., Sharma, V., Blumer, M., Morales, A.E., Ahmed, A.W., Kontopoulos, D.G., Hilgers, L., et al. (2023). Integrating gene annotation with orthology inference at scale. Science *380*, eabn3107.

91. Sun, Y.B. (2017). FasParser: a package for manipulating sequence data. Zool. Res. *38*, 110–112.

92. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589.

93. Cingolani, P., Platts, A., Wang, L.L., Coon, M., Nguyen, T., Wang, L., Land, S.J., Lu, X., and Ruden, D.M. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of Drosophila melanogaster strain w1118; iso-2; iso-3. Fly *6*, 80–92.

94. Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics *25*, 1754–1760.

95. Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., and Durbin, R.; 1000 Genome Project Data Processing Subgroup (2009). The sequence alignment/map format and SAMtools. Bioinformatics *25*, 2078–2079.

96. Danecek, P., Bonfield, J.K., Liddle, J., Marshall, J., Ohan, V., Pollard, M.O., Whitwham, A., Keane, T., McCarthy, S.A., Davies, R.M., and Li, H. (2021). Twelve years of SAMtools and BCFtools. GigaScience *10*, giab008.

97. Li, H., and Durbin, R. (2011). Inference of human population history from individual whole-genome sequences. Nature *475*, 493–496.

98. Belton, J.M., McCord, R.P., Gibcus, J.H., Naumova, N., Zhan, Y., and Dekker, J. (2012). Hi–C: a comprehensive technique to capture the conformation of genomes. Methods *58*, 268–276.

99. Haas, B.J., Salzberg, S.L., Zhu, W., Pertea, M., Allen, J.E., Orvis, J., White, O., Buell, C.R., and Wortman, J.R. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the program to assemble spliced alignments. Genome Biol. *9*, R7–R22.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Biological samples		
Blood samples of the olive ridley turtle	National Huidong Nature Sea Turtle Reserve, Guangdong Province, China	N/A
Deposited data		
Genome sequencing data	This study	GSA with accession CRA020668
Genome assembly data	This study	GWH with accession GWHFHHW00000000.1
Software and algorithms		
chopper (v.0.9.0)	De Coster et al.[58]	https://github.com/wdecoster/nanopack
Hifiasm (v.0.15.5-r352)	Cheng et al.[59]	https://github.com/chhylp123/hifiasm
purge_dups (v.1.2.5)	Guan et al.[60]	https://github.com/dfguan/purge_dups
Yahs (v.1.2)	Zhou et al.[61]	https://github.com/c-zhou/yahs
Juicebox (v.2.15)	Durand et al.[62]	https://github.com/aidenlab/Juicebox
RepeatModeler (v.2.0.3)	Flynn et al.[63]	https://github.com/Dfam-consortium/RepeatModeler
RepeatMasker (v.4.1.4)	Chen[64]	https://github.com/rmhubley/RepeatMasker
RepBase database (v.20181026)	Jurka et al.[65]	https://www.girinst.org/server/RepBase
TRF (v.4.10.0)	Benson[66]	https://github.com/Benson-Genomics-Lab/TRF
Augustus (v.3.4.0)	Stanke et al.[67]	https://github.com/nextgenusfs/augustus
Genscan	Burge et al.[68]	http://genes.mit.edu/GENSCAN.html
Glimmerhmm (v.3.0.1)	Majoros et al.[69]	https://github.com/kblin/glimmerhmm
GeMoMa (v.1.9)	Keilwagen et al.[70]	http://www.jstacs.de/index.php/GeMoMa
EvidenceModeler (v.1.1.1)	Haas et al.[70]	https://github.com/EVidenceModeler
SwissProt	Bairoch et al.[71]	https://ftp.ebi.ac.uk/pub/databases/uniprot/knowledgebase
NR	Pruitt et al.[72]	http://www.ncbi.nlm.nih.gov/RefSeq
Diamond (v.0.9.14)	Buchfink et al.[73]	https://github.com/bbuchfink/diamond
BUSCO (v.5.4.7)	Simão et al.[74]	https://busco.ezlab.org
Circos (v.0.69.8)	Krzywinski et al.[75]	http://circos.ca/software/download/circos
MCScanX	Wang et al.[76]	https://github.com/wyp1125/MCScanX
ChiPlot	N/A	https://www.chiplot.online
MitoHiFi (v.3.2.2)	Uliano-Silva et al.[77–79]	https://github.com/marcelauliano/MitoHiFi
Tutools	N/A	https://www.cloudtutu.com
MUSCLE (v.3.8.31)	Edgar[80]	http://www.drive5.com/muscle
MEGA (v.10)	Kumar et al.[81]	https://www.megasoftware.net
OrthoFinder (v.2.4.0)	Emms and Kelly[82]	https://github.com/davidemms/OrthoFinder
MACSE (v.2.06)	Ranwez et al.[83]	https://github.com/ranwez/MACSE_V2_PIPELINES
Gblocks (v.0.91b)	Castresana[84]	http://molevol.cmima.csic.es/castresana/Gblocks
RaxML (v.8.2.12)	Stamatakis[85]	https://github.com/stamatak/standard-RAxML
PAML package (v.4.9e)	Yang[86]	http://abacus.gene.ucl.ac.uk/software/paml.html
TimeTree database	Kumar et al.[87]	http://timetree.org

(*Continued on next page*)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
CAFÉ (v.4.2.1)	De Bie et al.[88]	https://github.com/hahnlab/CAFÉ
Metascape (v.1.0)	Zhou et al.[89]	https://metascape.org/gp/index.html
RectChr (v.1.37)	N/A	https://github.com/BGI-shenzhen/RectChr
TOGA (v.1.0)	Kirilenko et al.[90]	https://github.com/hillerlab/TOGA
FasParser2 (v.2.13.0)	Sun[91]	https://github.com/Sun-Yanbo/FasParser
AlphaFold2	Jumper et al.[92]	https://alphafold.com
SnpEff (v.5.1d)	Cingolani et al.[93]	https://github.com/pcingola/SnpEff
BWA (v.0.7.17)	Li and Durbin[94]	https://github.com/lh3/bwa
SAMtools (v.1.6)	Li et al.[95]	https://github.com/samtools/samtools
BCFtools (v.1.17)	Danecek et al.[96]	https://github.com/samtools/bcftools
PSMC (v.0.6.5)	Li and Durbin[97]	https://github.com/lh3/psmc
easy_psmc_plot	N/A	https://github.com/shengxinzhuan/easy_psmc_plot

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

A sub-adult olive ridley turtle with a fibroadenoma was rescued in 2021 near the National Huidong Nature Sea Turtle Reserve, Guangdong Province, China. After a period of care, the individual is now able to feed and move normally. Approximately 10 mL of fresh blood was collected from its jugular vein and rapidly frozen for further analysis. All experiment procedures including the blood sample collection of olive ridley turtle were conducted following the approval of the Animal Experiment Ethics Committee in the Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), China.

METHOD DETAILS

Genomic DNA extraction

The genomic DNA of blood was extracted by the CTAB method (Cetyl Tri-methyl Ammonium Bromide) and purified with Grandomics Genomic kit (Wuhan, China). The quality of genomic DNA was assessed using 1% agarose gels, Qubit® 4.0 Fluorometer (Invitrogen, USA), and NanoDrop™ One UV-Vis spectrophotometer (Thermo Fisher Scientific, USA), of which OD260/280 ranging from 1.8 to 2.0 and OD 260/230 is between 2.0∼2.2.

Library preparation and genomic DNA sequencing

The high-quality genomic DNA was used for PacBio single molecule real-time (SMRT) library preparation. Briefly, the genomic DNA sample was sheared by g-TUBEs (Covaris, USA) in 15kb fragments. Single-strand overhangs were removed, and DNA fragments were damage repaired, end repaired and A-tailing. Then, the fragments ligated with the hairpin adaptor for PacBio sequencing. Genome sequencing was performed on a PacBio Sequel II instrument with Sequencing Primer V5 and Sequel II Binding Kit 2.2 in Grandomics (Wuhan, China).

Hi-C library preparation and sequencing

The experimental procedure of Hi-C library preparation was carried out according to the method described previously.[98] Briefly, crosslinking the blood sample with 1% formaldehyde, digesting by the DNA New England Biolabs (NEB) buffer 3.1 and DPN II enzyme, making the DNA ends with biotin-14-dCTP, ligating the blunt-end fragments, and shearing the DNA into 400 bp fragments by sonication. Finally, the constructed Hi-C libraries were sequenced on Illumina Novaseq platform.

Genome assembly

The raw PacBio HiFi reads were first filtered using chopper (v.0.9.0)[58] with parameters of '–minlength 5000 –maxlength 25000'. The filtered high-quality HiFi reads were then assembled into contigs using Hifiasm (v.0.15.5-r352)[59] with default parameters. Next, the purge_dups (v.1.2.5)[60] was used to remove potential duplication. The Hi-C reads data was used to anchor contigs into chromosomes with the Yahs pipeline (v.1.2).[61] Finally, we used the Juicebox (v.2.15)[62] to correct assembly errors and manually adjusted the position of the scaffolds based on the Hi-C heatmaps.

Repeat annotation

De novo and homology-based prediction methods were used to annotate the repeat sequences of the *L. olivacea* genome. For *de novo* prediction, RepeatModeler (v.2.0.3)[63] was used to construct a *de novo* repeat library and RepeatMasker (v.4.1.4)[64] was used to

align against this library for searching transposon elements. For homologue-based prediction, RepeatModeler and RepeatMasker were used to predict the homologous transposon elements based on RepBase database (v.20181026).[65] TRF (v.4.10.0)[66] was used to predict the tandem repeats with default parameter.

Gene annotation

Base on repeat soft-masked genome, we performed protein-coding genes prediction using a combined method of *de novo* and homology-based prediction. First, Augustus (v.3.4.0),[67] Genscan (http://genes.mit.edu/GENSCAN.html),[68] Glimmerhmm (v.3.0.1)[69] were used to predict protein-coding genes in *de novo* prediction. Second, for the homology-based prediction, genome annotation Gff3 files and genome sequences files of closely related species including *C. mydas*, *D. coriacea*, *C. caretta*, *Mauremys mutica*, *Podarcis muralis*, *Sceloporus undulatus*, *Trachemys scripta elegans* were downloaded and used to protein gene prediction using GeMoMa (v.1.9).[70] Finally, all results obtained from both methods were integrated into the final gene set using EvidenceModeler (v.1.1.1).[99] The gene functions were further annotated by the SwissProt[71] and NR[72] databases using Diamond (v.0.9.14).[73]

Genome assessment and synteny

The annotation completeness of genome assembly was assessed based on the protein-coding gene set using BUSCO (v.5.4.7)[74] with sauropsida_odb10 gene sets. Circos (v.0.69.8)[75] was used to visualize the genome features. We also performed the genome synteny analysis of olive ridley turtle and its closely related species (*D. coriacea*, *C. mydas*, *E. imbricata*, and *C. caretta*) using MCScanX (https://github.com/wyp1125/MCScanX)[76] and we further visualized the collinearity region and detected the potential chromosome fusion and fission events using ChiPlot (https://www.chiplot.online/).

Mitochondrial genome assembly and haplotype identification

Using HiFi sequencing data, we assembled the complete mitochondrial genome of the olive ridley turtle using the MitoHiFi (v.3.2.2)[77–79] pipeline with default parameters. The mitochondrial genome circle-map was further visualized on Tutools platform (https://www.cloudtutu.com). Additionally, 97 olive ridley mitochondrial control region sequences (Figure S2) were retrieved from NCBI nucleotide database (https://www.ncbi.nlm.nih.gov/nuccore). All mitochondrial sequences were aligned using MUSCLE (v.3.8.31),[80] and the resulting multiple sequence alignments were manually trimmed to 787 bp D-loop regions. These trimmed sequences were then used to construct a neighbor-joining tree in MEGA (v.10)[81] for haplotype identification.

Gene family identification and phylogeny tree construction

We downloaded the protein coding sequences of 15 turtles (*Chelonoidis abingdonii*, GCF_003597395.1; *Chrysemys picta*, GCF_000241765.5; *C. mydas*, GCF_015237465.2; *C. caretta*, GCF_023653815.1; *D. coriacea*, GCF_009764565.3; *E. imbricata*, GCA_030012505.1; *Gopherus evgoodei*, GCF_007399415.2; *Gopherus flavomarginatus*, GCF_025201925.1; *M. mutica*, GCF_020497125.1; *Mauremys reevesii*, GCF_016161935.1; *Malaclemys terrapin*, GCF_027887155.1; *Platysternon megacephalum*, GCA_003942145.1; *Pelodiscus sinensis*, GCF_000230535.1; *Terrapene carolina triunguis*, GCF_002925995.2; *T. scripta elegans*, GCF_013100865.1) and one outgroup (*Alligator mississippiensis*, GCF_030867095.1*)* from NCBI assembly database (https://www.ncbi.nlm.nih.gov/datasets/genome/) and performed the comparative genomic analyses. The orthologous gene sets were identified from the protein-coding sequences of all 17 species using OrthoFinder (v.2.4.0)[82] pipeline with default parameters.

To reveal the *L. olivacea*'s phylogeny status in Testudinata, a phylogenetic tree was constructed based on above 7,148 single copy genes. First, multiple sequence alignment (MSA) was performed with MACSE (v.2.06)[83] at the codon level. Poorly alignment regions with gaps were removed using Gblocks (v.0.91b)[84] with parameters of '-t = c, -b5 =n'. All MSAs were then concatenated to one supergene and 4-fold degenerate synonymous sites (4dtv sites) were extracted based on this super-sequence. Subsequently, a ML phylogenetic tree was constructed based on 4dtv sites using RaxML (v.8.2.12)[85] with 1,000 bootstraps. The divergence time of each node was estimated using MCMCtree from the PAML package (v.4.9e).[86] Six calibrated divergence-time nodes (Table S11) were obtained from the TimeTree database.[87]

Expansion and contraction of gene families

Based on the time-calibrated phylogenetic tree, CAFÉ (v.4.2.1)[88] was used to investigate the significantly expanded, contracted, and rapidly evolving genes specific to the *L. olivacea* branch with the p-value < 0.01. The corresponding gene set of expanded and contracted gene families were further used for functional enrichment analysis with Metascape (v.1.0).[89] 329 expanded OR genes of *L. olivacea* were used for further phylogenetic analyses. Specifically, protein sequences of 329 OR genes and 1 outgroup protein sequences (vomeronasal type-2 receptor 1, *VMR2R1*) were first aligned using MUSCLE (v.3.8.31). The consensus alignment sequences next were used for neighbor-joining tree construction with MEGA (v.10). The location of each expanded OR genes was visualized using RectChr (v.1.3.7) (https://github.com/BGI-shenzhen/RectChr).

Positive selected genes and rapidly evolving genes

To explore the selection pressure of *L. olivacea*, the ratio of nonsynonymous substitutions to synonymous substitutions (dN/dS, ω) was estimated using CODMEL of PAML package. The branch-site model was used to test for positive selected genes (PSGs) identification with the *L. olivacea* set as foreground branch and other species as background branch. Briefly, alternative model allows

several particular sites on the foreground branch to be under positive selection ($\omega > 1$), whereas the null model assumes that sites of all branch evolve either neutrally ($\omega = 1$) or under purifying selection ($\omega < 1$). The branch model was implemented to test for rapidly evolving genes (REGs) identification with the *L. olivacea* set as foreground branch and other species as background branch. Specifically, the null model assuming all branches with the same evolving rate while the alternative model allowing the foreground (*L. olivacea*) and background branches being with different evolving rates. For both selection analyses, a likelihood ratio test (LRT) was performed to test whether the alternative model was significant (p-value < 0.05) as compared with null model. The p-value generated from LRT was further corrected by FDR (adjusted p-value < 0.05).

Pseudogene

We used the TOGA (v.1.0)[90] pipeline to identify pseudogenes in the *L. olivacea* genome with the reference of *C. mydas* (GCF_015237465.2). Those genes classified as ''lost (L)'' were considered the potential pseudogenes. The functional enrichment analysis of pseudogenes were further performed using Metascape.

Identification of specific amino acid substitutions

On the basis of 7,148 single-copy orthologous genes, we used FasParser2 (v.2.13.0)[91] to detect the specific amino acid site changes of three sea turtles (foreground group) compared with other 12 species (background group). Specifically, the specific amino acid substitutions should be completely identical within both groups and to differ between groups. The protein 3D structures of candidate genes were predicted and plotted using AlphaFold2.[92]

Heterozygosity and genetic load

We performed snp-calling pipeline and obtained high quality SNP/INDEL set followed the methods of Yang et al. (2022)[19] based on the reference genome and HiFi long reads data of *L. olivacea*. The SNP/INDEL density map was then plotted using RectChr. We calculated the heterozygosity from four different regions (whole-genome, repeat, non-exon, and exon region) based on the gene annotation file (GFF).

We further annotated the SNP/INDEL variant dataset using SnpEff (v.5.1d)[93] and obtained four types of coding variants including high-impact, moderate-impact, low-impact, and modifier-impact variants. To explore the potential deleterious variants whether affect the evolutionary adaptation of *L. olivacea*, we further checked the overlapping genes of high-impact mutations and performed the corresponding functional enrichment analysis.

Demographic history simulation

On the basis of the reference genome and HiFi long reads data of *L. olivacea,* whole genome's consensus sequences were obtained using BWA (v.0.7.17),[94] SAMtools (v.1.6)[95] and BCFtools (v.1.17).[96] We then used PSMC (v.0.6.5)[97] to reconstruct the demographic history of *L. olivacea* with 100 bootstrap replicates using the parameters of ''-N25 -t15 -r5 -b -p "4+25*2+4+6"''. To compare whether Pacific and Atlantic olive ridleys show similar variations in past effective population size, one olive ridley individual whole-genome sequencing data (SRR15328381)[25] from the Atlantic Ocean was retrieved from NCBI Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra) and analyzed with same pipeline as above. The generation time (g = 20) and mutation rate (u = 7.9e-9) were followed the previous study.[25] Finally, the result psmc files were visualized using easy_psmc_plot (https://github.com/shengxinzhuan/easy_psmc_plot).

QUANTIFICATION AND STATISTICAL ANALYSIS

Quantification and statistical analysis used in this study including the genome assembly, comparative genome analysis and genetic diversity analysis can be found in the method details.