

RESEARCH ARTICLE SUMMARY

EVOLUTION

Spatially resolved single-cell atlas reveals the macroevolutionary trajectory of animal hearts

Shanshan Lian†, Wei Wang†, Yaoxin Liu†, Meina Lu†, Wentao Han†, Jiaxin Luo†, Xuejiao Zhang†, Yue Pang†, Yuru Li† *et al.*



Full article and list of author affiliations:
<https://doi.org/10.1126/science.adw0855>

INTRODUCTION: The heart represents a key evolutionary innovation in the rise of complex animal life. Yet the cellular and molecular basis of heart diversity, its deep evolutionary origins, and the modes of convergent versus divergent evolution across phyla are largely unresolved. Addressing these fundamental questions is critical not only for understanding metazoan organ evolution but also for providing evolutionary context to human cardiac biology and disease. However, existing cardiac research leans narrowly on humans and a limited set of classic model organisms, leaving key groups, such as mollusks, arthropods, and basal chordates, severely underrepresented, which hinders a holistic view of cardiac origins and evolutionary principles. A comprehensive understanding of these unresolved questions therefore depends on integrating multimodal data from evolutionarily diverse species, an approach that remains challenging when applied to studies that involve nonmodel species across deep evolutionary timescales.

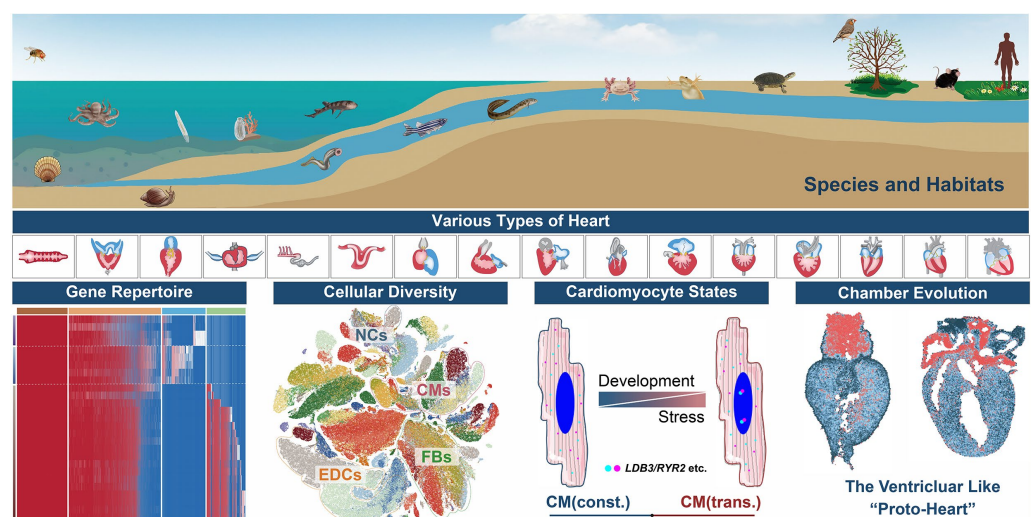
RATIONALE: To address these gaps, we integrated comparative genomics with bulk, single-cell, and spatial transcriptomics to profile adult hearts across a broad range of evolutionarily diverse taxa. Our sampling spanned protostomes (fruit fly, horseshoe crab, shrimp, crab, scallop, oyster, snail, and octopus) and deuterostomes (amphioxus, sea squirt, lamprey, skate, bamboo shark, zebrafish, lungfish, axolotl, tropical clawed frog, *Xenopus*, gecko, turtle, red junglefowl, zebra finch, mouse, rat, tree shrew, and human). This design enabled systematic interrogation of conserved, divergent, and convergently evolved features of the heart at the molecular, cellular, and tissue architecture levels. To empower the broader community, we have established an open database, HEART (<https://db.genomics.cn/stomics/heart/>), which offers both a resource and an analytical framework for cross-species exploration of cardiac molecular profiles.

RESULTS: Our analysis revealed that animal hearts share more than 60% of gene modules and a conserved cellular foundation. Notably, certain atrial and ventricular characteristics were inherited by both protostomes and deuterostomes, pointing to an ancient origin possibly traceable to the last common bilaterian ancestor. The early expansion of cardiac genes in chordates aligns with early chordate whole-genome duplication events. The diversity in cardiac architecture reflects lineage-specific genetic innovations, flexible cell-type combinations, and molecular-level functional adaptations, with both gene loss and gain events shaping cardiac evolution. Leveraging spatial transcriptomics, we identified an evolutionary conserved cardiomyocyte state that correlates with cardiac development and stress. Finally, our evidence suggests that the ancestral “proto-heart” consisted primarily of ventricular-like myocardium.

CONCLUSION: Our work reveals the macroevolutionary trajectory of animal hearts and illuminates the ancestral origin of their pumping function. By tracing distinct molecular trajectories across major cardiac cell types and identifying conserved chamber features shared by protostomes and deuterostomes, this study offers a framework for decoding heart evolution across bilaterians. Collectively, this work provides an open resource and a scalable research paradigm for investigating the evolutionary mechanisms of complex organs, exemplified by the heart, across deep evolutionary timescales. □

*Corresponding authors: Long Zhao (zhaolong@ouc.edu.cn); Wei Wang (ww8898@ouc.edu.cn); Ji-Feng Fei (jifengfei@gdph.org.cn); Shi Wang (swang@ouc.edu.cn); Silin Pan (pansilin@qdu.edu.cn); Qun Liu (liuqun@genomics.cn) †These authors contributed equally to this work. Cite this article as S. Lian *et al.*, *Science* 393, eadw0855 (2026). DOI: 10.1126/science.adw0855

Bilaterian heart evolution. Animal hearts evolved diverse anatomical structures and physiological functions to adapt to their habitats. A multiomics analysis across chordates, arthropods, and mollusks revealed that bilaterian hearts share a core gene repertoire, combined with evolutionary gained genes, driving cardiac cellular diversity and the emergence of atria from a ventricular foundation, along with an evolutionarily conserved cardiomyocyte state in response to cardiac development and stress. CM, cardiomyocyte; CM(const.), constitutive cardiomyocyte state; CM(trans.), transient cardiomyocyte state; EDC, endothelial cell; FB, fibroblast; NC, neural cell.



EVOLUTION

Spatially resolved single-cell atlas reveals the macroevolutionary trajectory of animal hearts

Shanshan Lian^{1†}, Wei Wang^{1*†}, Yaoxin Liu^{2†}, Meina Lu^{3†}, Wentao Han^{1†}, Jiaxin Luo^{1†}, Xuejiao Zhang^{3†}, Yue Pang^{4†}, Yuru Li^{1†}, Yufeng Li^{1,5}, Xin Dang⁶, Peiyan Wang⁷, Yating Qin^{8,9,10}, Ning Tan^{2,11}, Fang Zhang^{3,12}, Jin Zhang⁴, Shuai Gao¹³, Runyi Zhang^{1,5}, Lidong Guo⁸, Lei Jiang^{2,11}, Pengchong Zhang^{3,14}, Yuehuan Zhang¹⁵, Zhe Qu⁵, Shanshan Pan^{8,9}, Yue Zhang^{3,12}, Naina Hu^{1,5}, Yadong Chen^{8,9}, Yun Cao^{3,12}, Wenbin Li⁴, Xiaomei Chen^{1,5}, Junjie Shi^{8,9}, Shuaiqi Yang³, Yiran Li^{1,5}, Qiuyun Song^{8,9}, Lisui Bao³, Yuli Li^{1,5}, Xiawei Liu⁸, Jing Wang^{1,5}, Lingling Zhang^{1,5}, Rui Li⁸, Xin Yu⁸, Tao Yang^{16,17}, Jing Chen^{16,17}, Ling Li^{16,17}, Mengqi Zhang¹⁸, Lei Li⁸, Ruihua Zhang⁸, Hongyan Li³, Ying Gu¹⁹, Xun Xu¹⁹, Huanming Yang^{19,20,21,22}, Zhenmin Bao^{1,23,24,25}, C. Geoffrey Burns²⁶, Caroline E. Burns²⁶, Robert G. Kelly²⁷, Jianxun Wang⁷, Yang Zhao⁶, Ying Su^{3,12}, Guangyi Fan^{8,9,28,29}, Qun Liu^{8,9,10*}, Shi Wang^{1,5,23,24,25*}, Ji-Feng Fei^{2,30,31*}, Silin Pan^{13*}, Long Zhao^{3,14*}

Animal hearts display diverse anatomical structures during adaptive evolution. Here, we present a multiomics atlas of adult hearts from 27 species across chordates, arthropods, and mollusks. Joint analysis indicates that Bilateria hearts share a core gene repertoire, taking a stepwise “add-on” approach as a universal evolutionary strategy. The “proto-heart” is populated by key cell types, including cardiomyocytes, fibroblasts, endothelial cells, and neural cells, which maintained core signatures while evolving with shifts in living environments and corresponding adaptations in the cardiovascular system. Additionally, we reveal an evolutionarily conserved cardiomyocyte state dynamic potentially linked to cardiac development and stress responses. Finally, we identify a common molecular program underpinning chamber evolution from a ventricular foundation. This work establishes a resource for understanding the intrinsic mechanisms of heart evolution.

The heart represents one of the most important evolutionary innovations in the animal kingdom (1, 2). Fossil evidence indicates that a functional cardiovascular system had already emerged by the Early Cambrian period (3), coinciding with the Neoproterozoic Oxygenation Event and the Cambrian Explosion, a period of metazoan diversification that likely marks the heart’s first appearance (1, 4). Arthropods, mollusks, and chordates—the three dominant bilaterian lineages—diverged more than 500 million years ago, yet each evolved a heart specifically adapted to its ecological niche (5–8). Despite its ancient origin, how the heart evolved across major animal lineages, persisted through five mass extinctions, and adapted to diverse environments remains unresolved. Though animal hearts exhibit marked diversity in anatomical structure and physiological function, the underlying cellular and molecular basis remains poorly understood, as do the heart’s deep evolutionary origins and the extent of convergent or divergent evolution across phyla.

To decipher the ancestral molecular blueprint of the heart and elucidate the genetic programs underlying its variation, it is essential to broaden comparative analyses by introducing more species at key evolutionary nodes. In this study, we leveraged comparative genomics and single-cell and spatial transcriptomics to map the cellular architecture of adult hearts from 27 species across broad evolutionary taxa,

including previously overlooked arthropod and molluscan groups, to investigate the heart’s origin, evolution, and environmental adaptations. We have also established an open resource and analytical framework for exploring cardiac molecular profiles across species. Our analyses identify a conserved core gene set traceable to the last common bilaterian ancestor, reflected in cellular conservation, particularly among cardiomyocytes (CMs). We further reveal an evolutionarily conserved transient state of CMs, which is possibly correlated with oxygen conditions. Finally, we infer that the ancestral tubular heart consisted largely of ventricular-like myocardium and that increased cardiac complexity, including chamber emergence, arose through further specialization of these CMs, along with integration of nonmyocardial cells during evolution.

Results

A multiomics data resource for heart evolution across bilaterian species

Present understanding of heart evolution is limited by a narrow phylogenetic scope, with more than 95% of animal lineages, particularly nonmodel species, remaining poorly investigated (9) (fig. S1A). To address this gap, we established a multiomics resource spanning 27 bilaterian species across protostomes (nine species) and deuterostomes (18 species), covering diverse heart forms—from contracting vessels to chambered structures—and habitats, from aquatic to terrestrial environments (Fig. 1A). The dataset integrates comparative genomics with bulk, single-cell, and spatial transcriptomics (Fig. 1B and table S1).

Selected species included protostomes with a dorsal vessel (fruit fly), a tubular heart (horseshoe crab, white shrimp, and swimming crab), two-chambered heart (scallop, oyster, giant clams, and snail), and a three-chambered heart (octopus). Deuterostomes encompass species with contracting vessels (amphioxus), a tubular heart (sea squirt), a two-chambered heart (lamprey, skates, bamboo shark, and zebrafish), a three-chambered heart (lungfish, axolotl, tropical clawed frog, *Xenopus*, gecko, and turtle), and a four-chambered heart (chicken, zebra finch, mouse, rat, tree shrew, and human) (10–18) (Fig. 1, A and C, and table S1). This diverse set of species provided a diverse sample of animal hearts, representing various structural and functional characteristics in evolution.

To profile the organ-level transcriptome, we generated bulk RNA sequencing (RNA-seq) data from 23 species (Fig. 1B and table S1). Using single-cell RNA-seq (scRNA-seq), we obtained 399,483 cells from 14 species, ranging from 9438 to 60,471 cells for individual species (Fig. 1D and table S1). Subsequently, we generated high-resolution spatial transcriptomics data from tissue slices of 13 species and performed histological characterizations of cardiac spatial structures and CM sarcomere organization (fig. S1, B and C). We obtained highly accurate gene expression data, with a total of 685,023 spatial spots, ranging from 11,620 to 77,642 spots per individual animal heart (Fig. 1D, fig. S2, and table S1).

To empower the broader research community, we established an online database, HEART (Heart Evolution Atlas: A Repertoire across Tree-of-life; <https://db.genomics.cn/stomics/heart/>), which is a multi-species resource integrating single-cell and spatial transcriptomics data. This platform enables exploration of cardiac origin and evolution at the molecular, cellular, and organ levels.

Orchestration of a core gene repertoire during the evolution of bilaterian hearts

Animal hearts exhibit notable architectural diversity yet perform the same basic function of circulatory drive. Whether this functional conservation stems from shared genetic building blocks remains unclear. Our dataset offers an opportunity to address this question.

To enable robust cross-species comparison, we improved genome annotations for seven nonmodel species (fig. S3 and table S2). We performed ortholog inference hierarchically and manually verified for key

heart-related genes, identifying 21,473 gene families that presented in at least two species (table S3). Cross-referencing with transcriptome data yielded 15,577 heart-related gene families (Fig. 2A and table S3). Notably, 3456 (22.18%; CoreBA1, commonly shared) and 6367 (40.9%; CoreBA2, commonly shared with partial loss in certain lineages) families were shared across major groups (Fig. 2A, fig. S4, and table S3). Thus, more than 60% of cardiac gene families are conserved between protostomes and deuterostomes, despite their structural divergence, suggesting inheritance from an ancestral “proto-heart” and revealing a basic genetic blueprint that has been maintained throughout bilaterian heart evolution.

Beyond the conserved genetic core, we identified lineage-specific “add-on” gene families: 3086 in protostomes (CorePT) and 2668 in deuterostomes (CoreDT) (Fig. 2A, fig. S4, and table S3), indicating that both lineages genetically diverged by independently expanding the ancestral toolkit. Notably, chordates exhibited a punctuated, leap-like pattern of genetic accretion (fig. S5A), a pattern that aligns with the whole-genome duplication (WGD) events, especially the first two rounds of WGD events that drove the emergence of novel gene families and rapid increases in genetic complexity (19) (fig. S5A and table S4). The acquisition of key developmental regulators such as *MESP/MSGN* and *TGFBR2* in basal chordates marks the integration of critical transcription factors (TFs) and signaling pathways, underscoring the key genetic transition in chordate heart formation (table S4). Collectively, these findings reveal a globally orchestrated repertoire of cardiac genes, with more than 60% of gene families tracing back to an ancient origin shared across all species studied. Lineage-specific genetic additions and family gains further contributed to the diversification of heart form and function across bilaterians.

Deep evolutionary history of animal hearts

To reconstruct the deep evolutionary history of the bilaterian heart, we adopted a transcriptome age index (TAI) analysis (20, 21) on heart-specific gene-expression profiles across arthropods, mollusks, and chordates (Fig. 2, B and C, and table S5). A pronounced peak at phylostratum 6 (ps6) is observed across three major groups, with most of the contributing genes (87 to 95%) belonging to the CoreBA1 and CoreBA2 clusters, revealing a core genetic “toolkit” established rapidly in the prototype of the Bilateria heart (Fig. 2, B and C, and table S6). Functional enrichment highlighted key cardiogenic processes, including mitochondrial respiration, vasculature development, and ventricular cardiac muscle tissue morphogenesis, which involve genes such as *TGFBR*, *BAMBI*, *WIF*, *MYOD* (fig. S5B and table S6). At ps6, we observed the presence of *DGCR2*, a gene located in the 22q11.2 locus of the human genome and linked to major congenital heart disease seen in 22q11.2 deletion syndrome patients (22–24), which suggests a pathogenic mechanism traceable to the ancestral proto-heart.

Larger peaks subsequently emerged in each lineage and may contribute to heart morphological diversification, for example, ps9–ps10 in arthropods, ps8 and ps10–ps11 in mollusks, and ps9 in chordates

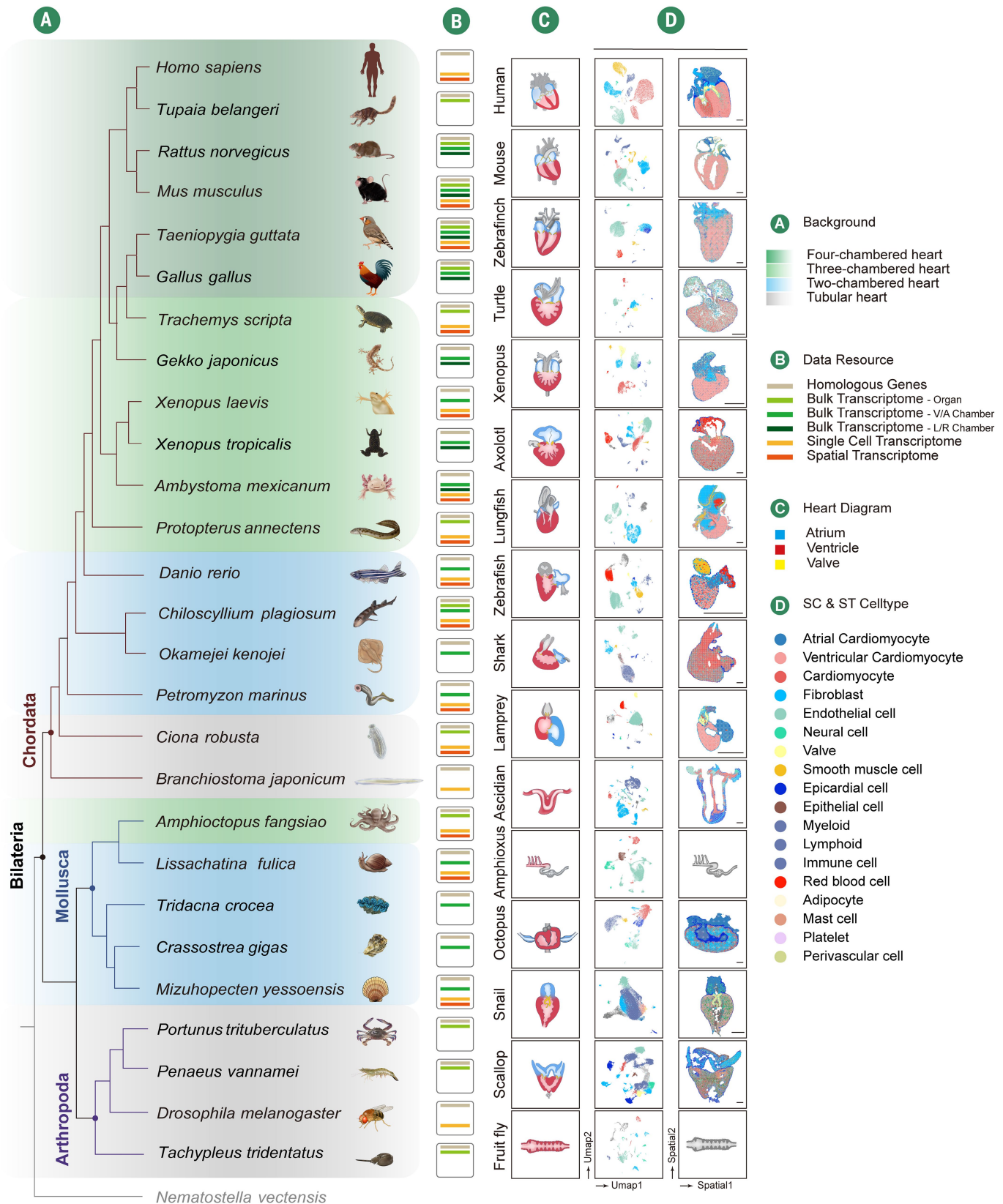
(Fig. 2B). This suggests that protostomes and deuterostomes reshaped their hearts along distinct evolutionary trajectories at different temporal points. In chordates, genes with high TAI contributions at ps9 were enriched for functions involved in heart processes, including G protein-coupled receptor signaling (25–27), neuropeptide signaling (28), response to hormone, regulation of heart rate, contraction, and chemical synaptic transmission (fig. S5B and table S6). Notably, 89% of these ps9 genes in chordates belong to the deuterostome-specific CoreDT cluster (Fig. 2C). These findings suggest a functional transition from a simple contractile tube to a signal-responsive, self-regulating organ, potentially driving lineage-specific heart diversification.

To investigate how heart complexity evolved along independent lineages, we analyzed gene gain and loss at key evolutionary nodes. Clear gene-gain signals were detected in Protostomia (for example, scallop with two-chambered heart, octopus with three-chambered heart) and Deuterostomia (for example, lamprey with two-chambered heart, lungfish with three-chambered heart, mouse with four-chambered heart) (Fig. 2D and table S7). This conserved pattern suggests that gene acquisition served as one of the driving forces during major morphological and functional transitions of the heart. We also compared the spatial expression patterns of genes from CoreBA1 and the gene-gain cluster (tables S3 and S7). Notably, in both protostomes and deuterostomes, we observed complementary spatial expression patterns between CoreBA1 and species-specific gained genes (Fig. 2E). This suggests that the conserved CoreBA1 genes likely established the primitive ventricular structure, which is also supported by the Gene Ontology (GO) enrichment of ps6 genes (fig. S5B), whereas the newly acquired genes contribute more to novel structural development and functional maturation of existing cardiac components.

Cross-species cellular diversity in heart evolution

Beyond genetic frameworks, whether diverse heart architectures arise primarily from cellular diversification or retain a common cellular foundation remains an open question. To explore cardiac evolution at the cellular level, we first characterized the cell-type composition of adult hearts across species. We identified 11 major cell types from single-cell and spatial transcriptomic data among individual 16 species, including 363,318 CMs (16/16 species), 261,580 endothelial cells (EDCs, 16/16), 161,460 fibroblasts (FBs, 16/16), 55,290 epicardial cells (EPCs, 13/16), 28,706 smooth muscle cells (SMCs, 12/16), 21,564 neural cells (NCs, 16/16), and 22,837 valve cells (9/16) (figs. S6 to S8 and tables S8 to S10). CMs, EDCs, and FBs constituted the three most abundant cell types, whereas the NCs, though relatively few per species, were consistently detected in both protostomes and deuterostomes (fig. S8B). The extensive presence of EDCs, FBs, CMs, and NCs across species is consistent with their essential roles in cardiac function and also suggests an early emergence during bilaterian proto-heart evolution, which is also supported by ancestral state reconstruction (fig. S9 and table S11). The preponderance of EDCs, FBs, and CMs is consistent with recent single-cell omics studies in fish and mammalian hearts

¹Fang Zongxi Center for Marine Evo-Devo and MOE Key Laboratory of Marine Genetics and Breeding, College of Marine Life Sciences, Ocean University of China, Qingdao, China. ²Guangdong Cardiovascular Institute, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Guangzhou, China. ³Key Laboratory of Evolution and Marine Biodiversity (Ministry of Education) and Institute of Evolution and Marine Biodiversity, Ocean University of China, Qingdao, China. ⁴College of Life Science, Liaoning Normal University, Dalian, China. ⁵Laboratory for Marine Biology and Biotechnology, Marine Science and Technology Center, Qingdao, China. ⁶State Key Laboratory of Natural and Biomimetic Drugs, Ministry of Education Key Laboratory of Cell Proliferation and Differentiation, Institute of Advanced Clinical Medicine, Center for Life Sciences, College of Future Technology, Peking University, Beijing, China. ⁷School of Basic Medicine, Qingdao University, Qingdao, China. ⁸BGI Research, Qingdao, China. ⁹Qingdao Key Laboratory of Marine Genomics, BGI Research, Qingdao, China. ¹⁰Department of Biology, University of Copenhagen, Copenhagen, Denmark. ¹¹Department of Cardiology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, Guangzhou, China. ¹²College of Marine Life Sciences, Ocean University of China, Qingdao, China. ¹³Women and Children's Hospital of Qingdao University, Qingdao, China. ¹⁴College of Fisheries, Ocean University of China, Qingdao, China. ¹⁵South China Sea Institute of Oceanology, Chinese Academy of Science, Guangzhou, China. ¹⁶China National GeneBank, BGI Research, Shenzhen, China. ¹⁷Guangdong Provincial Genomics Data Center, BGI Research, Shenzhen, China. ¹⁸MGI Tech, Shenzhen, China. ¹⁹BGI Research, Shenzhen, China. ²⁰College of Life Sciences, University of Chinese Academy of Sciences, Beijing, China. ²¹HIM-BGI Omics Center, Hangzhou Institute of Medicine (HIM), Chinese Academy of Sciences, Hangzhou, China. ²²James D. Watson Institute of Genome Sciences, Hangzhou, China. ²³Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou, China. ²⁴Key Laboratory of Tropical Aquatic Germplasm of Hainan Province, Sanya Oceanographic Institution, Ocean University of China, Sanya, China. ²⁵Qingdao Institute of Blue Seed Industry, Qingdao, China. ²⁶Division of Basic and Translational Cardiovascular Research, Department of Cardiology, Boston Children's Hospital, Harvard Medical School, Boston, MA, USA. ²⁷Aix-Marseille Université, CNRS UMR 7288, IBDM, Turing Centre for Living Systems, Marseille, France. ²⁸BGI Research, Sanya, China. ²⁹BGI-Shenzhen, Shenzhen, China. ³⁰Department of Pathology, Guangdong Provincial People's Hospital, Guangdong Academy of Medical Sciences, Southern Medical University, Guangzhou, China. ³¹The Innovation Centre of Ministry of Education for Development and Diseases, School of Medicine, South China University of Technology, Guangzhou, China. *Corresponding author. Email: zhaolong@ouc.edu.cn (L.Z.); ww8898@ouc.edu.cn (W.W.); jifengfei@gdph.org.cn (J.-F.F.); swang@ouc.edu.cn (S.W.); pansilin@qdu.edu.cn (S.P.); liuqun@genomics.cn (Q.L.) †These authors contributed equally to this work.



(11, 27, 29–34). Additionally, SMCs were detected only in chordates from sea squirt onward and in octopus among protostomes (fig. S8B). Epicardium is a layer of mesothelial tissue surrounding the heart that was previously thought to be unique to all vertebrates. Our data demonstrate that it may have been incorporated during the evolution of the chambered heart in vertebrates as well as mollusks (fig. S8B).

To assess cellular conservation, we integrated 479,483 cells from 16 species together using SATURN (35) (Fig. 3, A and B; fig. S10; and table S10). SAMap analysis and ancestral state reconstruction revealed that CMs are conservatively present across protostomes and deuterostomes (figs. S9 and S11A and table S11), suggesting their establishment in the ancestral bilaterian proto-heart as fundamental functional components. By contrast, EDCs, a major cardiac cell type (11, 36, 37), along with FBs and SMCs, are broadly present among vertebrates (figs. S9 and S11B and table S11), indicating evolutionary divergence of cardiac cellular lineages between protostome and deuterostome.

Given the high conservation of CMs and FBs across chordate hearts, we sought to identify the core gene sets defining these cell types to elucidate the molecular basis of their evolutionary stability. We integrated single-cell and spatial transcriptomic data to detect differentially expressed genes (DEGs), refined through weighted gene coexpression network analysis (WGCNA). In our analysis, the genes expressed in more than half of species within either protostome or deuterostome clades were designated as universal core genes for a particular cell type across species. Ultimately, we identified 97 CM-specific and 92 FB-specific core genes and their protein interaction modules (Fig. 3C, figs. S12 and S13, and table S12).

Cardiac contraction, a universal heart function driven by CMs, is molecularly supported by these core genes. Among the 97 CM-specific genes, mitochondrial-associated genes were prominently enriched (31/97), forming the top protein-protein interaction module (Fig. 3C, figs. S12A and S13A, and table S12) and underscoring the essential role of mitochondrial adenosine triphosphate (ATP) production in sustaining contraction. Sarcomere-related genes were also enriched, including cardiac muscle myosin, titin, myosin-binding protein, myomesin, troponin, tropomyosin, and actin (*MYH6/7/7B*, *MYL*, *TTN*, *MYBPC*, *MYOM*, *TNNT*, *ACTIN*) as CM signature genes, along with *SPTAN* and *SPTBN*, which are related to spectrin-based muscle cell membrane cytoskeleton. Although the core contractile machinery is largely conserved, certain components exhibit lineage-specific expression. For example, cardiac troponin T (*TNNT*) is highly expressed in CMs of 15 of 16 species except amphioxus, whereas the expression of troponin C (*TNNC*) appears generally restricted to deuterostomes (Fig. 3C and fig. S12A). In vertebrates, the inhibitory troponin complex subunit gene *TNNI1* is primarily expressed in lamprey and teleost fish, whereas *TNNI3* is highly expressed in amphibians and amniotes; notably, lungfish express both genes (38, 39) (Fig. 3, C and D; fig. S12A; and table S12). Taken together, these results demonstrate that diverse heart structures share a common cellular and molecular foundation, with lineage-specific genetic add-ons and flexible cell-type combinations collectively shaping cardiac evolution and adaptation.

Cellular evolution in the heart

Cardiac function has evolved through deep coadaptation with environmental shifts and corresponding adaptations in the circulatory system. Basal chordates such as tunicates and amphioxus use peristaltic waves in a simple tubular heart suited to filter-feeding in aquatic habitats, whereas mammals and birds developed high-performance hearts with complex structures and a sustained energy supply fitted for continuous oxygen delivery. scRNA-seq from protochordates to amniotes enables cellular deciphering of these evolutionary transitions.

Despite a universal core gene set that defines CMs, their molecular profiles vary considerably across chordates. Principal components analysis (PCA) of integrated single-cell data segregated metacells into three clades: Protochordata, Anamniota (fish and amphibians, hereafter

referred to as “Anamniota”), and Amniota (Fig. 3E). PC1 underlies the functional transition from peristaltic contraction to synchronized beating, accompanied by increased energy demands in chambered hearts (Fig. 3F). Consistent with this, the PC1(–) quadrant was enriched for genes involved in respiration, mitochondrial organization, and transport, highlighting mitochondrial enhancement as a key innovation in chordate CM evolution (table S13). The low expression of *TNNC* in amphioxus CMs (Fig. 3C) possibly reflects their rudimentary contractile apparatus, whereas the coordinated emergence of *TNNI*, *ACTIN*, *MYL*, *PLN*, and other muscle system genes delineates a likely genetic roadmap for the evolution of the contractile machinery (Fig. 3G and table S13). Protochordate CM signatures persist in the PC1(+) quadrant, particularly in amphioxus (Fig. 3F), with top-enriched Hedgehog signaling components (*GLI*, *PTCH1/2*, and *HHIP*, also *SMURF1/2*) (40) (fig. S14 and table S13), suggesting its regulatory importance in early chordate CM evolution.

Conversely, the PC1(–) and PC2(–) quadrants specifically delineate Amniota species, including zebra finch, turtle, and mouse (Fig. 3E). This Amniota-specific signature is identified by intersecting Amniota-specific genes with PC2(–) quadrant genes, reflecting cardiac functional adaptations associated with the transition from aquatic to terrestrial environments, which links to the emergence of Amniota (Fig. 3F). Within this quadrant, the apelin receptor (*APLNR*), a surface marker for in vitro induced cardiac progenitors and a key regulator of cardiovascular function, is notably prominent (41). *APLNR* mediates critical physiological processes, including enhanced cardiac contractility, angiogenesis, and regulation of energy metabolism, and is considered a promising therapeutic target for cardiovascular diseases (42). Its presence here correlates with multiple enriched biological processes, particularly those related to cardiovascular development (Fig. 3G and table S13). We also observed the emergence of genes involved in fine-tuning the nitric oxide biosynthetic process in Amniota CMs (43, 44) (Fig. 3G and table S13). Nitric oxide plays essential roles in heart physiology, including the regulation of myocardial contractility and coronary blood flow (45). The higher O₂ availability on land likely provided ample substrate for nitric oxide synthesis, driving the adaptation of this signaling pathway to support more sophisticated cardiovascular regulation in amniotes (46, 47). Additionally, the coordinated emergence of *IFNAR* and *IFNGR* in Amniota CMs implies the evolution of specialized inflammatory mechanisms potentially involved in postinfarction repair and regeneration. We also identified key developmental regulators, including *BMP5/6/7*, *FGF12*, *VEGFB*, *SOX17/18*, and *NOTCH4*, revealing the molecular evolution of heart development and organism-level homeostasis regulation in amniotes (Fig. 3G, fig. S14, and table S13). Beyond CMs, we extended the same analytical approach to FBs and EDCs, leading to the extensive identification of candidate genes implicated in heart evolution (figs. S15 to S17 and tables S14 and S15). The functional characterization of these genes offers a potential direction for future research.

We next examined the evolutionary conservation of transcriptional regulation and intercellular communication in the heart. Analysis revealed deep conservation of TFs, with more than 75% of heart-expressed TFs belonging to the CoreBA1 and CoreBA2 sets (fig. S18 and table S16). Cell type-specific gene regulatory networks constructed from fruit fly, mouse, and human single-cell data identified core TFs acting as key regulators across species (table S16). This prompted us to interrogate whether these regulatory hubs are themselves preserved across heart evolution. Analysis of the CoreBA1 and CoreBA2 sets revealed that CMs are a highly conserved cell type at the transcriptional regulatory level, with key regulators such as *FOXF1/2*, *GATA4/5/6*, and *MYOD* identified across protostomes and deuterostomes. Notably, *NR4A* and *SREBF* were simultaneously predicted from CM data in human, mouse, and fruit fly. This regulatory conservation extended to other major cardiac cell types, including FBs, EDCs, EPCs, and NCs (fig. S19A and table S16). As expected, TFs in the CoreDT set were exclusive to deuterostomes across all cell types (fig. S19B and table S16).

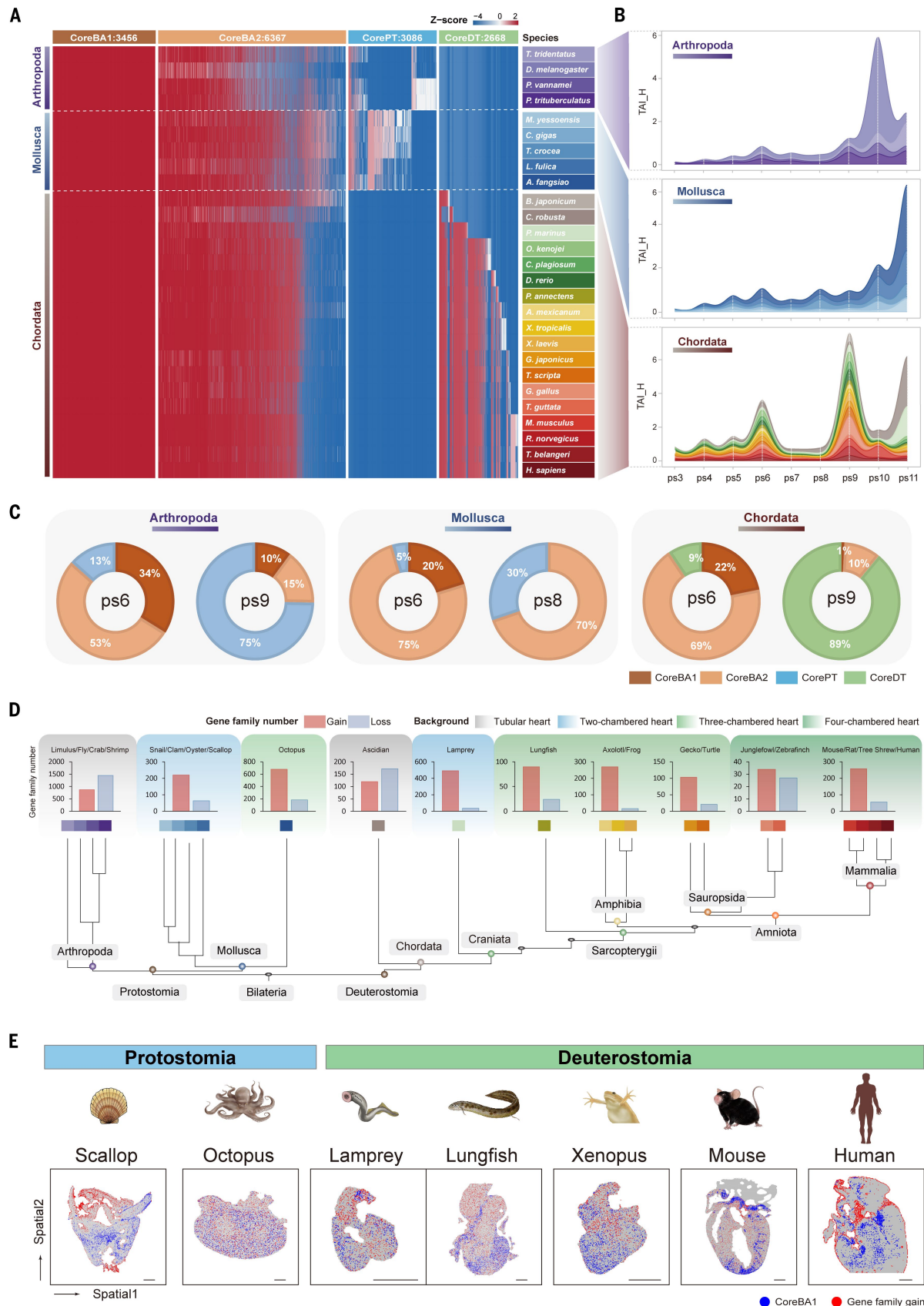


Fig. 2. Genetic basis of ancient origin and driving forces behind innovations in bilaterian cardiac structure. (A) Evolutionary clusters of orthogroups across 27 species. The cluster sizes (number of orthogroups) are labeled. The vertical color bar (left) indicates species, categorized into Arthropoda, Mollusca, and Chordata. The species identity of the color bars is annotated on the right side. The Z-score represents the distribution of orthogroups across species. CoreBA1, commonly shared; CoreBA2, commonly shared

with partial loss in certain lineage; CorePT, protostome-specific; CoreDT, deuterostome-specific. **(B)** Relative TAI contribution of each phylostratum (ps) for hearts of 27 species, grouped by Arthropoda (top), Mollusca (middle), and Chordata (bottom). Species colors correspond to those in (A). **(C)** Comparative composition of shared orthogroups across Arthropoda, Mollusca, and Chordata. Circular plots display the proportional distribution of evolutionary clusters (CoreBA1, CoreBA2, CorePT, CoreDT) for orthogroups conserved at a key phylostratum (ps6, and ps9 or ps8). **(D)** Analysis of gene family gain and loss at key evolutionary nodes marked by major innovations in cardiac chamber structure. Background colors represent different heart forms. Species colors correspond to those in (A). **(E)** Spatial expression of ancestral (CoreBA1, blue) and recently gained (species specific, red) gene families. Scale bars are 1000 μm .

We then extended our evolutionary inquiry to intercellular communication. CMs engage in bidirectional signaling with other major cardiac cell types. Our analysis in the CoreBA1 and CoreBA2 sets revealed that the ligand-receptor pairs mediating CM interactions with FBs, EDCs, EPCs, and NCs exhibit notable evolutionary conservation, surpassing even that of cell type-specific transcriptional regulators (figs. S20 and S21 and table S17). This establishes a core, conserved interactome for cardiac cellular cross-talk, which appears to be more stable over evolutionary time than the internal transcriptional programs of the individual cell types involved.

Identifying evolutionarily conserved transient states of CMs in bilaterian hearts

CMs, the principal cellular component of the heart, are a distinctive cell type that distinguishes the heart from other organs. To further explore the cellular identity conservation and/or divergence of CMs across species, we analyzed the spatial transcriptome data from 12 species. We revealed six conserved clusters (CM1 to CM6) with distinct chamber localizations: pan-cardiac (CM1 and CM2), atrial (CM3 and CM4), and ventricular (CM5 and CM6) (Fig. 4A and figs. S22 and S23A). These clusters exhibit a functional dichotomy: CM1, CM3, and CM5 express canonical contractile and metabolic genes (for example, *MYL*, *TNNT*), whereas CM2, CM4, and CM6 are enriched for ion transport and contraction-regulatory genes (for example, *CACNA1C*, *RYR2*, *LDB3*, *RBM20*) (Fig. 4B, fig. S23B, and table S18). Cross-species analysis showed that CM2, CM4, and CM6 form a tightly related group by t-distributed stochastic neighbor embedding (t-SNE) (fig. S24A), confirmed by high Pearson correlation (fig. S24B), indicating a unified, conserved transcriptional program.

We selected *Cacna1c* and *Ryr2* as markers (Fig. 4B and table S18) for the CM2, CM4, and CM6 clusters to map their spatial distribution. Although their expression in mouse heart spatial enhanced resolution omics sequencing (Stereo-seq) data lacked clear anatomical patterning (fig. S25A), subsequent validation by fluorescence in situ hybridization (FISH) revealed a distinctive transcriptional signature. Specifically, we observed a pronounced simultaneous aggregation of transcripts for both genes in the nuclei of a CM subpopulation (fig. S25, B and C). Compared with those CMs in which *Cacna1c* and *Ryr2* transcripts are only located in the cytoplasm, the cells with high expression of *Cacna1c* and *Ryr2* were identified as CM2, CM4, and CM6 cells that are widely distributed in the mouse heart. Because transcripts are usually exported to the cytoplasm for translation, we speculate that the unusual aggregation of transcripts in the nuclei of CM2, CM4, and CM6 cells may be a transient status that regulates protein synthesis, allowing CMs to adapt to the dynamic demands for gene products. Therefore, we designated CM2, CM, and CM6 as transient (trans.) CM states and, conversely, CM1, CM3, and CM5 as constitutive (const.) CM states.

To determine whether the cellular state shares the evolutionary conservation characteristics of CMs, we performed systematic quantification across multiple species. Our analysis confirmed the presence of CMs (trans.) in all 12 species examined, comprising 1.7 to 33.7% of all CMs, with a similar average prevalence of ~22% in both protostomes and deuterostomes (Fig. 4A and table S18). To functionally characterize these states, we performed GO and protein-protein interaction analysis on conserved CM genes, which identified five MCODE clusters (fig. S26, A and B, and table S19). The MCODE3 gene network was highly enriched for markers of the CMs (trans.), providing

independent network-based validation for their distinct molecular identity. To evolutionarily validate this state, we selected *RYR2* and *LDB3* (tables S18 and S20) as representative markers owing to their marked differential expression between CM (trans.) and CM (const.) cell states. FISH analysis across five representative species—mouse, axolotl, zebrafish, sea squirt, and scallop—confirmed nuclear aggregation of the transcripts of these two genes in a CM subpopulation (~8 to 28%) (Fig. 4, A and C, and table S21), recapitulating our initial finding in mouse heart (fig. S25). To provide auxiliary validation for this state, quantification of transcript abundance at either the whole-cell (Fig. 4C and table S21) or nuclear (fig. S26C and table S21) level demonstrated a divergence in expression between these two states. Furthermore, we confirmed the presence of this state in both adult and fetal human hearts (Figs. 4A and 5, A and B; fig. S26D; and tables S20 and S21), as well as in human cardiac organoids (~20% of CMs) (fig. S26, E and F, and tables S20 and S21). These results imply that nuclear aggregation of marker gene transcripts represents a defining feature of the CM (trans.) state.

Collectively, these findings suggest that subdividing the CMs into CMs (trans.) (CM2, CM4, and CM6) and CMs (const.) (CM1, CM3, and CM5) may represent a functionally relevant distinction in the heart, ensuring both stable basal function and dynamic adaptability to facilitate cardiac homeostasis. The conserved existence of the constitutive and transient states of CMs indicated the possible CM states of ancient hearts.

The dynamics of the transient CM state associated with cardiac development and stress

The CM state is closely associated with its function. We found that the proportion of CMs (trans.) was significantly higher in human fetal hearts than in adult, suggesting a postnatal decrease ($P = 0.0005$) (Figs. 4A and 5, A to C; fig. S26D; and tables S20 and S21). Subsequently, we validated this development-related decrease in mouse hearts, in which we observed an identical trend, with the proportion of CMs (trans.) being highest at embryonic day 14.5 (E14.5), decreasing at postnatal day 1 (P1), and then rapidly declining by P8, reaching the low level seen in adulthood (Fig. 5, D and E, and tables S20 and S21), suggesting a dynamic variation in the proportion of this cell state across different developmental stages.

Additionally, to determine whether the CM (trans.) state is related to the myocardial stress response, we examined the CM (trans.) states in a hypoxic environment. Under the condition of prolonged hypoxic treatment, the proportion of CMs (trans.) in adult mouse hearts indeed increased (Fig. 5, F and G, and tables S20 and S21). In the hearts of mice with myocardial infarction (MI), we quantified the proportion of CMs (trans.) in the border-zone (BZ) and remote-zone (RZ) regions of mouse hearts at 7 days post-MI. As expected, the proportion of CMs (trans.) in the BZ was significantly higher than that in the RZ and non-infarcted hearts ($P < 0.0001$) (Fig. 5, H to J, and tables S20 and S21). Furthermore, the cytoplasmic and nuclear transcripts of *Ryr2* and *Ldb3* in BZ CMs (trans.) were increased (fig. S26G). By contrast, another stress, transverse aortic constriction-induced hypertension, failed to induce an increase of this state (fig. S26, H and I, and tables S20 and S21). Collectively, our analysis reveals that CM (trans.) constitutes an evolutionarily conserved CM cell state, characterized by alterations in cellular abundance and nuclear aggregation of transcripts derived from distinct molecular networks, possibly reflecting a functional response to cardiac development and myocardial stress.

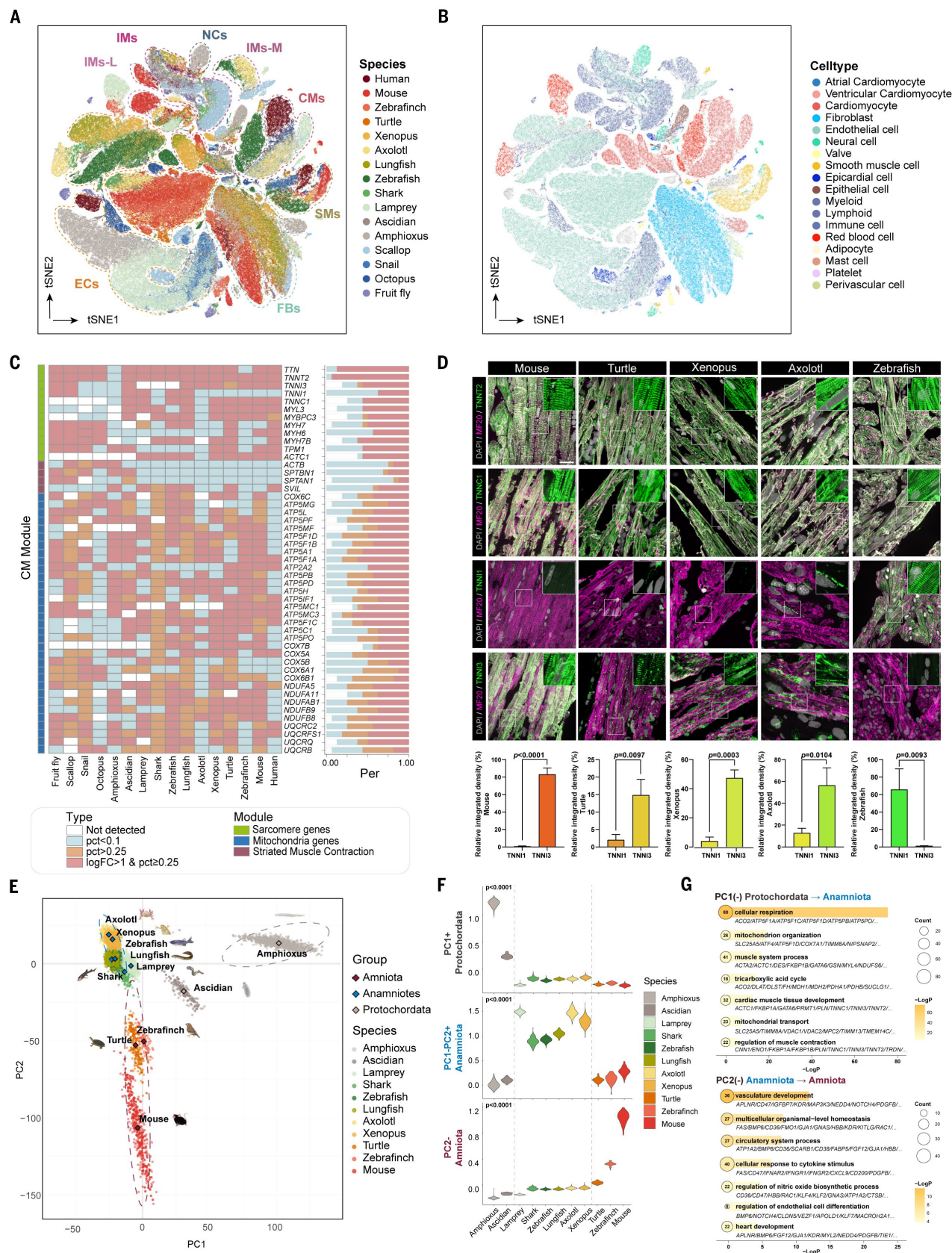


Fig. 3. Cross-species cardiac cellular landscape and evolution. (A and B) Integrated scRNA-seq analysis of hearts from 16 species. t-SNE visualization is colored by species (A) or by annotated cell types (B). (C) Cross-species core gene set for CMs. On the left is an expression-pattern heatmap with genes grouped by functional modules. The heatmap is color coded as follows: white (not detected, indicating either not found in genome or not captured or detected), light blue [percentage of cells expressing the given gene (pct) < 0.1], orange (genes with pct > 0.25 of the target cells), red (genes with pct > 0.25 of the target cells, and $\log_2FC > 1$, where FC is fold change). On the right is a

quantification of expression pattern distribution for each gene. Cell type-specific genes were defined using thresholds of $\log_2FC > 1$ and $pct > 0.25$ in single-cell transcriptome and spatial transcriptome integration clustering, and with module-to-hub association eigengene-based connectivity (kME) ≥ 0.3 in the coexpression network. For ancestral-type genes, such as a single invertebrate ortholog corresponding to multiple vertebrate paralogs, we have explicitly merged and filled in the presence calls in invertebrates. (D) Shown at the top are images of immunohistochemical staining of cardiac core genes. Shown at the bottom are bar plots that quantify the relative integrated density (%) of cardiac core genes in myocardium. $n = 3$ samples for each group. Scale bar is 20 μm . (E to G) Evolutionary trajectory of CMs. In (E), a PCA plot across species reveals subgroups of CMs (Protochordata, Anamniota, Amniota). Shown in (F) are module scores of the top 200 genes driving principal component (PC)-defined subgroups. Shown in (G) are biological processes enriched for PC-defined subgroups, ranked by statistical significance [$-\log_{10}(P \text{ value})$]. Data are represented as means $\pm$ SD. All P values were calculated using two-tailed unpaired Student's t test. $P < 0.05$ is considered statistically significant.

The convergent evolution of cardiac chambers

The independent emergence of atrial chambers from primitive tubular hearts in both protostomes and deuterostomes represents a notable case of convergent evolution. Although mammalian chamber formation has been extensively studied from developmental perspectives (48, 49) and cellular signatures of human fetal heart chambers have been recently characterized (27), the evolutionary-scale cellular signatures distinguishing atria from ventricles remain unexplored, a crucial gap in understanding chamber evolution. Using spatial omics data from maximal longitudinal heart sections across 14 species (fig. S27A) representing tubular and two-, three-, and four-chambered hearts, we analyzed cellular composition during chamber evolution. The chamber anatomical positions of atria and ventricles were defined according to acid fuchsin-orange G (AFOG) staining (figs. S1, B and C, and S27A). Our analysis revealed that the proportion of CMs in ventricles is higher than in atria (Fig. 6A, fig. S27B, and table S22). This aligns with the classical functional view of ventricles as pumping engines and atria as receiving chambers. Meanwhile, the higher proportion of non-CMs in protostome atria than in vertebrate atria was observed, which demonstrates that despite analogous chamber structures across protostomes and deuterostomes, their underlying cellular compositions and physiological implementations might differ substantially.

To investigate the molecular signatures that distinguished atria and ventricles during evolution, we integrated spatial omics data from 12 species and chamber-specific bulk RNA-seq data from 17 species (table S22). We identified 78 atrial-specific and 124 ventricular-specific genes conserved across protostomes and deuterostomes (fig. S28A and table S22), representing less than 0.1% of cardiac-expressed genes, with spatial transcriptomics validating conserved chamber-specific expression patterns in all 13 species examined (Fig. 6B and fig. S28B). PCA analysis revealed 20 key genes driving atrial-ventricular separation: 10 atrial-enriched (including *ALDH1A2*, *PITX*, *FLRT2*) and 10 in the ventricles (including *MYO18*, *LDB*, *PRDM16/MECOM*, *CACNA1C*) (Fig. 6C and table S23). *ALDH1A2* catalyzes retinoic acid synthesis, which is crucial for sino-atrial differentiation and establishes the inflow-outflow patterning of the heart (50–53), whereas *PITX* is a left-atrial marker and atrial fibrillation predisposition gene (54). These conserved signatures, although representing a small fraction of the transcriptome, constitute a highly conserved molecular core, which suggests that fundamental atrial-ventricular characteristics predate the protostome-deuterostome divergence. We therefore propose that atrial-ventricular features may have patterned the proto-heart before the evolution of distinct anatomical chambers. Consistent with this, previous research in zebrafish has demonstrated that regional physiological specialization within its undivided ventricle is governed by distinct transcriptional programs, which facilitate ventricular septation in amniotes (55, 56). We also observed that the conserved ventricular genes are predominantly associated with CMs, whereas atrial genes are largely expressed in FBs and EDCs (Fig. 6D and fig. S28C), which suggests that a core cellular consortium, consisting of CMs, FBs, and EDCs, constituted the primordial proto-heart, with their molecular patterning predating anatomical chamber segregation.

The ancestral identity of cardiac chambers

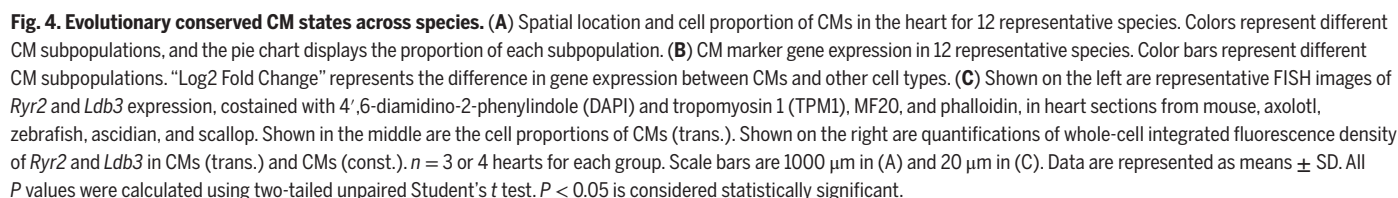
The evolutionary origin of multichambered hearts remains a subject of ongoing debate (57, 58). Key questions include whether all chambers

emerged simultaneously through patterning of a rudimentary peristaltic pump (53) or evolved sequentially through the addition of genetic modules (57) and what the evolutionary sequence of atrial and ventricular emergence might have been. Resolving these questions requires determining the ancestral identity of the proto-heart before chamber patterning.

To investigate chamber evolutionary history, we performed TAI analysis across 13 species with distinct atrial and ventricular structures. Ventricular genes consistently exhibited lower TAI values than atrial genes (Fig. 6E), indicating their more ancestral origin. Analysis of the core gene repertoire revealed distinct spatial expression patterns between CoreBA1 and CorePT in protostomes and CoreBA1 and CoreDT in deuterostomes (fig. S29A). Moreover, the CoreBA1 gene set, which is shared by protostomes and deuterostomes and may have been expressed in the ancestral bilateral proto-heart, is enriched in the ventricular region, whereas the CorePT and CoreDT gene sets, as later add-on cardiac genes in protostomes and deuterostomes, are enriched in the atrial region (fig. S29A). These findings collectively suggest that ventricular genes originated earlier in evolution, supporting the hypothesis that ventricular character emerged first, with atrial character represents a later evolutionary innovation.

The transition from the simple tubular heart of the basal chordate sea squirt to the chambered heart of the jawless lamprey illustrates a marked increase in chamber complexity. Although the intermediate forms between these structures have been lost to evolution, comparative analysis of the tubular heart with its evolutionary counterparts could still illuminate the primitive origins of the heart's chambers. Our spatial omics data reveal that ventricular-associated genes are more highly expressed in the myocardial layer of the sea squirt heart tube than in its pericardial sac, with the myocardial layer also exhibiting lower TAI values (Fig. 6F and fig. S29B). Previous lineage tracing studies in sea squirt (59) have established that both first and second heart precursors contribute to the tunicate heart, with first heart precursors biased toward the myocardial layer and second heart precursors toward the pericardial layer. These distinct heart field lineages date back to early chordate evolution (60) and predate the cardiac septation (61). We therefore propose that the multichambered vertebrate heart evolved from a tubular heart containing derivatives of both first and second heart precursors, with subsequent evolution involving increased progenitor heterogeneity and morphological rearrangement of multiple cardiac cell types.

Moreover, the observation of a stronger molecular similarity between tubular heart CMs and ventricular CMs across both major lineages further supports this model. In deuterostomes, sea squirt CMs showed a higher correlation with lamprey ventricular CMs [correlation coefficient (r) = 0.58] than atrial ones (r = 0.39). Similarly, in protostomes, fruit fly CMs were more similar to snail ventricular CMs (r = 0.55) than to atrial CMs (r = 0.22) (Fig. 6G and table S24). This ventricular bias was specific to CMs and not observed in FBs or EDCs, despite their early evolutionary appearance (Fig. 6G and table S24). Given this consistency, we propose that the heart before chamber septation was primarily composed of myocardial cells with ventricular identity, with atria emerging later through further differentiation of ventricular-like CMs, coupled with the integration and morphological rearrangement of nonmyocardial cells from more diverse progenitor sources.



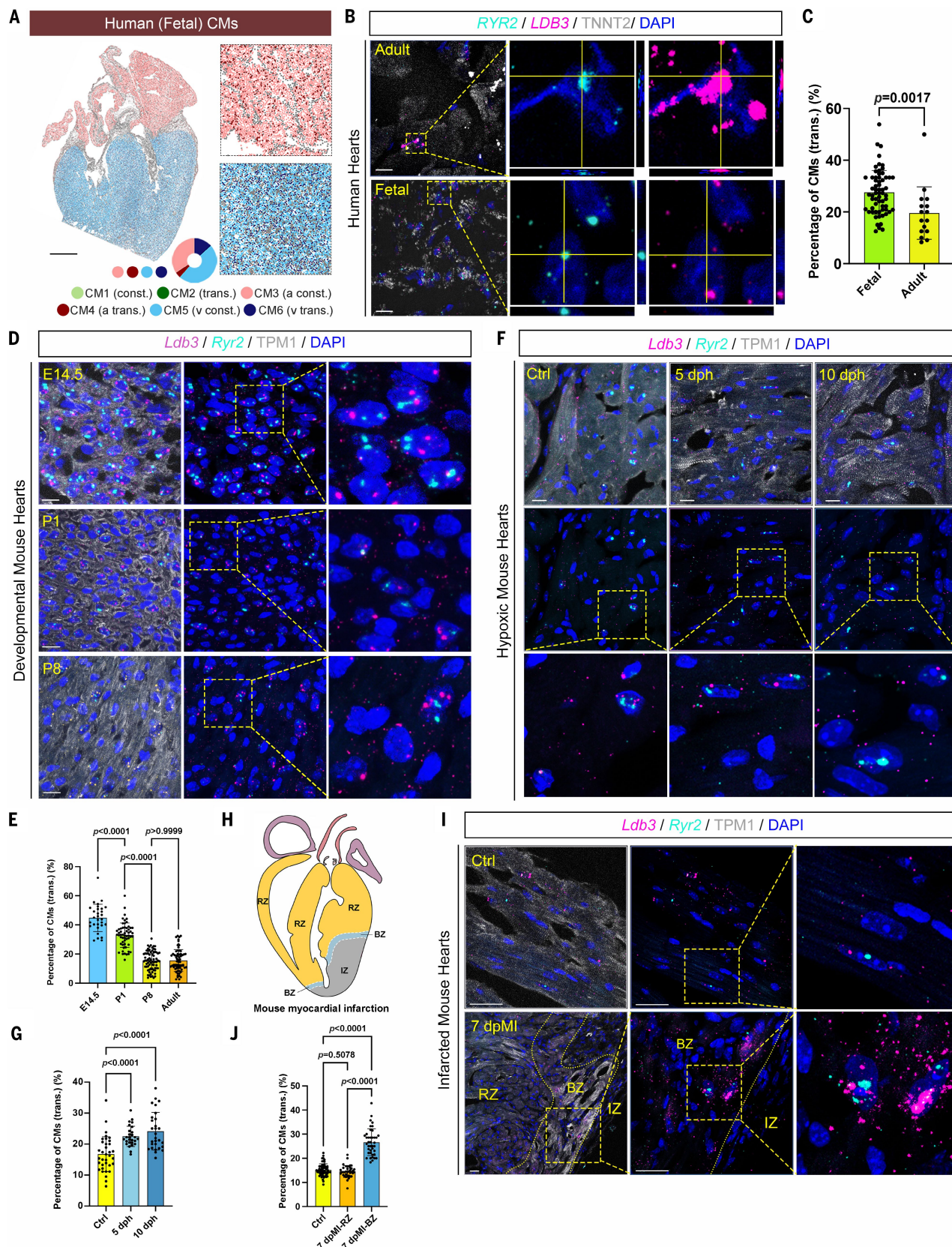


Fig. 5. The CM transient state is associated with cardiac development and stress. (A) Spatial location and cell proportion of CMs in the fetal human heart. Colors represent different cell types, and the pie chart displays the proportion of each cell type. (B) Representative FISH images showing *RYR2* and *LDB3* expression (combined with *TNNT2* and DAPI immunostaining) in adult ($n = 1$) and fetal ($n = 1$) human hearts. (C) Quantification of the percentage of CMs (trans.) in human hearts. (D) Representative FISH images for *Ryr2* and *Ldb3* with immunofluorescence costaining for DAPI and *TPM1* in embryonic (E14.5) and neonatal (P1 and P8) mouse hearts. $n = 3$ or 4 for each group. (E) Quantifica-

tion of CM (trans.) proportion in embryonic and neonatal mouse hearts. (F) Representative FISH images for *Ryr2* and *Ldb3* with immunofluorescence costaining for DAPI and TPM1 in adult hypoxic mouse hearts. $n = 3$ for each group. (G) Quantification of the percentage of CM (trans.) in adult hypoxic mouse hearts. (H) Schematic diagram of mouse heart after MI surgery. (I) Representative FISH images for *Ryr2* and *Ldb3* with immunofluorescence costaining for DAPI and TPM1 in adult infarcted mouse hearts. $n = 6$ to 8 for each group. BZ, border zone; IZ, infarcted zone; RZ, remote zone. (J) Quantification of the percentage of CMs (trans.) in adult infarcted mouse hearts. Scale bars are 1000 μm in (A); 10 μm in (B), (D), and (F); and 20 μm in (I). Data are represented as means $\pm$ SD. P values were calculated using two-tailed unpaired Student's t test (C), one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons (E), and Welch's ANOVA followed by Games-Howell multiple comparisons tests [(G) and (J)]. $P < 0.05$ is considered statistically significant.

Discussion

The heart has undergone extensive evolutionary adaptation, yet a comprehensive understanding of its genomic and cellular evolution across broad phylogenetic scales remains lacking. Existing cardiac research relies heavily on humans and model organisms such as mice and zebrafish, whereas the underrepresentation of groups like mollusks has hindered a holistic perspective on cardiac origins and evolutionary mechanisms (8, 62–64). This study addresses this gap by analyzing 27 species across arthropods, mollusks, and chordates, including previously overlooked groups and two protochordates (amphioxus and sea squirt). Our sampling covers diverse heart structures, from tubular to multichambered hearts, and adaptations to aquatic, terrestrial, and aerial environments, offering insights into the ancestral bilaterian heart and the molecular basis of its evolution.

Advances in sequencing technology now enable comparative genomics across species. However, the heart is a complex organ composed of functionally specialized cell types and thus requires multimodal data integration—including organ-specific, single-cell, and spatial transcriptomics—to resolve evolutionary changes at cellular resolution. Although single-cell atlases have been developed for organs such as the lung and liver from multiple species (65, 66), our study provides a cardiac-focused multiomics resource spanning 27 species. These data reveal evolutionary changes among cardiac cell types during circulatory system adaptation and identify conserved mechanisms of chamber evolution shared between protostomes and deuterostomes, offering initial characterization of the rudimentary bilaterian heart.

Our data delineate evolutionarily conserved cellular dynamics in the heart, illuminating the ancestral origin of its pumping function. We observed distinct molecular trajectories among major cardiac cell types during evolution. Elucidating the gene regulatory networks underlying these changes will enhance our understanding of cell type-specific functional adaptation and reveal candidate genes with potential biomedical relevance. Additionally, spatial transcriptomic data from 14 species, including three mollusks, 10 vertebrates, and one basal chordate, enabled the identification of conserved ventricular and atrial features within cellular mechanisms of chamber evolution shared across protostomes and deuterostomes.

We present an open resource and analytical framework for exploring cardiac cellular and molecular profiles across evolutionarily diverse species. Our analysis reveals several key findings. First, animal hearts share more than 60% of gene families and a conserved cellular foundation, with certain atrial and ventricular characteristics inherited by both protostomes and deuterostomes. This pattern suggests an ancient origin traceable to the last common bilaterian ancestor, consistent with fossil evidence from the Cambrian arthropod *Fuxianhuia protensa* (3). The substantial genetic commonality indicates that bilaterian hearts are unlikely to have evolved independently through parallel co-option of similar gene regulatory networks. The early expansion of cardiac genes in chordates aligns with the 1R and 2R and CR WGD events, yet the 3R WGD event, which occurred in teleost fish, did not coincide with a notable expansion of cardiac gene families, perhaps because its more recent occurrence placed the resulting gene duplicates into pre-existing families (67). Second, the diversity in cardiac architecture reflects lineage-specific genetic innovations, flexible cell-type combinations, and molecular-level functional adaptations among distinct cell types, which may collectively explain the cardiac adaptation to different

environments. Both gene loss and gain events, evident in our data, have contributed to cardiac evolution, and partial loss of shared CoreBA2 genes reshaped cardiac gene repertoires, whereas lineage-specific genetic acquisitions facilitated rapid structural innovation. Third, leveraging the strength of spatial transcriptomics, we identified a transitional CM state, designated as CM (trans.). Our analysis revealed that the CM (trans.) state exhibits a potential association with heart development or myocardial stress. During development, the proportion of CMs (trans.) was higher in both human and mouse embryonic hearts than in adult hearts. In the adult context, the proportion of CMs (trans.) increased in response to hypoxic exposure and MI but not under transverse aortic constriction stress. The mammalian embryonic heart inhabits a relatively hypoxic intrauterine environment compared with that after birth (68), and an oxygen concentration gradient (from low to high) has been reported across the infarcted myocardium, with hypoxia in the infarcted zone and BZ and normoxia in the RZ (69–71). Given these observations, the tissue oxygen condition could be a potential factor driving the dynamics of CMs (trans.). Notably, the presence of CMs (trans.) under low-oxygen conditions appears to correlate with retained regenerative capacity. It has been well documented that the neonatal mouse heart, which retains regenerative potential within the first week of life, and the adult zebrafish heart both exhibit a hypoxic myocardial microenvironment (72). Although the nuclear aggregation of this cell state might be influenced by various other factors, the distinct cytoplasmic aggregation in the BZ area of MI hearts observed under specific conditions may be associated with the formation of P-bodies and stress granules. The biological function of this state and the underlying molecular mechanisms need to be elucidated in the future. Fourth, our evidence indicates that the ancestral proto-heart was primarily composed of ventricular-like myocardium, supporting the hypothesis that heart evolution originated from contractile muscle cells representing the primordial forms of CMs (73). Taken together, these findings paint a picture of the pan-evolutionary history of animal hearts.

There are still challenges to overcome in comprehending the ground-work of heart evolution over deep time. Although we have included highly diverse species in our study, the species we can presently study cannot fully represent the cellular and molecular profile changes at evolutionary branching points that happened millions of years ago or those in key phyla that have become extinct, leaving gaps in our data resources. Technically, the capture bias and coverage of scRNA-seq is still a concern, although we have addressed this issue by optimizing sampling and experimental design as much as possible. scRNA-seq technology is highly effective for capturing diverse non-CM populations in the heart (33, 34). However, it presents challenges for CMs owing to their large size. Spatially resolved transcriptomics overcomes this limitation by preserving cellular architecture, thereby offering a superior approach for characterizing CM states (33, 74). A fundamental problem still to be overcome is that our knowledge about gene function is only from a limited number of classical model animals, which excludes many genes because of a lack of functional annotation. Linking key evolutionary genes to the acquisition of biological functions and identification of biomedically important genes will need extensive functional validation in future studies. Finally, understanding evolution from developmental data is a valuable goal to pursue in the next phase. Overall, this study offers a framework for decoding the intrinsic mechanisms of heart evolution across bilaterians, linking repertoires of core

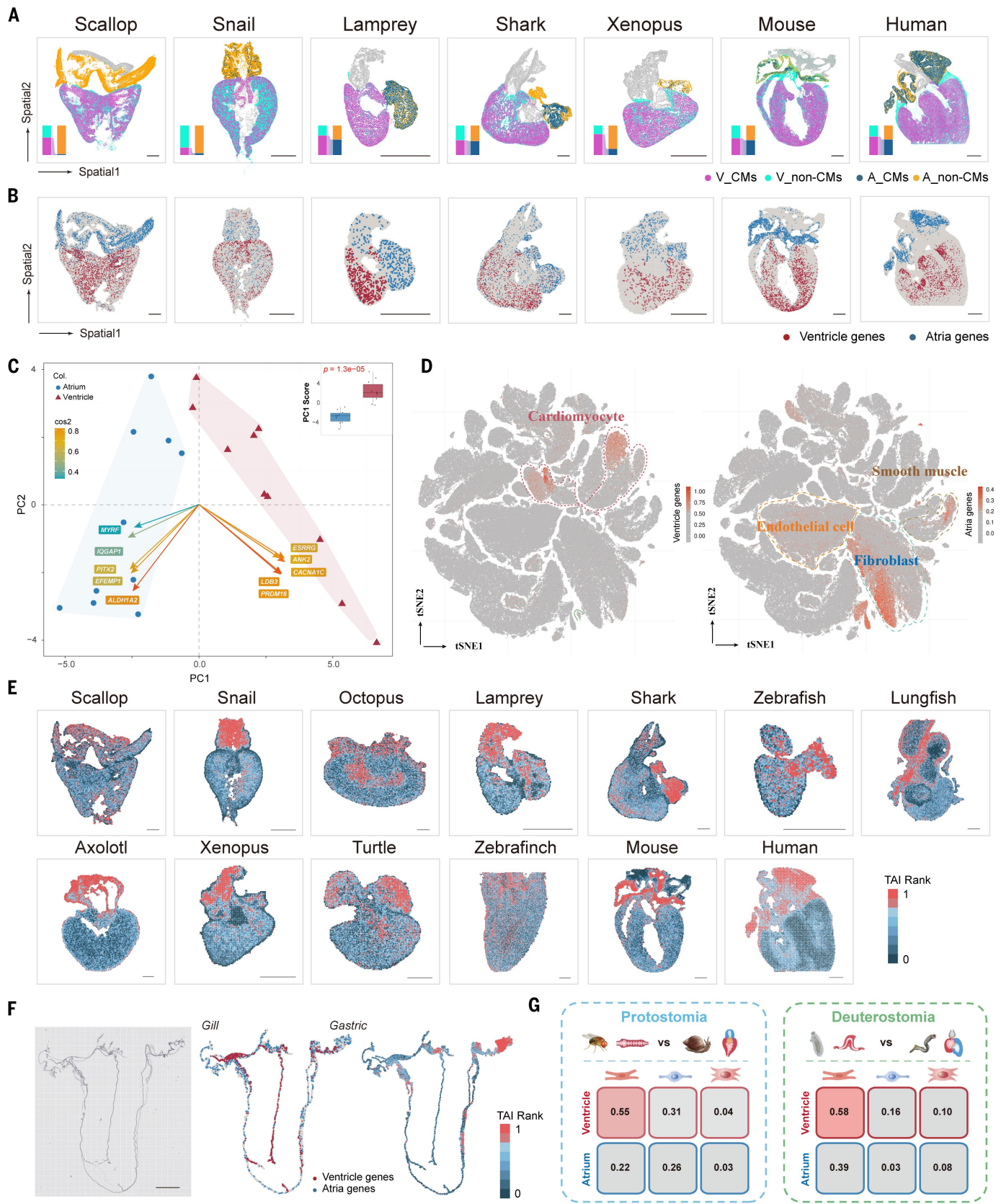


Fig. 6. Evolutionary origin of cardiac chambers. (A) Cellular composition of atrial and ventricular tissues. The bar charts show the ratio of myocardial to nonmyocardial cells in atrium (left semicircle) and ventricle (right semicircle) across species. (B) Spatial expression patterns of core atrial (blue) and ventricular (red) gene sets. Colors indicate the top 10% of pseudocells by AddModuleScore. (C) PCA plot of core chamber-specific genes. The inset boxplot shows the PC1 score distribution between atrial and ventricular samples (center line: median; box limits: upper and lower quartiles; whiskers: minimum and maximum values). Blue dots denote atrial samples (across species).

and red triangles denote ventricular samples. Arrows indicate the top 10 genes that contribute to atria or ventricles. **(D)** Projection of atrial and ventricular gene set scores onto the integrated single-cell data (t-SNE) embedding of 16 species. Color intensity reflects the AddModuleScore. **(E)** TAI at pseudocell resolution for atrial and ventricular tissues. A blue-to-red gradient indicates low-to-high TAI rank. **(F)** Spatial transcriptomic profiling of the sea squirt heart. Shown from left to right are a hematoxylin and eosin (H&E)-stained tissue section, spatial expression patterns of the core atrial and ventricular gene sets, and a pseudocell-level TAI map. **(G)** Cross-species similarity analysis of major cardiac cell types (CM, EDC, and FB) between tubular and two-chambered hearts. Shown on the left and right are Protostomes (fruit fly versus snail) and Deuterostomes (sea squirt versus lamprey), respectively. Scale bars are 1000 μm . Data are represented as means $\pm$ SD. All *P* values were calculated using two-tailed unpaired Student's *t* test. *P* < 0.05 is considered statistically significant.

gene, conservation of cell type, and dynamic progress of chambers during heart adaptive evolution.

Materials and methods

Experimental animals

Among the protostomes, there are four arthropods (horseshoe crab, fruit fly, white shrimp, and swimming crab) and five mollusks (snail, oyster, giant clams, scallop, and octopus). Horseshoe crab was a kind gift from H. Liu, Guangxi Academy of Marine Sciences; giant clams were a kind gift from Y. Zhang, Key Laboratory of Tropical Marine Bio-resources and Ecology, Chinese Academy of Sciences; fruit flies and snails were from the laboratory culture system from Ocean University of China; adult scallops were purchased from the Zhangzidao Fishery Group, China; and other animals, including white shrimp, swimming crab, oyster, and octopus, were purchased from a local fishery market (Qingdao, China). Among deuterostomes, together there are 17 chordates, including amphioxus, sea squirt, lamprey, skates, bamboo shark, zebrafish, lungfish, axolotl, tropical clawed frog, *Xenopus*, gecko, turtle, chicken, zebra finch, mouse, rat, and shrew. Amphioxus was a kind gift from Z. Gao, Ocean University of China; lampreys were cultured in the laboratory system, Lamprey Research Center, Liaoning Normal University; tropical clawed frog was a kind gift from X. Qi, Jinan University; zebra finch was a kind gift from Z. Wang, Zhengzhou University; zebrafish (AB, 4 to 6 months) were cultured in the laboratory system, Ocean University of China; axolotl (d/d, 18 months) were from the laboratory culture system, Guangdong Provincial People's Hospital; mouse (C57BL/6j, 9 to 10 weeks) and rat (Wistar, 7 to 8 weeks) were purchased from Jinan Pengyue Experimental Animal Breeding Co., Ltd.; tree shrew was purchased from Kunming Institute of Zoology, the Chinese Academy of Sciences; sea squirts were collected from coastal waters in Rongcheng, China; and other animals, including skates, bamboo shark, lungfish, *Xenopus*, turtle, gecko, and chicken, were purchased from a local commercial market (Qingdao, China). Detailed data information (biological replicates, accession numbers) is provided in table S1.

Procurement of donated tissues: Human fetal cardiac tissues were obtained from a fetus delivered vaginally at 14 weeks and 5 days gestation owing to an inevitable abortion. The pregnant woman had no complications, and the fetus was confirmed to have no structural malformations or chromosomal abnormalities. The human adult cardiac tissue was sourced from a 58-year-old male patient diagnosed with hypertrophic obstructive cardiomyopathy. Myocardial tissue was acquired from the left ventricular septal side during surgery.

Ethical statement: Animal care and experiments in this study were conducted in accordance with the Animal Welfare Guidelines of the National Institutes of Health (NIH) and the Laboratory Animal Guidelines of China. All experimental processes were approved by Research and Ethical Committee of Ocean University of China (OUC-AE-2024-124), Ethical Committee of Liaoning Normal University (LL2024221), Ethical Committee of Guangdong Provincial People's Hospital (KY2024-795-01), Research and Ethical Committee of BGI (FT19057-T1), and Ethics Review Committee of Qingdao Women and Children's Hospital (QFELL-YJ-2-24-213). The procurement of all human cardiac tissue samples was approved by the Research Ethics Committee, and written informed consent was obtained from the donors' families.

Genomic datasets

Out of 27 investigated animals, the fully sequenced genomes of 20 species were downloaded directly from National Center for Biotechnology Information (NCBI; <https://www.ncbi.nlm.nih.gov>), FlyBase (<https://flybase.org>), and MolluscDB database [<http://mgbase.qnlm.ac/>] (75). The genomes of the remaining seven species with relatively poor annotation were improved to varying degrees. We have reassembled and annotated the snail by adopting suitable assembly software (for example, NextDenovo for Mollusk) (76). By leveraging newly generated transcriptomes from this study, gene annotation of axolotl, octopus, swimming crab, white shrimp, and horseshoe crab integrates de novo assembly of these cardiac transcriptomes. For giant clam, owing to the absence of publicly available genome annotation data, we performed a comprehensive reannotation according to the released genome on NCBI. Transcriptome assembly was carried out with Trinity v2.1.1 (77, 78). TransDecoder v5.5.0 (parameters: -single_best_orf) was used for the in silico translation of the transcriptomes. The completeness of gene annotations was assessed using BUSCO v5.2.2 (benchmarking universal single-copy orthologs) (79).

Orthology inference and phylogenetic analysis

To reliably screen genes related to heart function across 27 investigated Bilateria animals, we used Orthofinder v2.5.2 (80) (parameters: -S blast -M msa) to infer gene orthogroups among all protein-coding genes. To avoid redundant gene homology calls, we selected the longest isoform for each gene to represent the respective gene. The resulting orthogroups were subsequently aligned using MAFFT v7.4.90 (81) (parameters: -localpair -maxiterate 1000), and for each alignment, an approximately maximum-likelihood gene tree was constructed using FastTree v2.1.10 (82). The alignments and their corresponding phylogenetic trees were then used in tree-based orthology inference program PhyloPyPruner (parameters: -min-len 100 -min-support 80 -prune MI). Multiple sequence alignments were performed for each single-copy gene using MAFFT v7.4.90 (81). The optimal substitution model of amino acid substitution was selected on the basis of the Akaike information criterion using ModelTest-NG v0.1.6 (83). The maximum likelihood phylogenetic tree was constructed with the best substitution models by RaxML v8.2.9 (84). Molecular dating was performed using MCMCtree implemented in PAML v4.8 (85). Fossil-derived timescale and evolutionary history were obtained from TIMETREE (<http://timetree.org>) (86).

Phylogenetic inference of heart-related gene family gains and losses

To obtain the gene repertoire associated with the hearts in Bilateria animals, we selected the atria, ventricles, or whole hearts of 24 of the 27 species for transcriptome sequencing. The raw reads were subsequently trimmed for quality using Trimmomatic v0.39 (87). Clean data of FASTQ format were mapped against the reference genomes using the STAR v2.7.7a (88) program with default parameters. We quantified expression of all genes by transcripts per million (TPM) reads. To ensure a meaningful cross-species comparison, our analysis focused specifically on gene families present in at least two species and obtained 21,473 orthogroups in total. To mine the heart-related gene family, we first selected the highest gene expression of the gene in each heart sample to represent expression level of that gene and then selected the highest gene expression level in each gene family to represent the gene expression level. Finally, we filtered gene families to retain only

those with a heart expression level (TPM > 5) in at least one species. This process yielded a final set of 15,577 heart-related gene families, which were categorized into four distinct clusters: CoreBA1, CoreBA2, CorePT, and CoreDT. We consider a gene family to be gained primarily by the absence of homologs outside the clade and the presence of homologs within the clade in at least one other lineage. Instead, we consider a gene family to be lost mainly when there are no homologs in the branch, but the homologs of the gene are present in the closest sister lineages of the branch.

Transcriptomic datasets

To comprehensively investigate cardiac evolution across bilaterians, we performed large-scale transcriptomic profiling, encompassing bulk RNA-seq, scRNA-seq, and spatial transcriptomics. Our dataset includes 174 bulk RNA-seq samples from 23 species with diverse cardiac structures. From these, we strategically selected 15 representative species for scRNA-seq, generating 38 libraries to capture key evolutionary transitions and adaptive traits. Furthermore, we conducted spatial transcriptomic analyses on 12 species, producing 14 tissue slices via Stereo-seq (including fruit fly and amphioxus). Additionally, we generated two spatial transcriptome slices for the sea squirt using NovaSeq 6000 technology.

Bulk RNA-seq: Tissues were dissected from adult animals. High-quality total RNA was extracted using the RNeasy Mini Kit (Qiagen, Hilden, Germany), which was determined to be the optimal protocol for our diverse samples. RNA libraries were then prepared with the MGIEasy Fast RNA Library Prep Kit for the DNBSEQ-T1 platform. Library quality and quantity were assessed on a Qseq400 bioanalyzer, and final sequencing was performed on the DNBSEQ-T1 platform to generate 100 base pair (bp) paired-end reads.

scRNA-seq: Fresh tissue blocks devoid of necrotic, hemorrhagic, or fibrous regions were dissected and minced into ~0.5-mm³ pieces in RPMI 1640 medium (Invitrogen). Tissues were enzymatically dissociated at a suitable temperature for 30 min with agitation using a cocktail containing 0.25% collagenase IV (Invitrogen, no. 17104-019), 0.2% collagenase II (BBI, no. A004174-0001), 0.1% Dispase II (Yeasten, no. 40104ES80), and 0.05% elastinase (Worthington, no. LS002292). The resulting cell suspension was sequentially filtered through 100- and 70-μm cell strainers (BD) and centrifuged at 300g for 5 min. After supernatant removal, the cell pellet was treated with a red blood cell lysis buffer (Thermo Fisher Scientific) on ice for 3 min, washed twice with phosphate-buffered saline (PBS) (Invitrogen), and finally resuspended in PBS containing 0.04% bovine serum albumin (BSA). Single-cell libraries were prepared using either the DNBelab C Series Kit (MGI, no. 1000021082) or the BMKMANU DG1000 GEM Kit, following the manufacturers' protocols. Briefly, single-cell suspensions were partitioned into droplets for cell lysis and barcoding. After emulsion breaking and bead collection, reverse transcription was performed, followed by cDNA amplification to generate barcoded libraries. The final libraries were sequenced on either the DNBSEQ-T1 platform (with a Read 1 of 41 bp and Read 2 of 100 bp) or the NovaSeq 6000 platform (PE150).

Spatial RNA-seq: Hearts from adult animals were isolated and rapidly fixed by submersion in methanol precooled to -20°C for 5 min. Tissues were snap-frozen in Tissue-Tek OCT (Sakura, no. 4583) using isopentane pre-chilled with liquid nitrogen and stored at -80°C until cryosectioning. Prior to sectioning, OCT-embedded tissues were equilibrated at -20°C for 30 min. Transverse cryosections of 10 μm thickness were collected at 10 μm intervals using a Leica CM1950 cryostat. After washing the Stereo-seq chip with NF-H₂O containing 0.05 U/μl RNase inhibitor (NEB, no. M0314L), cryosections were mounted onto individual chips and incubated at 37°C for 3 min. Fixation was performed with methanol at -20°C for 40 min, followed by staining with a nucleic acid dye (Thermo Fisher Scientific, no. Q10212) for imaging single-stranded

DNA. Whole-structure images were acquired using a Motic Custom PA53 FS6 microscope system with a 20× objective. Tissue sections were permeabilized with 0.1% pepsin (Sigma, no. P7000) in 0.01 M HCl at 37°C for 6 min. Captured RNA was reverse-transcribed at 42°C overnight using SuperScript II reverse transcriptase. Sections were washed twice with Wash Buffer and digested with Tissue Removal Buffer (10 mM Tris-HCl, 25 mM EDTA, 100 mM NaCl, 0.5% SDS) at 37°C for 30 min. The resulting cDNA was amplified using KAPA HiFi HotStart ReadyMix (Roche, no. KK2602) with 0.8 μM cDNA-PCR primer. Sequencing libraries were prepared through concentration quantification, fragmentation, PCR amplification, and purification. Purified products were used for DNA nanoball (DNB) generation and sequenced on an MGI DNBSEQ-T10 sequencer with 100 bp read length.

scRNA-seq and spatial RNA-seq for sea squirt: For single-cell RNA-seq, cell suspensions were fixed with methanol and resuspended in SSC buffer prior to capture using the BMKMANU DG1000 platform. Approximately 20,000 cells were used as input to generate Gel Bead-in-Emulsions (GEMs). The subsequent workflow included reverse transcription, emulsion breaking, cDNA cleanup and amplification, and library construction. Library quality was assessed using a Bioanalyzer and Qubit, followed by sequencing on the NovaSeq 6000 platform in PE150 mode. For spatial RNA-seq, hearts were isolated and perfused with 50% Tissue-Tek OCT from the myocardial tube end, embedded in OCT, and frozen on dry ice for ~5 min. Samples were stored at -80°C until sectioning. Cryosections of 10 μm thickness were mounted onto pre-cooled Tissue Optimization and Gene Expression Chips (BMKMANU no. S1000). Tissue optimization, fixation, H&E staining, imaging, and gene expression procedures were performed using the Tissue Optimization Kit (BMKMANU, no. ST03003) and Gene Expression Kit (BMKMANU, no. ST03002). Briefly, the Tissue Optimization Chip was incubated at 37°C for 1 min and fixed in precooled methanol at -20°C for 30 min. Sections were then stained with hematoxylin and eosin (H&E) and imaged at 40× magnification using a Panoramic MIDI microscope. Optimal permeabilization time for sea squirt heart sections was determined to be 20 min using 1 ml of permeabilization buffer. After a 0.1× SSC wash, the chip was incubated with first-strand cDNA synthesis mix and reverse-transcribed at 42°C for 90 min. After two washes, second-strand synthesis mix was added and incubated at 65°C for 30 min. cDNA was released using cDNA release buffer (0.08 M KOH, 100 mM Tris-HCl), amplified, and purified. Libraries were constructed using the Truseq Illumina library prep kit and sequenced on a NovaSeq 6000 with a minimum of 150,000 read pairs per spot. The sequencing protocol included: Read 1: 28 cycles; i7 index: 10 cycles; i5 index: 10 cycles; Read 2: 50 cycles, yielding 300 million to 400 million total reads.

TAI analysis of the heart and identification of key genes for heart evolution

Gene ages and the transcriptome age indexes were estimated using the phylostratigraphic approach as previously described (20). The phylostrata (ps) were defined according to the NCBI Taxonomy database, with psI representing the origin of cellular life and psII leading to each species. The putative homologous protein from each species was assigned to each ps level with hierarchical priority, through BLASTP with an e-value threshold of 1×10^{-3} according to previous studies (21). Then, myTAI package was used to calculate the TAI at both the organ level (bulk datasets) and cellular level (spatial transcriptomics data at bin100 resolution). For bulk datasets, by integrating the ps level of genes together with their expression levels (TPM), a weighted arithmetic mean of transcriptome age is obtained (89). To determine the ps level at which the heart most probably originated, we first calculated the TAI for each phylostratum and then used these values to derive the relative TAI contribution in heart (TAI-H) at each ps level (*N*) of a given species (*s*) as previously described (21). This contribution

is defined as $PS_{s(N)} = TAI_{s(N)} - TAI_{s(N-T)}$. To identify the key genes for heart evolution, the relative contribution of a specific gene to global TAI was assessed using the pMatrix function in the package myTAI (90), using the bootstrap approach with 1000 replicates to assess the standard error of global TAI. Genes with a high TAI contribution at ps6 shared by at least 15 species in chordates, 4 species in lophotrochozoans and 3 species in ecdysozoans were identified as core ps6 genes. Genes with high TAI contribution at ps9 shared by at least 13 species in chordates, 2 species in lophotrochozoans, and 2 species in ecdysozoans were identified as core ps9 genes. To test statistical significance for the difference in observed TAI profiles among different heart chambers, TAI values were first normalized by fold difference relative to either the average value of $TAI_{s(N)}$ across the heart chambers (for different chamber comparison); the statistical significance for TAI comparison of chambers were evaluated based on one-sided *t* testing. For spatial transcriptomics-based TAI analysis, we obtained the CPM matrix for each pseudocell from the spatial transcriptomics data of each species. Then, we used the myTAI package v0.9.3 to calculate the TAI score for each pseudocell of each species. Finally, we projected the TAI scores onto the corresponding spatial coordinates of each species using the ggplot2 package v3.5.1.

Single-cell transcriptome data analysis

scRNA-seq data were processed using standardized pipelines to ensure high-quality and comparable datasets across samples. Reads were aligned to corresponding reference genomes (table S2) using DNBelab_C_Series_HT_scRNA-analysis-software (v2.0.7) for most samples and Cell Ranger (v7.1.0) for the sea squirt dataset. For each sample, the single-cell count matrices were filtered to retain cells expressing more than 200 genes and genes detected in at least three cells. Potential doublets were identified and removed with DoubletFinder (v2.0.3), excluding ~5% of cells. Subsequent preprocessing was performed in Seurat, including data normalization, identification of highly variable genes (FindVariableFeatures, nfeatures = 3000), and feature scaling (ScaleData). PCA (RunPCA, npcs = 30) was used for initial dimensionality reduction. To correct for batch effects across samples, we applied Harmony integration (RunHarmony, grouping variable = "orig.ident"). The Harmony-corrected embeddings were then used for uniform manifold approximation and projection (RunUMAP, reduction = "harmony", dims = 1:30). Cell clustering was performed using FindClusters with a resolution of 0.5. DEGs for each cluster were identified with FindAllMarkers (min.pct = 0.25, logfc.threshold = 0.25, test.use = "wilcox").

Spatial transcriptome data analysis

Spatially resolved RNA-seq data were generated using the Stereo-seq platform. Raw sequencing data were processed with the Stereo-seq Analysis Workflow (<https://github.com/BGIResearch/SAW>) to map reads to their spatial coordinates and gene features. Spatial gene expression matrices were constructed using clean exon-mapped reads. For cell-type annotation, we first aggregated the cellbin-resolution data using spatial proximity-based clustering (SPC) to generate spatially coherent cellular units. This approach was applied to 10 species, whereas human and mouse data were analyzed at single-cell bin resolution, and zebrafish and ascidian data at bin50 resolution, to achieve optimal annotation quality. The aggregated cell-level data were normalized using the SCTransform function in Seurat (v5.0.1) (90). We then performed PCA and used the top 30 principal components for unsupervised clustering, independent of spatial information. Marker genes for each cluster were identified with the FindAllMarkers function in Seurat. Cell type annotation was conducted in two steps: (i) preliminary labeling via SAMap (v1.0.15) by transferring annotations from integrated single-cell and spatial transcriptomic datasets and (ii) manual refinement based on spatial context and known marker gene expression. Final annotations were mapped back to the original cellbin-resolution data for downstream analyses.

Ancestral state reconstruction analysis

Reconstructed ancestral states on the maximum likelihood topology (RaxML v8.2.9) using likelihood reconstruction methods in Mesquite 3.81 (91). The likelihood reconstruction finds, for each node, the state assignment that maximizes the probability of arriving at the observed states in the terminal taxa, given the model of evolution, and allowing the states at all other nodes to vary. The one parameter Markov k-state (Mk1) model was applied in the likelihood analysis and assumes a single rate for all character state transitions (92). Likelihood-ratio tests of respective nodes determined the best estimate of the state.

Cross-species analysis of the core gene set

The integration of the single cell atlas of the heart across various species was carried out using the SATURN tool v2023 (35). Initially, the ESM2B model was used to generate protein embeddings, which were subsequently averaged for all genes. This was followed by the application of SATURN with a specified number of macrogenes (2000) and highly variable genes (8000). To identify conserved core gene sets for CMs and FBs, we analyzed differential gene expression across species. DEGs for each cell type were identified using the FindAllMarkers function in Seurat v5.0.1 (parameters: $P < 0.05$, $\text{pct.1} > 0.25$, and $\log_2\text{FC} \geq 0.25$, where FC is fold change). Gene modules were detected using hdWGCNA v0.2.1, and genes with eigengene-based connectivity (kME) > 0.3 were considered core genes per species. Cross-species conserved core genes were defined as those expressed in $\geq 50\%$ of species within protostomes and deuterostomes, respectively. These genes were classified into three categories based on expression patterns: (i) highly expressed ($P < 0.05$, $\text{pct.1} > 0.25$, $\log_2\text{FC} \geq 1$, or $kME \geq 0.6$); (ii) broadly expressed ($P > 0.05$, $\text{pct.1} > 0.25$, $0.25 \leq \log_2\text{FC} < 1$, or $0.3 \leq kME < 0.6$); and (iii) lowly expressed ($\text{pct.1} > 0.25$) in the respective cell type. Finally, expression patterns of conserved core genes and partially absent genes (absent in at least one but not all species) were visualized using pheatmap v1.0.12. For ancestral-type genes, in cases where a single invertebrate ortholog corresponded to multiple vertebrate paralogs, we filled in the presence calls in invertebrates, consolidating them as representatives of the ancestral gene family.

Cross-species analysis of cell-type evolution

To investigate the evolutionary trajectory of cardiac cell types in chordates, we analyzed single-cell transcriptomic data from this study, focusing on CMs as an illustrative example. CM subsets were isolated from each species, and to mitigate data sparsity, we generated pseudo single-cell data by normalizing expression values and computing the average for every 10 cells. Pseudocells from all species were then integrated, retaining only genes detected in at least three pseudocells (missing values were imputed as zero). PCA (93) was performed on the integrated pseudocell matrix using the precomp function (FactoMineR v2.11) (94). The resulting distribution was visualized along the first two principal components (PC1 and PC2) using ggplot2 (v3.5.2), with each point representing a pseudocell colored by species. Based on the PCA clustering and established phylogeny, the 11 species were categorized into three major groups: Prochordates (ascidian, amphioxus), Anamniotes (lamprey, shark, zebrafish, lungfish, axolotl, *Xenopus*), and Amniotes (turtle, zebra finch, mouse), delineated by encircling contours. To identify genes driving the divergence among these groups, we quantified each gene's contribution to the principal components using the get_pca_var() function, which calculates the relative importance of genes (via var\$contrib) in defining the group identities.

GO enrichment analysis and protein-protein interaction network construction

For gene module and functional GO enrichment analysis, marker genes were detected using Seurat v5.0.1 with FindAllMarkers function (95). The top 100 DEGs based on fold change were further

selected to do functional enrichment analysis, and parameters, $pct.1 > pct.2$ and adjusted $P < 0.01$ were further applied in Stereo-seq datasets. Gene module analysis of each cluster was performed using Seurat v5.0.1 with the AddModuleScore function and visualized using ggplot2. Function enrichment analysis of each cell subtype was performed by the online tools, Metascape v3.5.20240901 (<https://metascape.org/>) (96), with the Human dataset as reference. All enrichment results and protein-protein interaction network figures were exported from Metascape for further statistical analysis and visualization in the study.

Chamber segmentation and cross-species analysis of atrial and ventricular cell proportions

Atrial and ventricular regions were delineated using cellbin-resolution spatial transcriptomic data. Spatial coordinates were segmented into atrial, ventricular, and other regions. Cells were classified as atrial CMs, ventricular CMs, atrial non-CMs, or ventricular non-CMs based on chamber location. Cell-type proportions were normalized to 1 and visualized using ggplot2 v3.5.1.

Identification and spatial expression of core atrial and ventricular gene sets

The identification of core atrial and ventricular gene sets was performed through a multistep approach integrating bulk RNA-seq and spatial transcriptomic data. First, conserved atrial and ventricular genes were identified from bulk RNA-seq data of 17 species. For species with more than three replicates (scallop, oyster, giant clams, snail, lamprey, zebrafish, axolotl, *Xenopus*, chicken, zebra finch, rat, mouse, tree shrew), differential expression analysis was performed using DESeq2 v1.38.8 (97). For species with single samples (skates, bamboo shark, tropical clawed frog, gecko), DEGs were called using GFold v1.1.4 (98). Genes with $\log_2FC > 1.5$ were considered atrial specific, and those with $\log_2FC < -1.5$ as ventricular specific. Species-specific gene IDs were converted to standardized cross-species gene names, and species-unique genes were removed. Conserved atrial genes were defined as those present in ≥ 4 of the 16 species in the atrial DEG list; conserved ventricular genes were defined similarly. Genes were further classified into protostome and deuterostome groups. Atrial DEGs were retained if expressed in ≥ 2 of 4 protostome species and ≥ 4 of 12 deuterostome species. The same approach was applied to ventricular DEGs. Second, conserved atrial and ventricular genes were also identified from spatial transcriptomics data (bin100 resolution) across 12 species. Spatial group files were converted to RDS format, and marker genes were identified using the FindAllMarkers function in Seurat v5.0.1. Gene IDs were standardized to cross-species names, and species-unique genes were excluded. Conserved atrial genes were defined as those expressed in ≥ 3 of 12 species; conserved ventricular genes required expression in ≥ 6 of 12 species. For lineage-specific conservation, atrial markers were retained if expressed in ≥ 2 of 3 protostome species and ≥ 3 of 9 deuterostome species. The same criteria were applied to ventricular markers. Finally, the core gene set was derived by integrating conserved atrial genes from both bulk RNA-seq and spatial transcriptomics, including lineage-specific genes from protostomes and deuterostomes. ATP synthase, COX, and ribosomal genes were excluded, except in amphibians where ribosomal genes were retained owing to high atrial expression. PCA was performed using the factoextra package (v1.0.7), and the top 70 contributing genes were selected as the core gene set. Expression patterns of these core gene sets were evaluated in using the AddModuleScore function in Seurat v5.0.1, and scores were visualized on cross-species t-SNE plots using ggplot2 v3.5.1.

Analysis of spatial patterns of core gene repertoire

Spatial expression patterns of core gene sets (CoreBA1, CoreBA2, CorePT, CoreDT) were analyzed using orthogroup information from the

cross-species gene family clustering tree. Gene lists for each core set were obtained from the single-copy gene annotation table. Expression levels were scored using the AddModuleScore function in Seurat v5.0.1 at bin100 resolution, and the top 10% of pseudocells by score were projected onto spatial coordinates. In addition, spatial expression of the gene family gain and CoreBA1 gene sets was similarly assessed. Protein IDs corresponding to the longest transcript of each orthogroup were obtained from the gene family tree and converted to gene IDs to generate species-specific gene-gain lists. Expression patterns of both the Gene Gain and CoreBA1 gene sets were evaluated using the AddModuleScore function. For CoreBA1, orthogroup information was specifically extracted from the clustering tree, and corresponding gene lists were retrieved from the single-copy gene annotation table. Module scores for both gene sets were computed and projected onto spatial coordinates, with the top 10% of pseudocells highlighted using ggplot2 v3.5.1. Spatial expression patterns of the core atrial and ventricular gene sets were analyzed following the same methodology described above.

Cross-species SAMap analysis of cell types and cardiac chambers

Cross-species similarity at the cell type and chamber levels was assessed using SAMap v1.0.15 (99). Mapping scores (0 to 1) were computed for all cross-species cluster pairs, and similarity matrices were visualized as heatmaps using pheatmap (v1.0.12). This analysis was applied to single-cell or spatial datasets including ascidian (Cro: SPC), lamprey (Lre: Cellbin), fruit fly (Dme: SC), and snail (Lfu: Cellbin). At the chamber level, the same pipeline was used to compare the ascidian or fruit fly cardiac tube with atrial-ventricular partitioned chambers across 13 multi-chambered species.

TF and cell-cell communication analysis

TF analysis: TFs for human, mouse, zebrafish, zebra finch, sea squirt, and fruit fly were initially obtained from AnimalTFDB v4.0 (100). To comprehensively identify TFs across species in the single-cell sequencing analysis, we performed additional annotation using eggNOG-mapper v2.1.3 (101). Then, we defined TF repertoires of other species by integrating homology information among the studied species with TF predictions from both sources. Using these TF sets, we reconstructed TF–gene regulatory networks (GRNs) from single-cell data using the GRNBoost module in pySCENIC v0.11.2 (102). Because each TF typically regulates numerous predicted target genes, we focused on the top 200 targets per TF to ensure interpretability in downstream analyses. We then quantified overlaps between core gene sets and TF-regulated GRNs, ranking TFs by the proportion of core genes among their predicted targets. Notably, in nearly all species examined, over 50% of the high-overlap TFs were associated with CoreBA1 and CoreBA2. We further analyzed TF–core gene overlaps at the species level and identified distinct phylogenetic patterns. Additionally, we identified evolutionarily conserved core-associated TFs across multiple lineages, defined as TFs linked to CoreBA1 (present in ≥ 3 species), CoreBA2 (≥ 3 species), CorePT (≥ 2 protostome species), and CoreDT (≥ 3 deuterostome species). For fruit fly, mouse, and human, regulon networks were constructed using the R package SCENIC (v1.1.3), which evaluates co-expression patterns between TFs and their putative target genes. Gene regulatory networks were built and scored using the package's default parameters.

Cell-cell communication analysis: To investigate intercellular signaling, we integrated single-cell and spatial transcriptomic data for each species. Based on established homologous relationships, all genes were converted to human ortholog identifiers. We then used the CellChat v2.0.0 (103) to predict and visualize biologically relevant cell-cell communication. Specifically, the createCellChat function was used to generate CellChat objects, followed by cell annotation and identification of overexpressed ligands or receptors. Communication

probabilities were inferred using the `computeCommunProb` function (`distance.use = TRUE`, `interaction.range = 250`, `scale.distance = 0.025`), and pathway-specific interaction networks were predicted with `computeCommunProbPathway` function. The cross-species expression patterns of key GRN-associated TFs and cell type-specific ligand-receptor genes identified in the above analyses are described in the “Cross-species analysis of the core gene set” section.

Histological sectioning, staining, and imaging

Cryosectioning: Cryosectioning was performed as previously described (104). Briefly, hearts were dissected and fixed with 4% paraformaldehyde (PFA) in PBS overnight at 4°C. After washing with PBS, the hearts were embedded in a 1.2% (wt/vol) agarose in 5% (wt/vol) sucrose-PBS solution, equilibrated in a 30% (wt/vol) sucrose-PBS solution overnight at 4°C, and frozen for cryosectioning using a Leica cryostat. Ten-micrometer cryosections were used in all experiments.

AFOG staining: AFOG staining was performed to visualize cardiac muscle fibers in orange and collagen in blue. Briefly, tissue sections were hydrated in dH₂O at 25°C and then incubated in Bouin's solution (Sigma, no. HT10132) for 2 hours at 60°C, followed by an additional hour at room temperature. After rinsing under running water for 30 min, slides were treated with 1% phosphomolybdic acid (Solarbio, no. G3472) for 5 min and rinsed again in dH₂O for 5 min. Staining was carried out in AFOG staining solution, containing 1 g of aniline blue (G-CLONE, no. CS2805), 2 g of orange G (BBI, no. 1936-15-8), and 3 g of acid fuchsin (G-CLONE, no. CS3244) dissolved in 200 ml of dH₂O (pH 1.09) for 10 min. After a dH₂O rinse, slides were dehydrated through a graded ethanol series (95 and 100%) and cleared in two changes of xylene (BBI), 2 min each. Finally, slides were mounted with Cytoseal (Thermo Fisher Scientific, no. 8310-4) and coverslipped.

Immunofluorescence staining: Cryosections were subjected to antigen retrieval by immersion in boiling citric acid buffer (10 mmol/liter). After washing with PBS, sections were permeabilized with 0.5% Triton X-100 in PBS and blocked with blocking buffer for 1 hour at room temperature. Subsequently, sections were incubated with primary antibodies diluted in PBS containing 5% BSA overnight at 4°C. After washing, sections were incubated with secondary antibodies for 1 hour at room temperature and mounted with VECTASHIELD mounting medium containing 4',6-diamidino-2-phenylindole (DAPI). The following primary antibodies were used: anti-ACTN2 (ABclonal, no. A3718, 1:200), anti-TPM1 (DSHB, no. CH1, 1:200), anti-myosin heavy chain monoclonal antibody (DSHB, MF20, 1:50), anti-TNNC1 (ABclonal, no. A1927, 1:200), anti-TNNT2 (DSHB, no. CT3, 1:200), anti-TNNI1 (ABclonal, no. A9664, 1:200), and anti-TNNI3 (ABclonal, no. A21246, 1:200). Secondary antibodies (1:200; Thermo Fisher Scientific) used in this study include Alexa Fluor 488 goat anti-mouse immunoglobulin G (IgG) (H+ L), Alexa Fluor 568 goat anti-mouse IgG (H+ L), Alexa Fluor 568 goat anti-rabbit IgG (H+ L), and Alexa Fluor 488 goat anti-rabbit IgG (H+ L).

FISH: FISH was performed to localize specific mRNAs in hearts. Digoxigenin- and fluorescein-labeled RNA probes were synthesized using the Roche kit (Roche, nos. 11685619910 and 11277073910) according to the manufacturer's instructions. Target gene fragments were cloned from heart cDNA libraries using primers listed in table S21. After permeabilization with 3% H₂O₂ and a wash with Triton X-100, the slides underwent a 16-hour hybridization at 68°C. Posthybridization washes were performed using SSCT and MABT buffers, followed by incubation with an anti-Dig/Fluo-Pod antibody (Roche, nos. 11207733910 and 11426346910). Subsequent washes with MABT and PBSTx buffers preceded the final staining reaction and double FISH procedure. For signal detection, the TSA PLUS FLUORESCENCE kit (Akoyabio, no. NEL741001KT) and Absin 3-Color IHC Kit (abs50089-100T) were used.

Imaging: Both immunofluorescence and FISH images were acquired using a spinning-disk confocal microscope (Dragonfly 200, Andor Oxford Instruments) and Zeiss confocal (LSM 900 and LSM980).

Statistical analysis: CM states were determined based on nuclear colocalization and enrichment of *Ldb3* and *Ryr2* in CMs. The proportion of each cell state was calculated relative to the total number of CMs. Gene expression levels were quantified by measuring the fluorescence area of gene signals. Sample sizes and the types of statistical tests are provided in the figure legends. Data are presented as means $\pm$ standard deviation (SD). A significance level of $P < 0.05$ was used. For comparisons between two groups, a two-tailed unpaired Student's *t* test was used. For comparisons among multiple groups, the homogeneity of variances was assessed first using the Brown-Forsythe test. The one-way analysis of variance (ANOVA) followed by Tukey's multiple comparisons test was applied when the assumption of homogeneity of variances was met; otherwise, Welch's ANOVA followed by Games-Howell multiple comparisons test was used. All statistical analyses were performed using GraphPad Prism 10.

Models for evaluating transient CM state

Identification and classification of CM states: To define and classify the CM (trans.) and CM (const.) CM states, a systematic pipeline was established, integrating spatial transcriptomics with the evaluation of transcript location via FISH. Cross-species spatial transcriptomic profiling identified *Cacna1c*, *Ryr2*, and *Ldb3* as evolutionarily conserved marker genes for the CM (trans.) state. To empirically identify these states, FISH was performed on cardiac tissues. The CM (trans.) state was primarily distinguished by the simultaneous and pronounced nuclear aggregation of these marker transcripts. To provide auxiliary validation, image-based fluorescence intensity profiling was subsequently conducted to measure transcript abundance across the whole-cell, nuclear (marked by DAPI), and cytosolic levels. Cells identified by this distinctive nuclear aggregation signature were further validated as the CM (trans.) state by their significantly elevated expression at either the whole-cell or nuclear level.

Mouse model of MI: An MI model was established in 8-week-old male C57BL/6 mice by permanent ligation of the left anterior descending coronary artery (LAD). Mice were anesthetized via intraperitoneal injection of ketamine (100 mg/kg) and xylazine (7.5 mg/kg). After loss of pedal reflexes was confirmed, animals were placed supine on a heating pad (37°C), intubated endotracheally, and connected to a miniventilator. A left thoracotomy was performed in the fourth intercostal space under sterile conditions. The pericardium was opened to expose the heart, and a 7-0 or 8-0 polypropylene suture was passed under the LAD ~2 to 3 mm from the tip of the left atrium and permanently ligated. Sham-operated mice underwent the same procedure without LAD ligation.

Mouse model of cardiac hypertrophy: Cardiac hypertrophy was induced in 8-week-old C57BL/6 mice via transverse aortic constriction. After anesthesia (ketamine/xylazine, intraperitoneal) and ventilatory support, a thoracotomy was performed to expose the aortic arch. A 27-gauge needle was placed parallel to the aorta, and a 6-0 suture was tied around both the needle and the aorta. The needle was then removed to create aortic stenosis. The thoracic incision was closed with 4-0 sutures, and mice were kept on a heating pad until recovery. Sham-operated animals underwent the same surgery without aortic constriction. Functional and morphological analyses were performed after four weeks.

Mouse model of hypoxia: Eight-week-old C57BL/6 mice (Jinan Pengyue Laboratory Animal Breeding Co., Ltd.) were acclimatized for 3 days and randomly assigned to high-altitude hypoxia (HH) or normobaric

normoxia (NN) groups. HH exposure was conducted in a decompression chamber (YUYAN INSTRUMENTS) for 5 or 10 days under the following conditions: simulated altitude 6000 m, inspired O₂ partial pressure 11 kPa, ascent rate 20 m s⁻¹, barometric pressure 54 kPa, temperature 25 °C, relative humidity 40%, and a 12-hour light-dark cycle. The chamber was ventilated at 25 air changes per minute. Control animals were maintained under NN conditions in an identical chamber for the same durations.

Fabrication and differentiation of human heart organoids: Human pluripotent stem cells were dissociated using Accutase (Thermo Fisher Scientific, no. A1110501), followed by neutralization with CDM medium (Baishou, no. 400105). The supernatant was removed, and the cells were resuspended in CDM medium containing 5 µM Y27632. The cells were then passaged at a density of 10,000 cells per well into 96-well low-attachment U-bottom plates, which was designated as differentiation day -2. After 12 to 24 hours of passaging, the medium was replaced with CDM medium to withdraw Y27632 (Selleck, no. S6390), and the CDM medium was refreshed daily. On day 0, the CDM medium was replaced with RPMI + B27- medium [RPMI 1640 basal medium (Gibco, no. 11875176) + 2% B-27 supplement (insulin-free, Gibco, no. 17504044) + 1% penicillin-streptomycin] supplemented with 12 µM CHIR99021 (Selleck, no. S1263) and 0.2 µM BI-1347 (Selleck, no. S1058). After 48 hours of differentiation, the medium was changed to RPMI + B27-. At 72 hours (day 3 of differentiation), by which time the cells should have differentiated into cardiac mesoderm, the medium was replaced with RPMI + B27- containing 5 µM IWR1 (Selleck, no. S7086). After 48 hours of IWR1 treatment, the medium was changed to RPMI + B27-, marking differentiation day 5. The cells were cultured continuously in RPMI + B27- for two additional days, during which they differentiated into cardiac progenitor cells. On differentiation day 7, the medium was switched to RPMI + B27 [RPMI 1640 basal medium + 2% B-27 supplement (with insulin, Gibco, no. A1895601) + 1% penicillin-streptomycin]. Beating organoids were typically observed after 2 to 3 days of culture in RPMI + B27. During stage III of differentiation, the RPMI + B27 medium was refreshed every 3 days, and samples were collected on differentiation day 20.

REFERENCES AND NOTES

1. E. Carlisle, Z. Yin, D. Pisani, P. C. J. Donoghue, Ediacaran origin and Ediacaran-Cambrian diversification of Metazoa. *Sci. Adv.* **10**, eadp7161 (2024). doi: [10.1126/sciadv.adp7161](https://doi.org/10.1126/sciadv.adp7161); pmid: [39536100](https://pubmed.ncbi.nlm.nih.gov/39536100/)
2. A. Wagner, The molecular origins of evolutionary innovations. *Trends Genet.* **27**, 397–410 (2011). doi: [10.1016/j.tig.2011.06.002](https://doi.org/10.1016/j.tig.2011.06.002); pmid: [21872964](https://pubmed.ncbi.nlm.nih.gov/21872964/)
3. X. Ma, P. Cong, X. Hou, G. D. Edgecombe, N. J. Strausfeld, An exceptionally preserved arthropod cardiovascular system from the early Cambrian. *Nat. Commun.* **5**, 3560 (2014). doi: [10.1038/ncomms4560](https://doi.org/10.1038/ncomms4560); pmid: [24704943](https://pubmed.ncbi.nlm.nih.gov/24704943/)
4. G. Shields-Zhou, L. Och, The case for a Neoproterozoic oxygenation event: Geochemical evidence and biological consequences. *GSA Today* **21**, 4–11 (2011). doi: [10.1130/GSATG102A.1](https://doi.org/10.1130/GSATG102A.1)
5. W. Joyce *et al.*, Genetic excision of the regulatory cardiac troponin I extension in high-heart rate mammal clades. *Science* **385**, 1466–1471 (2024). doi: [10.1126/science.ad8146](https://doi.org/10.1126/science.ad8146); pmid: [39325895](https://pubmed.ncbi.nlm.nih.gov/39325895/)
6. A. Joven, A. Elewa, A. Simon, Model systems for regeneration: Salamanders. *Development* **146**, dev167700 (2019). doi: [10.1242/dev.167700](https://doi.org/10.1242/dev.167700); pmid: [31332037](https://pubmed.ncbi.nlm.nih.gov/31332037/)
7. H. E. Anderson, L. Christiaen, Ciona as a simple chordate model for heart development and regeneration. *J. Cardiovasc. Dev. Dis.* **3**, 25 (2016). doi: [10.3390/jcdd3030025](https://doi.org/10.3390/jcdd3030025); pmid: [27642586](https://pubmed.ncbi.nlm.nih.gov/27642586/)
8. M. J. Wells, The cephalopod heart - the evolution of a high-performance invertebrate pump. *Experientia* **48**, 800–808 (1992). doi: [10.1007/BF02118412](https://doi.org/10.1007/BF02118412)
9. N. Eisenhauer, J. Hines, Invertebrate biodiversity and conservation. *Curr. Biol.* **31**, R1214–R1218 (2021). doi: [10.1016/j.cub.2021.06.058](https://doi.org/10.1016/j.cub.2021.06.058); pmid: [34637734](https://pubmed.ncbi.nlm.nih.gov/34637734/)
10. E. E. Kolesnikova, Anatomical and physiological peculiarities of the heart in jawless and jawed fish. *J. Evol. Biochem. Physiol.* **57**, 185–207 (2021). doi: [10.1134/S0022093021020022](https://doi.org/10.1134/S0022093021020022)
11. M. Litviňuková *et al.*, Cells of the adult human heart. *Nature* **588**, 466–472 (2020). doi: [10.1038/s41586-020-2797-4](https://doi.org/10.1038/s41586-020-2797-4); pmid: [32971526](https://pubmed.ncbi.nlm.nih.gov/32971526/)
12. W. F. Ponder, D. R. Lindberg, J. M. Ponder, *Biology and Evolution of the Mollusca*, vol. 1, (CRC Press, 2019).
13. R. E. Poelmann, A. C. Gittenberger-de Groot, Development and evolution of the metazoan heart. *Dev. Dyn.* **248**, 634–656 (2019). doi: [10.1002/dvdy.45](https://doi.org/10.1002/dvdy.45); pmid: [31063648](https://pubmed.ncbi.nlm.nih.gov/31063648/)
14. C. Gestal, S. Pascual, Á. Guerra, G. Fiorito, J. M. Vieites, Eds., *Handbook of Pathogens and Diseases in Cephalopods* (Springer, 2019).
15. A. Stephenson, J. W. Adams, M. Vaccarezza, The vertebrate heart: An evolutionary perspective. *J. Anat.* **231**, 787–797 (2017). doi: [10.1111/joa.12687](https://doi.org/10.1111/joa.12687); pmid: [28905992](https://pubmed.ncbi.nlm.nih.gov/28905992/)
16. C. I. Mosley, C. A. Mosley, "Comparative anesthesia and analgesia of reptiles, amphibians, and fishes" in *Veterinary Anesthesia and Analgesia*, K. A. Grimm, L. A. Lamont, W. J. Tranquilli, S. A. Greene, S. A. Robertson, Eds. (Wiley, 2015), pp. 784–799.
17. G. Zaccane, K. Dabrowski, M. S. Hedrick, J. M. O. Fernandes, J. M. Icardo, Eds., *Phylogeny, Anatomy and Physiology of Ancient Fishes* (CRC Press, 2015).
18. C. S. Wirkner, M. Tögel, G. Pass, "The arthropod circulatory system" in *Arthropod Biology and Evolution*, A. Minelli, G. Boxshall, G. Fusco, Eds. (Springer, 2013), pp. 343–391.
19. D. Yu *et al.*, Hagfish genome elucidates vertebrate whole-genome duplication events and their evolutionary consequences. *Nat. Ecol. Evol.* **8**, 519–535 (2024). doi: [10.1038/s41559-023-02299-z](https://doi.org/10.1038/s41559-023-02299-z); pmid: [38216617](https://pubmed.ncbi.nlm.nih.gov/38216617/)
20. T. Domazet-Lošo, D. Tautz, A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature* **468**, 815–818 (2010). doi: [10.1038/nature09632](https://doi.org/10.1038/nature09632); pmid: [21150997](https://pubmed.ncbi.nlm.nih.gov/21150997/)
21. J. Wang *et al.*, Evolutionary transcriptomics of metazoan biphasic life cycle supports a single intercalation origin of metazoan larvae. *Nat. Ecol. Evol.* **4**, 725–736 (2020). doi: [10.1038/s41559-020-1138-1](https://doi.org/10.1038/s41559-020-1138-1); pmid: [32203475](https://pubmed.ncbi.nlm.nih.gov/32203475/)
22. R. Wadey *et al.*, Isolation of a gene encoding an integral membrane protein from the vicinity of a balanced translocation breakpoint associated with DiGeorge syndrome. *Hum. Mol. Genet.* **4**, 1027–1033 (1995). doi: [10.1093/hmg/4.6.1027](https://doi.org/10.1093/hmg/4.6.1027); pmid: [7655455](https://pubmed.ncbi.nlm.nih.gov/7655455/)
23. S. Shifman *et al.*, A complete genetic association scan of the 22q11 deletion region and functional evidence reveal an association between DGCR2 and schizophrenia. *Hum. Genet.* **120**, 160–170 (2006). doi: [10.1007/s00439-006-0195-0](https://doi.org/10.1007/s00439-006-0195-0); pmid: [16783572](https://pubmed.ncbi.nlm.nih.gov/16783572/)
24. H. Ishiguro *et al.*, Replication study for associations between polymorphisms in the CLDN5 and DGCR2 genes in the 22q11 deletion syndrome region and schizophrenia. *Psychiatr. Genet.* **18**, 255–256 (2008). doi: [10.1097/YPG.0b013e328306c7dc](https://doi.org/10.1097/YPG.0b013e328306c7dc); pmid: [18797402](https://pubmed.ncbi.nlm.nih.gov/18797402/)
25. J. Pfleger, K. Gresham, W. J. Koch, G protein-coupled receptor kinases as therapeutic targets in the heart. *Nat. Rev. Cardiol.* **16**, 612–622 (2019). doi: [10.1038/s41569-019-0220-3](https://doi.org/10.1038/s41569-019-0220-3); pmid: [31186538](https://pubmed.ncbi.nlm.nih.gov/31186538/)
26. S. Liu, P. J. Anderson, S. Rajagopal, R. J. Lefkowitz, H. A. Rockman, G protein-coupled receptors: A century of research and discovery. *Circ. Res.* **135**, 174–197 (2024). doi: [10.1161/CIRCRESAHA.124.323067](https://doi.org/10.1161/CIRCRESAHA.124.323067); pmid: [38900852](https://pubmed.ncbi.nlm.nih.gov/38900852/)
27. K. Kanamaru *et al.*, Spatially resolved multiomics of human cardiac niches. *Nature* **619**, 801–810 (2023). doi: [10.1038/s41586-023-06311-1](https://doi.org/10.1038/s41586-023-06311-1); pmid: [37438528](https://pubmed.ncbi.nlm.nih.gov/37438528/)
28. J. P. Goetze *et al.*, Cardiac natriuretic peptides. *Nat. Rev. Cardiol.* **17**, 698–717 (2020). doi: [10.1038/s41569-020-0381-0](https://doi.org/10.1038/s41569-020-0381-0); pmid: [32444692](https://pubmed.ncbi.nlm.nih.gov/32444692/)
29. W. Feng *et al.*, Single-cell transcriptomic analysis identifies murine heart molecular features at embryonic and neonatal stages. *Nat. Commun.* **13**, 7960 (2022). doi: [10.1038/s41467-022-35691-7](https://doi.org/10.1038/s41467-022-35691-7); pmid: [36575170](https://pubmed.ncbi.nlm.nih.gov/36575170/)
30. A. L. Koenig *et al.*, Single-cell transcriptomics reveals cell-type-specific diversification in human heart failure. *Nat. Cardiovasc. Res.* **1**, 263–280 (2022). doi: [10.1038/s44161-022-00028-6](https://doi.org/10.1038/s44161-022-00028-6); pmid: [35959412](https://pubmed.ncbi.nlm.nih.gov/35959412/)
31. M. Asp *et al.*, A spatiotemporal organ-wide gene expression and cell atlas of the developing human heart. *Cell* **179**, 1647–1660.e19 (2019). doi: [10.1016/j.cell.2019.11.025](https://doi.org/10.1016/j.cell.2019.11.025); pmid: [31835037](https://pubmed.ncbi.nlm.nih.gov/31835037/)
32. E. N. Farah *et al.*, Spatially organized cellular communities form the developing human heart. *Nature* **627**, 854–864 (2024). doi: [10.1038/s41586-024-07171-z](https://doi.org/10.1038/s41586-024-07171-z); pmid: [38480880](https://pubmed.ncbi.nlm.nih.gov/38480880/)
33. L. Li *et al.*, An organ-wide spatiotemporal transcriptomic and cellular atlas of the regenerating zebrafish heart. *Nat. Commun.* **16**, 3716 (2025). doi: [10.1038/s41467-025-59070-0](https://doi.org/10.1038/s41467-025-59070-0); pmid: [40253397](https://pubmed.ncbi.nlm.nih.gov/40253397/)
34. B. Hu *et al.*, Origin and function of activated fibroblast states during zebrafish heart regeneration. *Nat. Genet.* **54**, 1227–1237 (2022). doi: [10.1038/s41588-022-01129-5](https://doi.org/10.1038/s41588-022-01129-5); pmid: [35864193](https://pubmed.ncbi.nlm.nih.gov/35864193/)
35. Y. Rosen *et al.*, Toward universal cell embeddings: Integrating single-cell RNA-seq datasets across species with SATURN. *Nat. Methods* **21**, 1492–1500 (2024). doi: [10.1038/s41592-024-02191-z](https://doi.org/10.1038/s41592-024-02191-z); pmid: [38366243](https://pubmed.ncbi.nlm.nih.gov/38366243/)
36. A. R. Pinto *et al.*, Revisiting cardiac cellular composition. *Circ. Res.* **118**, 400–409 (2016). doi: [10.1161/CIRCRESAHA.115.307778](https://doi.org/10.1161/CIRCRESAHA.115.307778); pmid: [26635390](https://pubmed.ncbi.nlm.nih.gov/26635390/)
37. D. A. Skelly *et al.*, Single-cell transcriptional profiling reveals cellular diversity and intercommunication in the mouse heart. *Cell Rep.* **22**, 600–610 (2018). doi: [10.1016/j.celrep.2017.12.072](https://doi.org/10.1016/j.celrep.2017.12.072); pmid: [29346760](https://pubmed.ncbi.nlm.nih.gov/29346760/)
38. K. C. Park, D. C. Gaze, P. O. Collinson, M. S. Marber, Cardiac troponins: From myocardial infarction to chronic disease. *Cardiovasc. Res.* **113**, 1708–1718 (2017). doi: [10.1093/cvr/cvx183](https://doi.org/10.1093/cvr/cvx183); pmid: [29016754](https://pubmed.ncbi.nlm.nih.gov/29016754/)
39. W. Joyce *et al.*, A revised perspective on the evolution of troponin I and troponin T gene families in vertebrates. *Genome Biol. Evol.* **15**, evac173 (2023). doi: [10.1093/gbe/evac173](https://doi.org/10.1093/gbe/evac173); pmid: [36518048](https://pubmed.ncbi.nlm.nih.gov/36518048/)
40. S. Yue *et al.*, Requirement of Smurf-mediated endocytosis of Patched1 in sonic hedgehog signal reception. *eLife* **3**, e02555 (2014). doi: [10.7554/eLife.02555](https://doi.org/10.7554/eLife.02555); pmid: [24925320](https://pubmed.ncbi.nlm.nih.gov/24925320/)

41. Y. Y. Lam *et al.*, APLNR marks a cardiac progenitor derived with human induced pluripotent stem cells. *Heliyon* **9**, e18243 (2023). doi: [10.1016/j.heliyon.2023.e18243](https://doi.org/10.1016/j.heliyon.2023.e18243); pmid: [37539315](https://pubmed.ncbi.nlm.nih.gov/37539315/)
42. N. A. Chapman, D. J. Dupré, J. K. Rainey, The apelin receptor: Physiology, pathology, cell signalling, and ligand modulation of a peptide-activated class A GPCR. *Biochem. Cell Biol.* **92**, 431–440 (2014). doi: [10.1139/bcb-2014-0072](https://doi.org/10.1139/bcb-2014-0072); pmid: [25275559](https://pubmed.ncbi.nlm.nih.gov/25275559/)
43. P. Sangwung *et al.*, KLF2 and KLF4 control endothelial identity and vascular integrity. *JCI Insight* **2**, e91700 (2017). doi: [10.1172/jci.insight.91700](https://doi.org/10.1172/jci.insight.91700); pmid: [28239661](https://pubmed.ncbi.nlm.nih.gov/28239661/)
44. J. S. Isenberg *et al.*, CD47 is necessary for inhibition of nitric oxide-stimulated vascular cell responses by thrombospondin-1. *J. Biol. Chem.* **281**, 26069–26080 (2006). doi: [10.1074/jbc.M605040200](https://doi.org/10.1074/jbc.M605040200); pmid: [16835222](https://pubmed.ncbi.nlm.nih.gov/16835222/)
45. A. Ghasemi, S. Jeddi, Quantitative aspects of nitric oxide production in the heart. *Mol. Biol. Rep.* **49**, 11113–11122 (2022). doi: [10.1007/s11033-022-07889-x](https://doi.org/10.1007/s11033-022-07889-x); pmid: [36107377](https://pubmed.ncbi.nlm.nih.gov/36107377/)
46. K. Huynh, Interferon response at the border zone of the infarcted heart. *Nat. Rev. Cardiol.* **21**, 740 (2024). doi: [10.1038/s41569-024-01078-y](https://doi.org/10.1038/s41569-024-01078-y); pmid: [39289539](https://pubmed.ncbi.nlm.nih.gov/39289539/)
47. K. R. King *et al.*, IRF3 and type I interferons fuel a fatal response to myocardial infarction. *Nat. Med.* **23**, 1481–1487 (2017). doi: [10.1038/nm.4428](https://doi.org/10.1038/nm.4428); pmid: [29106401](https://pubmed.ncbi.nlm.nih.gov/29106401/)
48. A. Krishnan *et al.*, A detailed comparison of mouse and human cardiac development. *Pediatr. Res.* **76**, 500–507 (2014). doi: [10.1038/pr.2014.128](https://doi.org/10.1038/pr.2014.128); pmid: [25167202](https://pubmed.ncbi.nlm.nih.gov/25167202/)
49. I. Esteban *et al.*, Pseudodynamic analysis of heart tube formation in the mouse reveals strong regional variability and early left-right asymmetry. *Nat. Cardiovasc. Res.* **1**, 504–517 (2022). doi: [10.1038/s44161-022-00065-1](https://doi.org/10.1038/s44161-022-00065-1); pmid: [39195950](https://pubmed.ncbi.nlm.nih.gov/39195950/)
50. K. Niederreither *et al.*, Embryonic retinoic acid synthesis is essential for heart morphogenesis in the mouse. *Development* **128**, 1019–1031 (2001). doi: [10.1242/dev.128.7.1019](https://doi.org/10.1242/dev.128.7.1019); pmid: [11245568](https://pubmed.ncbi.nlm.nih.gov/11245568/)
51. J. Xavier-Neto *et al.*, A retinoic acid-inducible transgenic marker of sino-atrial development in the mouse heart. *Development* **126**, 2677–2687 (1999). doi: [10.1242/dev.126.12.2677](https://doi.org/10.1242/dev.126.12.2677); pmid: [10331979](https://pubmed.ncbi.nlm.nih.gov/10331979/)
52. J. B. Moss *et al.*, Dynamic patterns of retinoic acid synthesis and response in the developing mammalian heart. *Dev. Biol.* **199**, 55–71 (1998). doi: [10.1006/dbio.1998.8911](https://doi.org/10.1006/dbio.1998.8911); pmid: [9676192](https://pubmed.ncbi.nlm.nih.gov/9676192/)
53. N. Rosenthal, J. Xavier-Neto, From the bottom of the heart: Anteroposterior decisions in cardiac muscle differentiation. *Curr. Opin. Cell Biol.* **12**, 742–746 (2000). doi: [10.1016/S0955-0674\(00\)00162-9](https://doi.org/10.1016/S0955-0674(00)00162-9); pmid: [11063942](https://pubmed.ncbi.nlm.nih.gov/11063942/)
54. Y. Tao *et al.*, Ptx2, an atrial fibrillation predisposition gene, directly regulates ion transport and intercalated disc genes. *Circ. Cardiovasc. Genet.* **7**, 23–32 (2014). doi: [10.1161/CIRCGENETICS.113.000259](https://doi.org/10.1161/CIRCGENETICS.113.000259); pmid: [24395921](https://pubmed.ncbi.nlm.nih.gov/24395921/)
55. C. Mosimann *et al.*, Chamber identity programs drive early functional partitioning of the heart. *Nat. Commun.* **6**, 8146 (2015). doi: [10.1038/ncomms9146](https://doi.org/10.1038/ncomms9146); pmid: [26306682](https://pubmed.ncbi.nlm.nih.gov/26306682/)
56. S. S. Ranade *et al.*, Myocardial reprogramming by HMGN1 underlies heart defects in trisomy 21. *Nature* **647**, 979–987 (2025). doi: [10.1038/s41586-025-09593-9](https://doi.org/10.1038/s41586-025-09593-9); pmid: [41125893](https://pubmed.ncbi.nlm.nih.gov/41125893/)
57. M. S. Simões-Costa *et al.*, The evolutionary origin of cardiac chambers. *Dev. Biol.* **277**, 1–15 (2005). doi: [10.1016/j.ydbio.2004.09.026](https://doi.org/10.1016/j.ydbio.2004.09.026); pmid: [15572135](https://pubmed.ncbi.nlm.nih.gov/15572135/)
58. V. M. Christoffels *et al.*, Chamber formation and morphogenesis in the developing mammalian heart. *Dev. Biol.* **223**, 266–278 (2000). doi: [10.1006/dbio.2000.9753](https://doi.org/10.1006/dbio.2000.9753); pmid: [10882515](https://pubmed.ncbi.nlm.nih.gov/10882515/)
59. W. Wang *et al.*, A single-cell transcriptional roadmap for cardiopharyngeal fate diversification. *Nat. Cell Biol.* **21**, 674–686 (2019). doi: [10.1038/s41556-019-0336-z](https://doi.org/10.1038/s41556-019-0336-z); pmid: [31160712](https://pubmed.ncbi.nlm.nih.gov/31160712/)
60. X. Huang *et al.*, Single-cell profiling of the developing embryonic heart in *Drosophila*. *Development* **150**, dev201936 (2023). doi: [10.1242/dev.201936](https://doi.org/10.1242/dev.201936); pmid: [37526610](https://pubmed.ncbi.nlm.nih.gov/37526610/)
61. A. Stolfi *et al.*, Early chordate origins of the vertebrate second heart field. *Science* **329**, 565–568 (2010). doi: [10.1126/science.1190181](https://doi.org/10.1126/science.1190181); pmid: [20671188](https://pubmed.ncbi.nlm.nih.gov/20671188/)
62. F. Darwin, On the structure of the snail's heart. *J. Anat. Physiol.* **10**, 506–510 (1876). pmid: [17231103](https://pubmed.ncbi.nlm.nih.gov/17231103/)
63. M. J. Wells, P. J. S. Smith, The performance of the octopus circulatory system: A triumph of engineering over design. *Experientia* **43**, 487–499 (1987). doi: [10.1007/BF02143577](https://doi.org/10.1007/BF02143577)
64. M. Lu *et al.*, Comparative analysis of the cardiac structure and transcriptome of scallop and snail, perspectives on heart chamber evolution. *Mar. Life Sci. Technol.* **5**, 478–491 (2023). doi: [10.1007/s42995-023-00202-0](https://doi.org/10.1007/s42995-023-00202-0); pmid: [38045548](https://pubmed.ncbi.nlm.nih.gov/38045548/)
65. B. Wu *et al.*, Single-cell analysis of the amphioxus hepatic caecum and vertebrate liver reveals genetic mechanisms of vertebrate liver evolution. *Nat. Ecol. Evol.* **8**, 1972–1990 (2024). doi: [10.1038/s41559-024-02510-9](https://doi.org/10.1038/s41559-024-02510-9); pmid: [39152328](https://pubmed.ncbi.nlm.nih.gov/39152328/)
66. Y. Li *et al.*, Origin and stepwise evolution of vertebrate lungs. *Nat. Ecol. Evol.* **9**, 672–691 (2025). doi: [10.1038/s41559-025-02642-6](https://doi.org/10.1038/s41559-025-02642-6); pmid: [39953253](https://pubmed.ncbi.nlm.nih.gov/39953253/)
67. M. Qi, J. Clark, E. R. R. Moody, D. Pisani, P. C. J. Donoghue, Molecular dating of the teleost whole genome duplication (3R) is compatible with the expectations of delayed rediploidization. *Genome Biol. Evol.* **16**, evae128 (2024). doi: [10.1093/gbe/evae128](https://doi.org/10.1093/gbe/evae128); pmid: [38913570](https://pubmed.ncbi.nlm.nih.gov/38913570/)
68. J. Piquereau, R. Ventura-Clapier, Maturation of cardiac energy metabolism during perinatal development. *Front. Physiol.* **9**, 959 (2018). doi: [10.3389/fphys.2018.00959](https://doi.org/10.3389/fphys.2018.00959); pmid: [30072919](https://pubmed.ncbi.nlm.nih.gov/30072919/)
69. M. L. Rexius-Hall *et al.*, A myocardial infarct border-zone-on-a-chip demonstrates distinct regulation of cardiac tissue function by an oxygen gradient. *Sci. Adv.* **8**, eabn7097 (2022). doi: [10.1126/sciadv.abn7097](https://doi.org/10.1126/sciadv.abn7097); pmid: [36475790](https://pubmed.ncbi.nlm.nih.gov/36475790/)
70. B. H. Davis *et al.*, Effects of myocardial infarction on the distribution and transport of nutrients and oxygen in porcine myocardium. *J. Biomech. Eng.* **134**, 101005 (2012). doi: [10.1115/1.4007455](https://doi.org/10.1115/1.4007455); pmid: [23083196](https://pubmed.ncbi.nlm.nih.gov/23083196/)
71. W. L. Rumsey, M. Pawlowski, N. Lejvardi, D. F. Wilson, Oxygen pressure distribution in the heart in vivo and evaluation of the ischemic “border zone”. *Am. J. Physiol.* **266**, H1676–H1680 (1994). doi: [10.1152/ajpheart.1994.266.4.H1676](https://doi.org/10.1152/ajpheart.1994.266.4.H1676); pmid: [8184947](https://pubmed.ncbi.nlm.nih.gov/8184947/)
72. B. N. Puente *et al.*, The oxygen-rich postnatal environment induces cardiomyocyte cell-cycle arrest through DNA damage response. *Cell* **157**, 565–579 (2014). doi: [10.1016/j.cell.2014.03.032](https://doi.org/10.1016/j.cell.2014.03.032); pmid: [24766806](https://pubmed.ncbi.nlm.nih.gov/24766806/)
73. N. H. Bishopric, Evolution of the heart from bacteria to man. *Ann. N. Y. Acad. Sci.* **1047**, 13–29 (2005). doi: [10.1196/annals.1341.002](https://doi.org/10.1196/annals.1341.002); pmid: [16093481](https://pubmed.ncbi.nlm.nih.gov/16093481/)
74. F. Liu *et al.*, MolluscDB 2.0: A comprehensive functional and evolutionary genomics database for over 1400 molluscan species. *Nucleic Acids Res.* **53**, D1075–D1086 (2025). doi: [10.1093/nar/gkai1026](https://doi.org/10.1093/nar/gkai1026); pmid: [39530242](https://pubmed.ncbi.nlm.nih.gov/39530242/)
75. J. A. Palmer, N. Rosenthal, S. A. Teichmann, M. Litvinukova, Revisiting cardiac biology in the era of single cell and spatial omics. *Circ. Res.* **134**, 1681–1702 (2024). doi: [10.1161/CIRCRESAHA.124.323672](https://doi.org/10.1161/CIRCRESAHA.124.323672); pmid: [38843288](https://pubmed.ncbi.nlm.nih.gov/38843288/)
76. H. Zhong *et al.*, Benchmarking cross-species single-cell RNA-seq data integration methods: Towards a cell type tree of life. *Nucleic Acids Res.* **53**, gkai1316 (2025). doi: [10.1093/nar/gkai1316](https://doi.org/10.1093/nar/gkai1316); pmid: [39778870](https://pubmed.ncbi.nlm.nih.gov/39778870/)
77. B. J. Haas *et al.*, De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat. Protoc.* **8**, 1494–1512 (2013). doi: [10.1038/nprot.2013.084](https://doi.org/10.1038/nprot.2013.084); pmid: [23845962](https://pubmed.ncbi.nlm.nih.gov/23845962/)
78. M. G. Grabherr *et al.*, Full-length transcriptome assembly from RNA-seq data without a reference genome. *Nat. Biotechnol.* **29**, 644–652 (2011). doi: [10.1038/nbt.1883](https://doi.org/10.1038/nbt.1883); pmid: [21572440](https://pubmed.ncbi.nlm.nih.gov/21572440/)
79. R. M. Waterhouse *et al.*, BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol. Evol.* **35**, 543–548 (2018). doi: [10.1093/molbev/msx319](https://doi.org/10.1093/molbev/msx319); pmid: [29220515](https://pubmed.ncbi.nlm.nih.gov/29220515/)
80. D. M. Emmis, S. Kelly, OrthoFinder: Phylogenetic orthology inference for comparative genomics. *Genome Biol.* **20**, 238 (2019). doi: [10.1186/s13059-019-1832-y](https://doi.org/10.1186/s13059-019-1832-y); pmid: [31727128](https://pubmed.ncbi.nlm.nih.gov/31727128/)
81. K. Katoh, D. M. Standley, MAFFT multiple sequence alignment software version 7: Improvements in performance and usability. *Mol. Biol. Evol.* **30**, 772–780 (2013). doi: [10.1093/molbev/mst010](https://doi.org/10.1093/molbev/mst010); pmid: [23329690](https://pubmed.ncbi.nlm.nih.gov/23329690/)
82. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLOS ONE* **5**, e9490 (2010). doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490); pmid: [20224823](https://pubmed.ncbi.nlm.nih.gov/20224823/)
83. D. Darriba *et al.*, ModelTest-NG: A new and scalable tool for the selection of DNA and protein evolutionary models. *Mol. Biol. Evol.* **37**, 291–294 (2020). doi: [10.1093/molbev/msz189](https://doi.org/10.1093/molbev/msz189); pmid: [31432070](https://pubmed.ncbi.nlm.nih.gov/31432070/)
84. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014). doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033); pmid: [24451623](https://pubmed.ncbi.nlm.nih.gov/24451623/)
85. Z. Yang, PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24**, 1586–1591 (2007). doi: [10.1093/molbev/msm088](https://doi.org/10.1093/molbev/msm088); pmid: [17483113](https://pubmed.ncbi.nlm.nih.gov/17483113/)
86. S. Kumar, G. Stecher, M. Suleski, S. B. Hedges, TimeTree: A resource for timelines, timetrees, and divergence times. *Mol. Biol. Evol.* **34**, 1812–1819 (2017). doi: [10.1093/molbev/msx116](https://doi.org/10.1093/molbev/msx116); pmid: [28387841](https://pubmed.ncbi.nlm.nih.gov/28387841/)
87. A. M. Bolger, M. Lohse, B. Usadel, Trimmomatic: A flexible trimmer for Illumina sequence data. *Bioinformatics* **30**, 2114–2120 (2014). doi: [10.1093/bioinformatics/btu170](https://doi.org/10.1093/bioinformatics/btu170); pmid: [24695404](https://pubmed.ncbi.nlm.nih.gov/24695404/)
88. A. Dobin *et al.*, STAR: Ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2013). doi: [10.1093/bioinformatics/bts635](https://doi.org/10.1093/bioinformatics/bts635); pmid: [23104886](https://pubmed.ncbi.nlm.nih.gov/23104886/)
89. H. G. Drost, A. Gabel, J. Liu, M. Quint, I. Grosse, myTAI: Evolutionary transcriptomics with R. *Bioinformatics* **34**, 1589–1590 (2018). doi: [10.1093/bioinformatics/btx835](https://doi.org/10.1093/bioinformatics/btx835); pmid: [29309527](https://pubmed.ncbi.nlm.nih.gov/29309527/)
90. Y. Hao *et al.*, Dictionary learning for integrative, multimodal and scalable single-cell analysis. *Nat. Biotechnol.* **42**, 293–304 (2024). doi: [10.1038/s41587-023-01767-y](https://doi.org/10.1038/s41587-023-01767-y); pmid: [37231261](https://pubmed.ncbi.nlm.nih.gov/37231261/)
91. W. P. Maddison, D. R. Maddison, MESQUITE: A modular system for evolutionary analysis (2009); <https://www.mesquiteproject.org>
92. P. O. Lewis, A likelihood approach to estimating phylogeny from discrete morphological character data. *Syst. Biol.* **50**, 913–925 (2001). doi: [10.1080/106351501753462876](https://doi.org/10.1080/106351501753462876); pmid: [12116640](https://pubmed.ncbi.nlm.nih.gov/12116640/)
93. J. Lever, M. Krzywinski, N. Altman, Principal component analysis. *Nat. Methods* **14**, 641–642 (2017).
94. S. Lê, J. Josse, F. Husson, FactoMineR: An R package for multivariate analysis. *J. Stat. Softw.* **25**, 1–18 (2008).
95. Y. Hao *et al.*, Integrated analysis of multimodal single-cell data. *Cell* **184**, 3573–3587.e29 (2021). doi: [10.1016/j.cell.2021.04.048](https://doi.org/10.1016/j.cell.2021.04.048); pmid: [34062119](https://pubmed.ncbi.nlm.nih.gov/34062119/)
96. Y. Zhou *et al.*, Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. *Nat. Commun.* **10**, 1523 (2019). doi: [10.1038/s41467-019-09234-6](https://doi.org/10.1038/s41467-019-09234-6); pmid: [30944313](https://pubmed.ncbi.nlm.nih.gov/30944313/)

97. M. I. Love, W. Huber, S. Anders, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8); pmid: [25516281](https://pubmed.ncbi.nlm.nih.gov/25516281/)
98. J. Feng *et al.*, GFOLD: A generalized fold change for ranking differentially expressed genes from RNA-seq data. *Bioinformatics* **28**, 2782–2788 (2012). doi: [10.1093/bioinformatics/bts515](https://doi.org/10.1093/bioinformatics/bts515); pmid: [22923299](https://pubmed.ncbi.nlm.nih.gov/22923299/)
99. A. J. Tarashansky *et al.*, Mapping single-cell atlases throughout Metazoa unravels cell type evolution. *eLife* **10**, e66747 (2021). doi: [10.7554/eLife.66747](https://doi.org/10.7554/eLife.66747); pmid: [33944782](https://pubmed.ncbi.nlm.nih.gov/33944782/)
100. W. K. Shen *et al.*, AnimalTFDB 4.0: A comprehensive animal transcription factor database updated with variation and expression annotations. *Nucleic Acids Res.* **51**, D39–D45 (2023). doi: [10.1093/nar/gkac907](https://doi.org/10.1093/nar/gkac907); pmid: [36268869](https://pubmed.ncbi.nlm.nih.gov/36268869/)
101. C. P. Cantalapiedra, A. Hernández-Plaza, I. Letunic, P. Bork, J. Huerta-Cepas, eggNOG-mapper v2: Functional annotation, orthology assignments, and domain prediction at the metagenomic scale. *Mol. Biol. Evol.* **38**, 5825–5829 (2021). doi: [10.1093/molbev/msab293](https://doi.org/10.1093/molbev/msab293); pmid: [34597405](https://pubmed.ncbi.nlm.nih.gov/34597405/)
102. N. Kumar, B. Mishra, M. Athar, S. Mukhtar, Inference of gene regulatory network from single-cell transcriptomic data using pySCENIC. *Methods Mol. Biol.* **2328**, 171–182 (2021). doi: [10.1007/978-1-0716-1534-8_10](https://doi.org/10.1007/978-1-0716-1534-8_10); pmid: [34251625](https://pubmed.ncbi.nlm.nih.gov/34251625/)
103. S. Jin, M. V. Plikus, Q. Nie, CellChat for systematic analysis of cell-cell communication from single-cell transcriptomics. *Nat. Protoc.* **20**, 180–219 (2025). doi: [10.1038/s41596-024-01045-4](https://doi.org/10.1038/s41596-024-01045-4); pmid: [39289562](https://pubmed.ncbi.nlm.nih.gov/39289562/)
104. L. Zhao *et al.*, Notch signaling regulates cardiomyocyte proliferation during zebrafish heart regeneration. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 1403–1408 (2014). doi: [10.1073/pnas.1311705111](https://doi.org/10.1073/pnas.1311705111); pmid: [24474765](https://pubmed.ncbi.nlm.nih.gov/24474765/)
105. S. Lian *et al.*, Cleaning scripts and codes for spatially-resolved single-cell atlas reveals the macro-evolutionary trajectory of animal hearts. Zenodo (2026); <https://doi.org/10.5281/zenodo.21199119>.

ACKNOWLEDGMENTS

We thank Z. Wang (Zhengzhou University) for providing zebra finch, X. Qi (Jinan University) for providing tropical clawed frog, Z. Gao (Ocean University of China) for providing amphioxus, H. Liu (Guangxi Academy of Marine Sciences) for providing horseshoe crab, and L. Lu (Ocean University of China) for providing the decompression chamber. We are grateful to M. Levine (Princeton University), L. Christiaen (University of Bergen), K. He (Guangzhou University), J.-W. Xiong (Nanchang University), L. Chen (Center for Excellence in Molecular Cell Science, Chinese Academy of Sciences), W. Wang (Northwestern Polytechnical University), K. Wang (Northwestern Polytechnical University), D. Larhammar (Uppsala University), and S. H. Church (New York University) for their extensive discussions and feedback. We acknowledge the support of the High-Performance Biological Supercomputing Center and core facilities in Institute of Evolution & Marine Biodiversity (IEMB) at the Ocean University of China for this research and China National GenBank (CNCB). **Funding:** This work was funded by Laoshan Laboratory (LSKJ202203204 to L.Z.); The National Key R&D Program of China (2023YFA1800600 to J.-F.F.); The National Natural Science Foundation of China (32130107 to S.W., 32170541 and 31970506 to L.Z., 32270870 to W.W., 31970475 and 32170832 to Y.S., 32441089 to L.B., 32222085 to Y.L.L., 32422088 to J.W., 82471753 to Q.L., 3240210275 to N.H., and 82271725 and 82471753 to S.L.P.); The National Key R&D Program of China (2021YFA0805000 to J.-F.F.); Guangdong Provincial Genomics Data Center (2021B1212100001 to G.F.);

Fundamental and Interdisciplinary Disciplines Breakthrough Plan of the Ministry of Education of China (JYB2025XDXIM801); The Fundamental Research Funds for Central Universities China (202012004 to L.Z., 862101225004 to W.W., and 202562018 to L.L.Z.); Beijing Natural Science Foundation (L232088); The State Key Laboratory of Medical Proteomics (SKLP-0202513); Major Program for the Fundamental Research of the Natural Science Foundation of Shandong Province (ZR2025ZD06); Shandong Provincial Natural Science Fund for Distinguished Young Scientists Fund Program (ZR2024JQ022); Key R&D Project of Shandong Province (2021ZLGX03 and 2022ZLGX01); Science and Technology Innovation Project of Laoshan Laboratory (LSKJ202203001); Blue Seed Industry Science and Technology Innovation Project (QDLYY-2024011); Outstanding Talent Team Project of Hainan Province (HNYT20240001); PI Project of Southern Marine Science and Engineering Guangdong Laboratory (GML20220018); Young Elite Scientists Sponsorship Program by CAST (2023QNRC001 to W.H.); and the Taishan Scholar Project Fund of Shandong Province of China. **Author contributions:** Conceptualization and funding acquisition: L.Z., G.F., W.W., S.W., J.-F.F., S.L.P.; Resources and sample collection and preprocessing: S.L., M.L., X.Z., S.L.P., Y.S., F.Z., P.Z., Y.H.Z., Z.Q., Y.F.L., N.H., X.C., Yi.L., Y.X.L., L.J., Y.P., Y.Q., S.G., J.Z., W.L., M.Z.; scRNA-seq experiments: S.S.P., Y.D.C., Q.S., J.S., Yu.L.; Stereo-seq experiments: X.L., R.L., X.Y.; Genome annotation and orthogroup assignment: W.H.; Software and computational analysis: S.L., M.L., Q.L., W.H., Y.X.L., J.L., Yu.L., J.W., R.Z., Yu.Z., Y.L.L., Le.L., R.H.Z.; Database construction: L.G., T.Y., J.C., Li.L.; Investigation: X.Z., M.L., Y.F.L., Y.C., Yi.L., X.D., P.W., S.Y., N.T.; Experiment supervision: L.Z., W.W., S.L., J.-F.F., Ya.Z., J.X.W., H.L.; Data interpretation: L.Z., S.W., W.W., S.L., J.-F.F., L.B., L.L.Z.; Visualization: S.L., M.L., Q.L., X.Z., W.H., Y.X.L., J.L.; Writing – original draft: W.W., S.L., M.L., S.W., L.Z., Y.S., W.H., G.F., J.-F.F., S.L.P., Q.L., J.L., Y.X.L.; Writing – review & editing: L.Z., W.W., S.W., S.L., J.-F.F., M.L., S.L.P., G.F., Y.S., Q.L., R.G.K., C.G.B., C.E.B., Y.G., X.X., H.Y., Z.B. The summary figure was prepared by S.L. and M.L. using Adobe Photoshop and Adobe Illustrator. **Competing interests:** The authors declare no competing interests. **Data, code, and materials availability:** All raw data generated by bulk RNA-seq, scRNA-seq and Stereo-seq have been deposited to the China National GeneBank (CNCB) Nucleotide Sequence Archive under accession no. CNP0006537. Additional data, including processed data generated by scRNA-seq and Stereo-seq and associated analysis protocols and software used in this study can be accessed at <https://db.genomics.cn/stomics/heart/>. The published human scRNA-seq data are available at <https://www.ebi.ac.uk/ena/browser/view/ERP123138>. The data analysis scripts are available to download at GitHub (<https://github.com/ouclian/EvoHeart>) and Zenodo ([105](https://doi.org/10.5281/zenodo.21199119)). The published scRNA-seq and Stereo-seq data of zebrafish are available in the NCBI database under accession no. PRJNA1233465. All additional information required to reanalyze the data reported in this paper is available in the supplementary materials or public repositories. Requests for further information can be directed to the lead contact L.Z. All materials used in this study are available upon request. Requests should be directed to corresponding author L.Z. **License information:** Copyright © 2026 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.adw0855

Figs. S1 to S29; Tables S1 to S24; MDAR Reproducibility Checklist

Submitted 21 January 2025; resubmitted 29 December 2025; accepted 6 July 2026; published online 23 July 2026

10.1126/science.adw0855



Spatially resolved single-cell atlas reveals the macroevolutionary trajectory of animal hearts

Shanshan Lian, Wei Wang, Yaoxin Liu, Meina Lu, Wentao Han, Jiaxin Luo, Xuejiao Zhang, Yue Pang, Yuru Li, Yufeng Li, Xin Dang, Peiyan Wang, Yating Qin, Ning Tan, Fang Zhang, Jin Zhang, Shuai Gao, Runyi Zhang, Lidong Guo, Lei Jiang, Pengchong Zhang, Yuehuan Zhang, Zhe Qu, Shanshan Pan, Yue Zhang, Naina Hu, Yadong Chen, Yun Cao, Wenbin Li, Xiaomei Chen, Junjie Shi, Shuaiqi Yang, Yiran Li, Qiuyun Song, Lisui Bao, Yuli Li, Xiawei Liu, Jing Wang, Lingling Zhang, Rui Li, Xin Yu, Tao Yang, Jing Chen, Ling Li, Mengqi Zhang, Lei Li, Ruihua Zhang, Hongyan Li, Ying Gu, Xun Xu, Huanming Yang, Zhenmin Bao, C. Geoffrey Burns, Caroline E. Burns, Robert G. Kelly, Jianxun Wang, Yang Zhao, Ying Su, Guangyi Fan, Qun Liu, Shi Wang, Ji-Feng Fei, Silin Pan, and Long Zhao

Science **393** (6815), eadw0855. DOI: 10.1126/science.adw0855

Editor's summary

The heart is a vital organ specialized for pumping blood, and early forms of the cardiovascular system have been around since at least the Early Cambrian Period. To understand the mechanisms driving the evolution of the heart, Lian *et al.* performed spatially resolved single-cell analysis on hearts from 27 different animal species such as mollusks, arthropods, fish, birds, and mammals, including humans. Through this analysis, the authors identified evolutionarily conserved elements as well as changes in gene expression patterns and cell composition that emerged with increased complexity of the heart during evolution. They also discovered a transitional cardiomyocyte state that appears to be stimulated by hypoxia and to promote cardiac regeneration. —Yevgeniya Nusinovich

View the article online

<https://www.science.org/doi/10.1126/science.adw0855>

Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2026 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works