Received: 17 August 2022 | Revised: 10 January 2023 | Accepted: 13 February 2023

DOI: 10.1002/jmor.21569

REVIEW

JOURNAL OF morphology WILEY

Evolution of homology: From archetype towards a holistic concept of cell type

Leonid Y. Rusin[1,2]

[1]Laboratory for Mathematic Methods and Models in Bioinformatics, Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Moscow, Russia

[2]EvoGenome Analytics LLC, Odintsovo, Moscow Region, Russia

Correspondence
Leonid Y. Rusin, Institute for Information Transmission Problems (Kharkevich Institute), Russian Academy of Sciences, Bolshoy Karetny per., 19, bld. 1, Moscow 127051, Russia.
Email: rusin@iitp.ru

Funding information
Foundation for Assistance to Small Innovative Enterprises

Abstract

The concept of homology lies in the heart of comparative biological science. The distinction between homology as structure and analogy as function has shaped the evolutionary paradigm for a century and formed the axis of comparative anatomy and embryology, which accept the identity of structure as a ground measure of relatedness. The advent of single-cell genomics overturned the classical view of cell homology by establishing a backbone regulatory identity of cell types, the basic biological units bridging the molecular and phenotypic dimensions, to reveal that the cell is the most flexible unit of living matter and that many approaches of classical biology need to be revised to understand evolution and diversity at the cellular level. The emerging theory of cell types explicitly decouples cell identity from phenotype, essentially allowing for the divergence of evolutionarily related morphotypes beyond recognition, as well as it decouples ontogenetic cell lineage from cell-type phylogeny, whereby explicating that cell types can share common descent regardless of their structure, function or developmental origin. The article succinctly summarizes current progress and opinion in this field and formulates a more generalistic view of biological cell types as avatars, transient or terminal cell states deployed in a continuum of states by the developmental programme of one and the same omnipotent cell, capable of changing or combining identities with distinct evolutionary histories or inventing ad hoc identities that never existed in evolution or development. It highlights how the new logic grounded in the regulatory nature of cell identity transforms the concepts of cell homology and phenotypic stability, suggesting that cellular evolution is inherently and massively network-like, with one-to-one homologies being rather uncommon and restricted to shallower levels of the animal tree of life.

KEYWORDS
cell-type evolution, developmental programme, gene regulation, morphotype, multicellularity, phenotypic stability, urmetazoan

1 | INTRODUCTION

Over 170 years ago, Richard Owen coined the term homology as 'identity of structure regardless of form and function' (Owen, 1848). The descendent Darwinian distinction between homology as structure and analogy as function is founded on this recognition and over a century has shaped the evolutionary paradigm of biology. The today's concept of body plan relies entirely on this distinction to place the diversity of complex organisms into a comparative framework in terms of their outward morphology (phenotype). On the way from Owen's rational archetype towards the cladistic ground plan (or ground pattern; Ax, 1984; Hennig, 1965; Scholtz, 2004; Yeates,

1995), the view of organismal structural identity has developed into a dynamic and versatile concept based on phylogenetic principles to capture a set of inheritable characters that describe a stem evolutionary lineage or crown group ancestor (Budd & Jensen, 2000). The identity of structure as a ground measure of evolutionary relatedness—and, hence, homology—has become the axis of comparative developmental biology as well, where the first and fundamental discoveries were made towards the understanding of universal patterns of embryogenesis. It was then recognized that a discrete animal body plan emerges only at a mid-embryonic stage of development, which was named the phylotypic stage (Duboule, 1994; Raff, 1996; Slack et al., 1993) to reflect the phylum-level conservation of morphological structures that are fated to become definitive organs or body regions not necessarily sharing the same function in adults. The recognition of deep structural homology of the embryo has driven the fields of comparative anatomy and embryology for decades towards the common truth that tissues and organs originating from the same primordium (Anlagen) are evolutionarily related (de Beer, 1954; Kowalevsky, 1866). At the more elementary level of the cell, its structural characteristics have traditionally been interpreted to define cell morphotype (Valentine, 2003). Observable morphotypes (like astrocyte, archaeocyte, rod, smooth muscle cell and so on) naturally delimited distinct cell types and were considered homologous and directly comparable across evolutionary lineages. Although the ancestry of individual cell types continued to be largely debatable, it was considered common between lineages and was largely linked to cell function.

The stance changed with the advent of single-cell genomics, whose technical innovations made it possible to directly measure and profile the genetic background of a phenotype at unparalleled, single-cell resolution. Recent advances in microfluidics, nucleic acid sequencing and bioinformatics have led to the emergence of high-throughput methods for estimating the transcriptome (a set of active/expressed/transcribed genes) in several thousand-to-million individual cells per experiment (Jaitin et al., 2014; Klein et al., 2015; Kolodziejczyk et al., 2015; Rao et al., 2021; Trapnell, 2015; Zheng et al., 2017). In contrast to traditional bulk transcriptomics, where gene transcripts are detected for a pool of cells to obtain sample-averaged estimates, single-cell approaches (mainly, single-cell RNA sequencing [scRNA-Seq]) supply an inventory to unmask molecular identity of individual cells within potentially heterogeneous cell populations of tissues, organs or a whole body (Shapiro et al., 2013; Stegle et al., 2015; Tanay & Sebé-Pedrós, 2021; Trapnell, 2015). This technological transition overturned the classical perception of cell identity and homology, with profound implications ensuing in our understanding of the general principles that underlie the evolution of cell, as well as ontogenesis and origin of multicellular life. This article succinctly summarizes the current progress and opinion in this field, and formulates a more generalistic view of biological cell types to highlight how the new logic grounded in the regulatory nature of cell evolution transforms the interpretation of cell-type identity, its evolutionary and ontogenetic transitions, changes the view of cell phenotypic stability, as well as elucidates the mechanisms of morphogenesis and body plan evolution, particularly in animals (Metazoa).

2 | CELL TYPES AS BASIC UNITS OF COMPARISON

A hallmark of multicellularity is the coexistence of a vast diversity of different cell types that implement different functions within a multicelled organism. This phenomenon had been emerging multiple times in the history of life, while true multicellularity only established in eukaryotes and marked the origins of several stem lineages—animals, fungi, land plants and green, red and brown algae (Burki et al., 2020; Knoll, 2011; Nedelcu, 2019; Niklas & Newman, 2013), thereby arguing that major selective advantages underlie this fundamental transition. Unlike temporary cellular aggregates in various protozoans, truly multicellular forms are clonal (e.g., Ros-Rocher et al., 2021). This means that all cell types in a multicelled colony are realized on the basis of a common genome via cell differentiation during successive divisions of the zygote. At least with animals, it is now well understood that major gene families driving true multicellularity had emerged already in their unicellular ancestors (e.g., Mikhailov et al., 2009; Richter et al., 2018; Ros-Rocher et al., 2021; Sebé-Pedrós et al., 2013) and preconditioned the primary attributes of a spatially stable colony—cell adhesion, communication and transdifferentiation (seamless, within cell-cycle switching of morphotypes). Therefore, the cell capacity to differentiate into cell types was a game-changing innovation that preceded and fuelled the milestone transition to complex multicellular architectures consisting of tissues and organs, whereas the cell types themselves comprise the natural building blocks that bridge the molecular and phenotypic dimensions, thus acquiring key significance as elementary units of organismal evolution. Throughout this text, 'coloniality' refers to the primary level of organization of animals as forms with true obligate multicellularity in the life cycle.

3 | MOLECULAR SIGNATURES OF CELL IDENTITY

For decades, cell types have been delimited phenotypically (i.e., as morphotypes) on the basis of their structural and physiological traits by means of microscopy and functional characterization. Therefore, the discovery, description, classification and evolutionary interpretation of cell diversity relied upon nonuniform (often study-specific) sets of nonstrictly formalized characters and character states, which introduced noticeable subjectivism in the cell type definition. The advent of genomics has made it possible to capture the genetic basis of cell phenotype and facilitated hypotheses claiming that discrete morphotypes are realized and maintained via co-ordinated activation of certain biomolecules that carry specific functions or subfunctions and interact to form morphotype-specific molecular machines or modules (Alberts, 1998; Hartwell et al., 1999; Pereira-Leal et al., 2006, 2007). In this view, cell–cell similarity can be elicited by juxtaposing the expression levels of individual genes—discrete, directly comparable biological entities, thereby rendering the comparison procedure more formalizable, testable and objective.

The molecular modules underlying cell identity are organized into complex gene regulatory networks (GRNs)—sets of interacting components (genes and their products, RNAs or proteins), which control the expression of target genes (e.g., Erwin & Davidson, 2009). On top of this regulatory hierarchy are genes encoding transcription factors (TFs, also termed *trans*-regulators or terminal selectors), the proteins that specifically recognize and bind short conserved strings in genomic DNA (TF-binding sites or motifs clustered in so-called enhancer and promoter regions, collectively also referred to as *cis*-regulators), whereby they form complexes with other proteins (other TFs and cofactors) and, ultimately, with the RNA polymerase II enzyme as part of the promoter-occupying basal transcriptional apparatus, enabling the latter to initiate transcription of downstream-regulated genes into their messenger RNAs; these target genes may also encode TFs (mediating their own GRNs or providing autoregulatory feedback) or end-function genes (also termed effector genes), which finally enact the observed cell morphotype-specific traits (structural and physiological; Hobert, 2008; Peter & Davidson, 2015; Spitz & Furlong, 2012; Wolpert et al., 2019).

The critical role of terminal selectors in dictating cell identity has been extensively documented and experimentally verified in various biological settings, including the conversion of fibroblasts to striated muscle cells (Davis et al., 1987), neural and ectoderm specification (Deneris & Hobert, 2014; Dillon et al., 2022; Leon et al., 2022), as well as the prominent involvement of homeodomain TFs (e.g., HOX proteins) in a plethora of biological processes during embryo- and organogenesis, axial and tissue patterning, limb formation and so on (e.g., Bürglin & Affolter, 2016). Many TFs predate the Metazoa in evolutionary origin (e.g., Brunet & King, 2017, 2022; de Mendoza & Sebé-Pedrós, 2019; Fairclough et al., 2013; Grau-Bové et al., 2017; López-Escardó et al., 2019; Sebé-Pedrós & de Mendoza, 2015; Sebé-Pedrós et al., 2016) and act deep within primal cell plasticity control mechanisms. It has been demonstrated that forced expression/activation or repression/disturbance of (sometimes single) key TFs is sufficient to alter cell identity, change cell-fate decisions along the differentiation route or even return the cell to totipotency, a baseline, early blastomere-like state with unrestricted potential to reproduce all cell lineages and give rise to a clonal embryoid (e.g., Amadei et al., 2022; de Silva et al., 2022; DuBuc et al., 2020; Graf & Enver, 2009; Lau et al., 2022; Minnoye et al., 2020; E. S. Wong et al., 2020).

This well-argued logic of GRN-driven cell identity specification implies the existence of 'signature genes' that maintain a cell type and distinguish it from other types at the level of gene expression. Among those, terminal selectors will serve as an apt proxy of active GRNs mediating the activation of various effector genes. Taking the reasoning that TFs are of an ancient age, and that regulatory connections are more evolutionarily constrained (vs. end-function gene usage), they are expected to comprise the bona fide genetic markers of evolutionary relatedness and, therefore, better elucidate cell-type homology. The combination of terminal selectors with more numerous effector genes, all of which exhibit cell type-specific (statistically correlated) expression, reflects the identity of a given cell type and can be compared between types to infer relatedness and interpret differences in the end-function complement of cell type-specifying genes. It is important to note that in cross-species comparisons, homology can only be inferred from the shared signature genes that comply with the orthology requirement, where each gene is represented by descendants of a gene lineage that vertically mirrors the evolutionary path of speciation (orthologues), excluding those of any parallel lineages resulting from gene family-specific duplications (paralogues). After putative homology has been established, the paralogous structure of signature genes can be interpreted between homologous cell types to explore various aspects of cell type evolution, including adaptive shifts in cell function or physiology driven by gene family-specific events like duplications or recombination/hybridization-associated transfers. Not straying beyond the scope into more technical details (e.g., see Tanay & Sebé-Pedrós [2021]), orthology identification remains a state-of-the-art of any analytic pipeline and benefits from the choice of more functionally and structurally conserved genes, such as TFs, other regulator or key effector genes.

The GRN-based view of cell identity has met with empirical appraisal at the level of single cells in the first comprehensive whole-body scRNA-Seq studies of bilaterian and nonbilaterian animals (e.g., J. Cao et al., 2017; Sebé-Pedrós, Chomsky, et al., 2018, Sebé-Pedrós, Saudemont, et al., 2018). Since their divergence in the Precambrian over half a billion years ago (Dohrmann & Wörheide, 2017), the sponges (Porifera), ctenophores (comb jellies; Ctenophora), placozoans (Placozoa), cnidarians (Cnidaria) and bilaterians (Bilateria) have inherited an extensive common suite of TFs, chromatin modifiers and remodellers, suggesting the existence of pan-metazoan (and older) regulatory mechanisms for orchestrating cell-type specification and maintenance (Putnam et al., 2007; Ryan et al., 2013; Srivastava et al., 2008, 2010). In scRNA-Seq assays, it was found that TFs are indeed much more cell-type specific (statistically significantly enriched) in their expression compared with all other genes, with each cell type being uniquely marked by some active TF(s). This situation is consistent with the GRN-based logic and underpins the existence of cell-type signature genes, among which the terminal selectors operate as drivers of cell identity and can be potentially used to discern cell types and their affinities. An appealing expectation would be that the cell-type boundaries derived from scRNA-Seq data inherently reflecting the activity of 'molecular machines' would also demarcate the observed cell morphotypes, thus establishing transparent homology of basic biological units—the cells—at the three basic levels of organization—genotype, regulation and phenotype. Instead, as explicated below, the fundamental discovery has been that this expectation is not fulfilled, thereby disclosing intrinsic gaps in our understanding of the biological nature of cell type and a demand for its conceptual rethinking.

4 | SHAKING THE OLD KNOWLEDGE: MORPHOTYPE DOES NOT DESCRIBE CELL IDENTITY

A revelation of the pioneering whole-body scRNA-Seq assays in animals was the inference of distinct, transcriptionally coherent clusters of signature gene-expressing single cells (single-cell

expression profiles) that correspond to morphologically indistinguishable cell types within the same organism. This implies that similar (or identical) cell morphologies may implement different functions and—potentially—have different origins.

Instantiating the findings in nonbilaterians, a sponge exhibited several transcriptionally distinct but phenotypically cryptic types of choanocytes, archaeocytes, pinacocytes and collagen cells; a placozoan—a cryptic diversity of peptidergic and ciliated epithelial cells; a ctenophore—genetically heterogeneous muscle, digestive and epithelial cells, diverse neuronal types with no coherent expression of canonical pan-neuronal bilaterian or cnidarian signature genes (despite having clearly neuronal morphologies; see Burkhardt & Jékely [2021] for an exceptional later finding of a bilaterian-like neuronal type from the same single-cell data), as well as the majority of types that could not be assigned to any known function, with many being strongly associated with unannotated, often ctenophore lineage-specific, proteins (Sebé-Pedrós, Chomsky, et al., 2018); a cnidarian—heterogeneity within gastrodermis, epidermis, gland/secretory cells and, especially, neurons, the latter cryptic repertoire being considered a lineage-specific innovation (Sebé-Pedrós, Saudemont et al., 2018).

The situation is similar with bilaterian animals. For instance, a relatively simple animal, the nematode, reveals heterogeneity of body wall muscles, intestine, sheath cells and, to a large extent, the neurons (J. Cao et al., 2017). As complexity increases towards vertebrates (craniotes), organisms begin to exhibit rich heterogeneity landscapes for most cell populations, as documented in high-resolution atlases of single-cell transcriptomes of tissues, organs, whole organisms or ontogenetic stages (embryonic, larval or organogenetic). The following references are intended to acquaint the interested reader with the recent single-cell evidence on animal cell diversity in both model and nonmodel species, which would not be appropriate to detail in a single review but is still citable in terms of the number of publications. The wealth of scRNA-Seq data already generated is expected to grow rapidly and already covers major animal lineages: Porifera (Musser et al., 2021; Sebé-Pedrós, Chomsky, et al., 2018), Ctenophora and Placozoa (Sebé-Pedrós, Chomsky, et al., 2018), Cnidaria (Chari et al., 2021; Hu et al., 2020; Levy et al., 2021; Sebé-Pedrós, Saudemont, et al., 2018; Siebert et al., 2019), Acoela (Duruz et al., 2021; Hulett et al., 2022), Plathelminthes (Fincher et al., 2018; Li et al., 2021; Plass et al., 2018), Annelida (Achim et al., 2018; Shao et al., 2020), Arthropoda (Allen et al., 2020; Brückner et al., 2021; Croset et al., 2018; Davie et al., 2018; Dillon et al., 2022; Hung et al., 2020; Karaiskos et al., 2017; Rust et al., 2020; Slaidina et al., 2020), Nematoda (J. Cao et al., 2017; Packer et al., 2019), Echinodermata (Foster et al., 2020; Massri et al., 2021; Meyer et al., 2022; Paganos et al., 2021), Urochordata (C. Cao et al., 2019; Horie et al., 2018; Sladitschek et al., 2020), as well as the most densely sampled Craniota (Briggs et al., 2018; J. Cao et al. 2019, 2020; Farrell et al., 2018; Han et al., 2020; Hodge et al., 2019; Pijuan-Sala et al., 2019; Regev et al., 2017; Shafer et al., 2022; The Tabula Muris Consortium, 2018; Tosches et al., 2018; D. E. Wagner et al., 2018).

Further ground-shaking insight into the evolutionary structure of 'phenotypic' cell types was obtained in the first systematic attempts to juxtapose 'molecular' cell types between phylogenetic lineages. There, a whole new array of connections between traditional and cryptic morphotypes were uncovered, where some reflected textbook views and others expressed never-expected similarities or disproved the expected ones.

Comparisons on a gross timescale of divergence between sponges, ctenophores, cnidarians, placozoans and bilaterians relied upon the notion that phenotype-specifying signature genes (especially, TFs, such as the HOX cluster) are often coregulated and positionally linked within stretches of genomic DNA (termed syntenies), which are preserved in animal genomes at a remarkable level of conservation (Engström et al., 2007; Irimia et al., 2012; Putnam et al. 2007, 2008; Simakov et al. 2013, 2015), with some being obviously inherited from even deeper, unicellular roots of the Metazoa (Simakov et al., 2022). By quantifying single-cell co-expression of genes strictly belonging to microsyntenic linkages largely conserved for over 600–650 Myr, sponge archaeocytes and pinacocytes were shown to be enriched in such 'relic expression' compared with other well-delimited cell types and to share some of strong and highly expressed linkages, thus suggesting a related and more ancient nature of these cell types (Zimmermann et al., 2019). Further, the expression pattern of a subset of linkages strictly preserved between the sponge, placozoan and cnidarian exposed a similarity between sponge choanocytes, placozoan peptidergic cells of undefined function and cnidarian digestive filaments.

The more ancient origin of archaeocytes also finds support at the level of comparison of sponges against unicellular animal relatives from the Holozoa lineage (Sogabe et al., 2019). Archaeocytes specifically upregulate the oldest, premetazoan genes mainly involved in cell proliferation, gene transcription and translation, similar to animal stem cells and proliferating stages in holozoans, including a colonial choanoflagellate. Choanocytes and pinacocytes, on the other hand, are enriched for markers of cell adhesion, signalling and polarity, in agreement with their role as epithelial cells. Opposing with the traditionally embraced homology of sponge choanocyte and the choanoflagellate (which lies in the crux of some textbook and modern hypotheses of transition to multicellularity; Arendt et al., 2015; Brunet & King, 2017; Metschnikoff, 1886; Nielsen, 2008; Remane, 1963; Ruppert et al., 2004; Salvini-Plawen, 2009), those were found to be the least similar, with choanocyte expressing mostly metazoan- and sponge-specific genes, suggesting its later origin and independently acquired (convergent) collar morphotype (see also Mah et al., 2014).

Remarkably, some lower-metazoan cell types revealed certain components of synaptic communication integral to the molecular machinery of bilaterian or cnidarian neurons. In sponges, presynaptic genes were found to be specifically co-expressed in a sensory contractile type of epithelial pinacocytes, as well as in a new type of multipolar secretory cells (originally named 'neuroid cells'), which were verified to possess vesicular apparatus and tightly enwrap choanocyte microvilli and cilia; choanocytes and apopylar cells, in turn, complement this 'proto-neuronal' system by co-expressing scaffolding and receptor proteins of the 'postsynapse' (Musser et al.,

2021; E. Wong et al., 2019). In placozoans, some lower-frequency cell types uniquely activate putative regulatory neuropeptides shown to be implicated in placozoan locomotion behaviours via ciliary beating control and to occur in elongated, vesicle-rich epithelial cells; immunostaining assays indicate that these cells co-express synaptobrevin, complexin and synaptophysin, the protein components of vertebrate synapses (Nikitin, 2015; Sebé-Pedrós, Chomsky, et al., 2018; Senatore et al., 2017; Smith et al., 2014). In cnidarians, nematocytes upregulate orthologues of a voltage-gated calcium channel and use it to control discharge response, thus revealing excitatory neuronal properties (Sebé-Pedrós, Saudemont, et al., 2018; Weir et al., 2020).

A recent systematic comparison of whole-body transcriptomic atlases between a sponge, cnidarian and five bilaterians—two flatworms (a tricladid planarian and a trematode) and three vertebrates (embryonic stages of a bony fish, an amphibian and mammal), which took a special algorithmic approach to account for the paralogous signature gene complement in cell types was able to reliably establish three clusters (families, with many-to-many similarity connections between members) within this metazoan cell diversity—neural, contractile and stem cells (Tarashansky et al., 2021). In consistency with the above mentioned findings, the neural family has united bilaterian and cnidarian neurons, vertebrate brain tissues, cnidarian nematocytes, as well as sponge choanocytes and apopylar cells.

The contractile family has included bilaterian myocytes, cnidarian myoepithelial cells known to possess contractile myofibrils (Buzgariu et al., 2015), as well as sponge pinacocytes and myopeptidocytes, both of which earlier implicated to be involved in sponge contractility (Musser et al., 2021; Sebé-Pedrós, Chomsky, et al., 2018). Notably, transcriptomic similarity between bilaterian and nonbilaterian contractile cell types extends beyond the core contractile system to also include orthologues of the adhesion complex that facilitates cell interactions, actomyosin networks that drive contractility, as well as some signalling pathways that mediate contraction. This finding indicates that these molecular modules, that is, their corresponding GRNs, were inherently co-ordinated to establish a contractile cell already in earliest metazoans.

The most compact (and fully interconnected) family has comprised invertebrate multipotent stem cells—flatworm neoblasts, cnidarian interstitial cells and sponge archaeocytes. Its signature orthology groups are enriched for genes involved in RNA processing, translation and posttranslational modification, as well as—but to a lesser extent—in cell cycle and DNA replication events, the latter two categories being rather more expected for actively dividing cells; several known stem cell-associated TFs, chromatin modifiers and remodellers have also been detected. Overall, these results suggest that totipotency may be largely conferred on stem cell at the deeply conserved level of epigenetic control, besides genome-level transcription regulation by TFs (Alié et al., 2015; Juliano et al., 2010; Lau et al., 2022; X. Wu et al., 2021; Xu et al., 2022).

Deep molecular roots of animal cell stemness also become elucidated from a comparison of whole-body single-cell atlases of two hydrozoan cnidarians that diverged over 200 Mya and represent different, polypoid and medusoid, metagenetic life cycle generations (Cazet et al., 2022). The study employed the cell type 'alignment' procedure developed in Tarashansky et al. (2021), to quantitatively match cell types between species, and used a phylostratigraphic approach (Domazet-Lošo & Tautz, 2010; Domazet-Lošo et al., 2007) to assign clade-specific ages to genes maintaining the cell types. Among the three hydrozoan embryonic cell lineages (germ layers), ecto-, endoderm and interstitial cells, the latter revealed highest transcriptomic conservation, with nearly each interstitial type finding a match in another species, in some cases with clear one-to-one similarity. High-potency interstitial types—germ cells, early neuron and nematocyte progenitors, and, especially, stem cells—were found to upregulate the oldest, premetazoan gene families, with the largest fraction of them shared between medusoid totipotent i-cells and polypoid i-cells and progenitor types. On the contrary, mature neuron and nematocyte types had either one-to-many or many-to-many patterns of cross-species similarity and were enriched in younger, metazoan- (neuron) or mostly hydrozoan-specific (nematocyte) genes. The same applies to both epithelial lineages: ecto- and endoderm appeared to be enriched for hydrozoan-specific genes, with ectoderm being the youngest layer with fewest cross-species matches. Despite dramatic differences in epithelial morphologies between the polyp and medusa, however, those do share some epithelia-specific gene families, thus suggesting common origin of their underlying molecular machinery. Generally, this situation discloses that more ancient and affine progenitor cells produce evolutionarily younger and more genetically heterogeneous cells, which may or may not have a direct match (one-to-one homologue) in another species, largely irrespective of morphotype, even at a moderate evolutionary distance between two hydrozoans.

The existence of several genetically related families among the diversity of animal cell types also becomes clear from the first attempt to build a proxy of cell type phylogeny as a dendrogram of pair-wise cell-type expression profile similarities for strictly orthologous genes within and across species (Wang et al., 2021). The study covered seven representatives of the Planulozoa (the group uniting Cnidaria and Bilateria)—a cnidarian, flatworm, nematode, tunicate and three vertebrates, a bony fish and two mammals—to reveal that neurons and muscles (striated and smooth, separately) form the most compact clades in terms of uniting traditional morphotypes across species, in contrast to the most heterogeneous epithelial and stromal types, thus suggesting strong evolutionary conservation of neural and contractile GRNs. Smooth muscles exhibit relatedness to a subset of stromal types. This is consistent with the single-cell-based evidence that a planarian flatworm muscle also acts as a vertebrate connective tissue (especially fibroblasts) via extracellular matrix secretion and activation of key patterning pathways for maintaining tissue architecture and regeneration control, which promoted a hypothesis of common origin of muscle and stromal cells from a mesodermal cell type that combined contractile, secretory and patterning activities (Cote et al., 2019).

Cnidarian gastrodermis, flatworm phagocytes and nematode pseudocoelomocytes reveal similarity to vertebrate immune cells to form a dense 'immune' clade, thus suggesting common origin of immunity and digestion (Hartenstein & Martinez, 2019). This evidence fully recapitulates the classical Metchnikoff's phagocytic

theory of immunity, where he posited that the roots of cellular immunity lie in the common origin of macrophages and gastric cells of lower metazoans (Metchnikoff, 1901; Metschnikoff, 1884). The tunicate's membership in this clade for the first time predicts an immune system in urochordate larvae.

The variety of observed cross-clade connections presents an intricate evolutionary pattern behind cell morphotype relationships, which is yet to be disentangled and interpreted (see Section 7 for a discussion on cell—cell comparability). Such are the similarities between some cryptic flatworm parenchymal types and immune cells; vertebrate fetal stroma and proliferating cells (rather than adult stroma); and so on. Complementing those is the recently reported distinct bilateral single-cell-level heterogeneity of embryonic coelom in a sea star, which unexpectedly exposes no equivalents of the left-side pouches to any of the embryonic cell types in a sea urchin, at the same time showing strong affinity between the right-side pouches and the urchin's special primary mesenchyme cells (Meyer et al., 2022). Meanwhile, the left embryonic coelom is known in all echinoderms to give rise to hydrocoele, a water vascular system of the adult animal (Harrison & Ruppert, 1991).

Of special notice in this section is the commonly seen genetic heterogeneity of cell types, with many, if not most, of them displaying one- or many-to-many similarity connections across other types within and between species. This means that certain portions of their underlying GRNs become shared by cell types via particular evolutionary mechanisms, whereas, at least, the 'many-to-many' structure of similarity cannot be explained by linear evolutionary divergence of cell types with time. A marginal illustration are epi- and endothelia (Cazet et al., 2022; Wang et al., 2021): these cell types are highest derived in terms of both their underlying gene content and age, display fewest one-to-one cross-species similarities and many inter-type connections, suggesting their role as a major reservoir of cell type innovation. The epithelial 'phylogeny' has been particularly demonstrated to have a more nonbinary, network-like structure, where strongest connections across phyla do not span continuous paths in the species tree, thus precluding its explanation by vertical evolutionary descent. Genetic heterogeneity inherently marks novel cell types, as can be illustrated, for instance, with a report on sea urchin's larval pigment cells: those are traditionally considered a morphological and functional novelty of echinoid echinoderms and also form a distinct cluster in the urchin's single-cell atlas, while they turn to comprise a 'hybrid' of immune mesenchyme and neuronal-type GRNs in comparison with a sea star (Meyer et al., 2022). The salient examples of cell type novelty emerging via assembly of different labour-executing components can be observed among the exceptional diversity of exocrine glands in insects (Brückner et al., 2021; Kishi & Parker, 2021).

5 | DEMISE OF CLASSIC ANLAGEN: CELL LINEAGE IS NOT EQUIVALENT TO EVOLUTIONARY DESCENT

The notion that differentiation of phenotypically (i.e., morphologically and functionally) similar or identical cells is not strictly governed by ontogenetic time course or a cell-lineage precursor has been around for decades, while always provoking controversy and debate in various contexts in developmental and evolutionary science. This lability is well documented even for relatively compact ontogenetic cell lineages in a eutelic organism, the nematode *Caenorhabditis elegans*, where all cell fates are precisely known: for example, as with same-lineage differentiation of phenotypically distinct muscles and neurons (Sulston et al., 1983). More complex tissue-specific lineages in vertebrates (Emerson et al., 2013) and arthropods (Bayraktar & Doe, 2013) are also known to produce markedly different (in terms of both structure and regulator gene expression) retinal cell types or neuronal and glial cells from common progenitors, respectively.

Further, equivalent cell phenotypes can derive from ontogenetic sublineages as distant as different embryonic germ layers. A good illustration is the de novo embryonic development of some of the pharyngeal neurons from endoderm in the sea urchin, despite neural cells being well-documented ectodermal derivatives (Wei et al., 2011). This 'ectopic' neuronal specification was experimentally justified by detecting pluripotency factor expression in the foregut, which persists even after normal endoderm specification by a well-studied GRN and maintains a downstream neural GRN, thus allowing for the formation of neuroendoderm. A more recent study of bilaterian ultrafiltration excretory organs showed that all basic components of proto-, metanephridia and the kidney—from podocytes and flame cells to pore cells—share a conserved set of active TFs and TF-mediated end-function genes, irrespective of the component's origin from ecto- or mesoderm (Gąsiorowski et al., 2021). Besides, grounding in the presence of the Pax-Six-Eya TF network with a well-studied role in sensory organ development (Fortunato et al., 2014), as well as the activity of virtually all of the nephridial TFs also present in otic vesicle, sensory hair or support cells of the inner ear, the entire nephridial complex has been suggested to constitute an evolutionarily sister subfamily to otic mechanosensory hair cells, with their common ancestor hypothetically traced back to an ectodermal cell type combining water propulsion, mechanosensation, reabsorption and ultrafiltration functions associated in modern morphotypes with motile cilia and microvillar collar (Arendt, 2021).

In nondeterministic development, cell fates in germ layers are not strongly restricted. This follows from classical transplantation experiments in mice, showing that transplanted cells adopt the fates of a recipient germ layer (Tam & Gad, 2004), as well as from research into basal animals, which exhibit no progressive fate determination in postgastrulation layers, with cells being able to readily transdifferentiate between types (Nakanishi et al., 2014; Sogabe et al., 2019). Eventually, any potency for extensive regeneration would not be feasible without breaking cell lineage boundaries by precursor, pluripotent or stem cells during their differentiation towards the target tissue cell types.

Most of the studies in evolutionary developmental biology (evo–devo) rely on the molecular line of evidence to verify or detect evolutionary affinity between differentiating cells and interpret their homology. Recognizing that special experimental settings and targeting of particular biological objects are required for generating

a result, one should admit that these studies have only scratched the surface of the real evolutionary complexity of ontogenesis. Today, the emerging high-throughput single-cell and computational analytic techniques are paving the way for systematic discovery of cell homology on a tissue- or whole-organism scale, including cases of nonsuspected relatedness. For instance, such is the detected affinity between some of the endodermal cells in zebrafish and small secretory cells of ectodermal epithelium in clawed frog (Tarashansky et al., 2021), with the latter comprising a not-so-long discovered morphological and functional cell phenotype with a vital immune activity in the frog's embryonic ectoderm (Dubaissi et al., 2014). This affinity was inferred from the presence of numerous shared signature genes, including conserved TFs and effector genes involved in vesicular protein trafficking, which all turned to have well-documented roles in the specification of secretory cells.

Based on the foregoing, the knowledge grasped from earlier studies and augmented with today's single-cell-based evidence spells the demise of the germ-layer concept of homology from comparative anatomy and embryology (de Beer, 1954; Kowalevsky, 1866) and exposes a demand for new and holistic evolutionary concepts.

6 | AN EVOLUTIONARY CONCEPT OF CELL TYPE

The corpus of modern evidence on evolutionary and ontogenetic relations between animal cell types reveals a situation not envisaged ever before: cells may or may not be related irrespective of both their phenotype and cell-lineage origin—within an individual and across species, similar morphotypes may appear nonrelated, whilst morphologically and developmentally (also at a germ-layer level) distinct cells may exhibit relatedness. The level at which cell–cell similarity becomes evident in an objective and testable way is that of gene expression, and it crosses through the planes of phenotypic and developmental similarities. This new fact of life renders the Owen's definition of homology at the basic level of the cell an elusive concept, as identity is to be defined here regardless of all—structure, form or function (and development), the scenario that needs to be addressed from a theoretical standpoint.

6.1 | Regulatory isolation drives cell-type evolution

Clonally multicellular organisms, such as Metazoa, reproduce the colonial body each time anew from a single cell, with all cell types being necessarily rebuilt within lifetime on the basis of the same genome. Naturally, specification of each cell type is to be defined by selective expression, regulation and interaction of key molecular modules, which realize and maintain a cell state. Simplistically, this selectivity means that certain portions of the common genome become accessible or inaccessible to the cell transcription machinery to process cell type-specific genetic information at a given timepoint in a given cell neighbourhood. A succession of such spatio-temporally cohered events drives the formation of a spatially stable colony through time and is referred to as the *developmental programme*. Obviously, precise deployment of genomic instructions is primal to achieving organismal integrity, as well as avoiding survival threats of the exploitation by selfish cell types with altered development, as occurs in cancer (de Silva et al., 2022; Liang et al., 2015; Libby & Ratcliff, 2014), which imputes a principal role to gene regulation in enacting the whole programme.

The spectacular regularity of cell differentiation and tissue patterning observed in animal development, especially during embryo- and organogenesis (e.g., Arthur, 1997; Kowalevsky, 1866; Martin-Duran & Marletaz, 2020; Richardson & Keuck, 2002; Temereva & Malakhov, 2012), supplemented with the growing genomic and transcriptomic evidence on molecular markers and specifiers of differentiated tissues and isolated cell populations (e.g., Alié & Manuel, 2010; Arendt et al., 2004; Hirano et al., 2013; Lamb, 2013; Riesgo et al., 2014; Royuela et al., 2000; Ryan et al., 2013; Seipel & Schmid, 2005; Steinmetz et al., 2012; Yanay et al., 2008), have guided conceptual views to propose that a stable, recurrent cell phenotype is uniquely defined by its molecular fingerprint, a set of co-expressed genes belonging to coregulated functionally cohesive molecular modules (Achim & Arendt, 2014; Arendt, 2005, 2008)—a statement largely confirmed from single-cell data, as well as to formulate a mechanistic model to explain the evolution of this fingerprint via changes in its regulatory 'core'.

This core regulatory complex (CoRC) is viewed as a set of co-interacting TFs, their cofactors and implicated genomic *cis*-regulatory elements, which control and determine cell-type-specific gene expression (Arendt et al., 2016, 2019). Thereby, a CoRC resembles a GRN (elementary or consisting of nested, downstream-regulated GRNs) mediating a set of end-function genes. Importantly, this GRN shows only partial (or no) overlap in components with other-cell GRNs, thus conferring the cell type an identity and (quasi-) independence to evolve separately from other cells in the process of genetic individuation (G. P. Wagner, 2014) via selection acting on this GRN-specific phenotypic traits. The end-function gene modules coregulated by the same CoRC constitute the cell type apomeres (by analogy to cladistic apomorphies defining newly acquired, clade-specific traits). When evolutionary change impacts apomeres but not the CoRC, it drives the diversification of cell type into subtypes, all of which comprise the bona fide homologous members of one cell-type family, irrespective of concurrent phenotypic divergence (underlain by the diverging downstream-regulated apomere genes). On the other hand, changes in the CoRC lead to the diversification of novel, sister cell types (Arendt, 2008), which themselves comprise bona fide homologues with respect to each other. Such changes may occur via sequence divergence, duplication or loss of individual TFs, as well as their co-option (recruitment) from other-cell-type CoRCs, the latter process resulting in recombination (hybridization) of different CoRCs and the origin of novel cell types by fusion (Arendt et al., 2016; Oakley, 2017; Schlosser, 2021).

A CoRC-based model provides an elegant conceptual definition of the cell type as a population of cells separated from other cells in a

multicellular colony via regulation-determined, selective access to common genomic information. By analogy to the speciation process, which usually occurs via reproductive isolation of entire genomes, cell types evolve within a cellular colony via regulatory isolation of genomic regions involved in the deployment of cell-type-specific gene expression pattern. The actual molecular regulators of gene transcription, therefore, define cell type identity, whereas evolutionary change in this regulatory core drives the birth and divergence of cell types, determines their genealogy and, hence, innate homology. By drawing the line between the regulatory and end-function complements of the cell type's active gene kit, the model mechanistically decouples cell type identity from phenotype, essentially allowing for the divergence of related morphotypes beyond recognition. In terms of the molecular mechanism, a succession of cell-fate decisions in ontogenesis is not equivalent to the evolutionary individuation and diversification of cell types, which explicitly decouples the developmental cell lineage from cell type phylogeny.

6.2 | Epigenetic control is integral to deployment and evolution of cell types

Ontogenesis, especially early development, is known to critically rely upon epigenetic control of gene expression (e.g., Bird, 2002; Cedar & Bergman, 2012; Chen et al., 2022; Du et al., 2022; Mariño-Ramírez et al., 2005). This regulation layer is, probably, at least as archaic as is TF-mediated regulation based on genome-fixed features (i.e., binding specificity of enhancer or promoter DNA primary sequences), whereas it exerts control in a dynamic, reversible fashion over the cell or colony's lifespan via a variety of nonmutational mechanisms acing at the genetic- (DNA methylation), chromatin- (histone modifications, commonly via acetylation and methylation; higher-order spatial chromatin remodelling) or posttranscriptional (non-coding RNA-mediated processes) levels. During DNA methylation and histone modification, distinct epigenetic marks are enzymatically introduced as modest chemical modifications on DNA and chromatin's histone proteins, respectively, which regulate the activation or repression of downstream genes via modulation of the intermolecular interactions between the DNA strands and the protein transcription machinery. This epigenetic landscape controls the responsiveness of TF-binding sites and, hence, their ability to regulate gene transcription. Proper remodelling or complete resetting of the landscape stand as the major factors of epigenetic reprogramming, which enables the cell to enter toti- or pluri(multi-)potency and safeguards the colony's proper subsequent differentiation during development. In vertebrates, this reprogramming is essential to convert terminally differentiated gametes to the totipotent zygote producing an embryo, whereas totipotency itself can be captured and maintained in vitro from embryonic cells by sustained chemical inhibition of just a few key enzymes, mainly chromatin modifiers implicated in gene repression via histone methylation and deacetylation (Xu et al., 2022). Recent cellular transcriptomic evidence reveals chromatin modifiers among the markers of animal stem cells on a broad evolutionary timescale, which suggests a pan-metazoan epigenetic control of cell stemness (see above; Cazet et al., 2022; Sogabe et al., 2019; Tarashansky et al., 2021). It is becoming clear that epigenetic modifications are universally employed by most animals and beyond for executing deeply conserved regulatory functions, including expression control of evolutionarily old genes, as well as that these marks play vital evolutionary roles by establishing epigenetic memories inheritable across generations (Keller et al., 2016; Vogt, 2022; Zemach et al., 2010). In jawed vertebrates (at least osteognathostomes), parental epigenetic settings are reset in the starting embryo via global DNA rehypermethylation or active local methylation of enhancer regions, which leads to their 'dememorization' and prevents premature, ectopic (and fatal to the embryo) firing of adult, tissue-specific enhancers and genes in early embryogenesis (X. Wu et al., 2021). Widespread, active enhancer demethylation only occurs later, during the phylotypic stage of development, when the vertebrate body plan and definitive organs are shaped, and is almost exclusively targeted towards genes with well-established critical roles in morphogenesis, including key developmental pathways, such as Wnt, Notch and transforming growth factor-β (Bogdanović et al., 2016). This wave of epigenetic remodelling coincides with the major transcriptomic and morphological transitions during the most conserved, phylotypic stage of ontogenesis (Drost et al., 2017; Levin et al., 2016; Martín-Zamora et al., 2023; Uesaka et al., 2022; L. Wu et al., 2019), thus suggesting the existence of an, at least, pan-vertebrate regulatory logic based on DNA methylation as a primal, upstream regulator of phylotypic enhancer activity.

Importantly, DNA methylation and demethylation, as well as histone modifications, exhibit clear signs of their genomic target sequence specificity, meanwhile being implemented via distinct enzyme systems. During enhancer dememorization, embryonic enhancers become less methylated and remain operational likely due to their lower CG-content versus adult enhancers, which essentially enables an 'epigenetic gate' to separate embryonic and adult gene transcription in development (X. Wu et al., 2021). Specifically, demethylated phylotypic enhancers display higher evolutionary conservation of primary sequence versus early (blastula or gastrula) or late (adult stages) enhancers (Bogdanović et al., 2016). This comes in line with the evidence showing that transcription of major developmental genes is regulated by poised enhancers, which are genetically distinct from other distal *cis*-regulatory elements and become specifically bookmarked in progenitor cells with unique epigenetic features that could contribute to their privileged regulatory properties during subsequent differentiation. These enhancers reveal high sequence conservation for a number of vertebrate clades, and some are evolutionarily conserved across all vertebrates (Crispatzu et al., 2021).

The dependence of epigenetic control on sequence-embedded genomic features clearly suggests its involvement in the process of genetic individuation as an important factor affecting regulatory isolation of incipient cell types. The epigenetic layer of regulation not only switches enhancers on or off in a stochastic, nonselective manner, but also specifically affects particular sequence contexts,

thus changing the interplay between active terminal selectors and their temporal co-ordination in the cell. A phenotypic (functional) product of such variation will be subject to natural selection and, hence, will exert a feedback impact on the relevant enhancers' sequence composition and syntax by favouring certain conditions, whereby converting epigenetic effects into inheritable genetic properties, which can be fixed in evolution by a population of cells diverging into a cell type. The same logic applies to promoter regions, eventually making epigenetics an inextricable component of the organism's developmental programme, with an imprint in genomic and cell type evolution.

6.3 | From convention to reality: Cell types as avatars of omnipotent cell

The epigenetic and genomic layers of gene regulation act via radically distinct mechanisms but complement each other in an intricate, agonistic system to shape cell types during both ontogenesis and evolution. Current advances in cellular reprogramming and comparative transcriptomics, briefly outlined above, clearly disclose concrete molecular players enacting cell differentiation in both regulation layers and demonstrate the spectacular plasticity of the cell—its ability to morph between phenotypes, as well as molecular identities, which may or may not be evolutionarily related. During normal differentiation of totipotent early-stage blastomeres towards terminally committed functional somatic cells, many kinds of cells can undergo reprogramming and dedifferentiate into an intermediate, progenitor-like state (completely or partially, adopting a stem- or only immature condition), whereby they can re-enter the cell cycle and replenish the supply of other functional cells, for instance, lost upon injury (Arthur-Farraj et al., 2012; Brawley & Matunis, 2004; Jopling et al., 2010; Kusaba et al., 2014; Rompolas et al., 2013; Sheng et al., 2009; Tata et al., 2013). In basal metazoans, the sponges, cells routinely transdifferentiate between types with maximal transcriptomic (hence, evolutionary) disparity: for example, evolutionarily derived choanocytes exist in a metastable state for a few hours only and morph back into multipotent, ancient archaeocytes at no prior cell division, the latter transdifferentiating into a range of other cell types; juvenile epithelial cells also use an archaeocyte intermediate to produce multiple other cells 'bridging' the sponge germ layers (Nakanishi et al., 2014; Sogabe et al., 2019). Any stem cell is capable of self-renewal and acts like a hub that connects various sublineages and routes of differentiation, with its degree of potency becoming gradually restricted and more tissue-specific as more intermediate identities are committed along the route from the totipotent zygote. On the other hand, one mature cell can convert into another mature cell directly, without reprogramming into an intermediate. Direct conversion is widely employed in vivo even in vertebrates, at least, in response to tissue damage or other physiological stress (Tarlow et al., 2014; Thorel et al., 2010; Yanger et al., 2013) and was shown to be readily attainable in a range of experiments with genetic and epigenetic factor induction (reviewed in Ladewig et al., 2013), suggesting that it is commonplace in simpler metazoans. Terminally differentiated gametes are reset, genetically and/or epigenetically, to fully regain totipotency before embryonic genome activation in all animals reproducing via germline.

The remarkable capacity of the cell to seamlessly morph between identities of various evolutionary age and origin, changing guise and function, coupled with knowledge of the molecular mechanics behind cell transformations in development, regeneration and evolution, lead us to recognize the classical, intuitive cell types (like neuron, muscle, stem cell and so on) as nothing more than a convention. Essentially, biological cells can be thought of as *avatars* (in Hinduism, avatar is the manifestation of a deity in animal form; figuratively, an archetype, an embodiment of an abstract concept), transient or terminal states deployed in a continuum of states by the developmental programme of one and the same *omnipotent cell*.

In 1957, Conrad Waddington introduced an elaborated version of his 'epigenetic landscape', a theoretical framework to describe the terrain of developmental choices negotiated by embryonic tissues on their routes to become a mature embryo (Waddington, 1957). Over decades since then, the Waddington diagram has been repurposed to describe cell-fate decisions and has become a widely used metaphor for thinking about hierarchical lineage segregation and determination in vivo and in vitro. The term 'epigenetic landscape' was originally conceived to separate the 'genetic' component acting in evolution and moulding the landscape from the choices taken in development when the landscape is fixed. Past half a century, this designation has literally acquired the new meaning, however, in a new context that integrates evolution with development and implicates that hierarchical and unidirectional thinking is not fully applicable to development and even less so—to evolution at the level of the cell.

The diagram on Figure 1 illustrates a more realistic concept, which centres around not the hierarchy of fate but the deployment of cell types in ontogenesis and their origin in evolution. Although it is challenging to pictorially represent the organism's developmental programme, the co-ordinated dynamics of gene transcription lies in the crux of cell metamorphoses, both in ontogenetic and evolutionary dimensions. Generally, this programme is a modular hierarchy of segments, where an elementary segment comprises a regulator–effector(s) linkage and can itself be part of a higher-rank, upstream regulation segment; the segments are interlinked by regulatory (positive, negative, bi- or unidirectional) connections into a complex, modular regulatory network which changes dynamically with time as new regulators come into play (e.g., by signal induction) or vanish (e.g., by factor depletion). A cell type is, therefore, an instantaneous product of gene transcription realized by the developmental programme in a multidimensional vector space or, simplistically, on a landscape of permissible cell states. Peaks on this landscape correspond to stable cell states (distinct cell types deployed by time-steady regulation modules) and the valleys—to intermediate ones, which the cell needs to cross to reach other stable states (transient cell types deployed by regulation modules in a process of their rewiring during differentiation, reprogramming or direct conversion). The whole colony formation is, therefore, described by a manifold of trajectories that the developmental programme is able to travel between the peaks

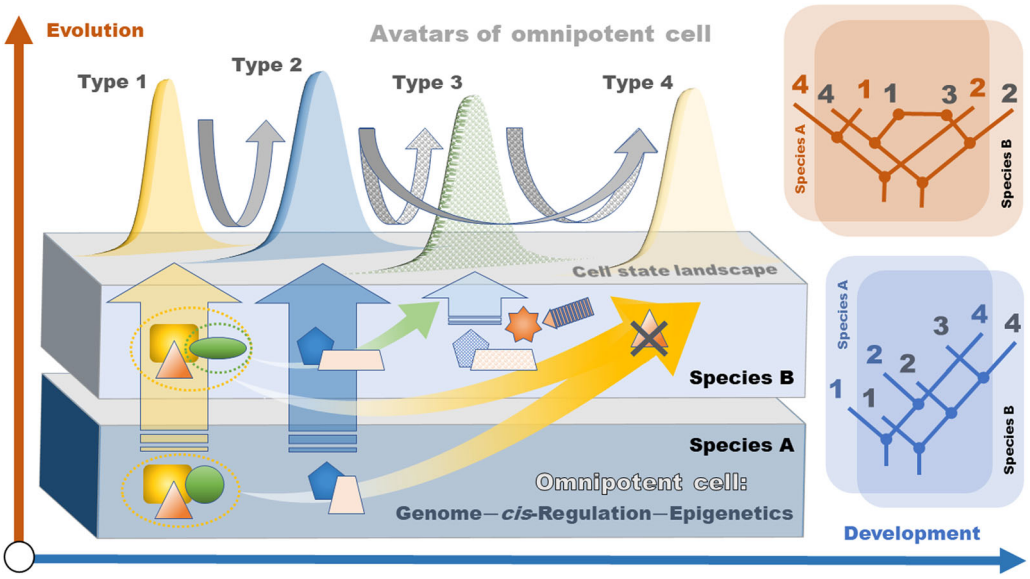

FIGURE 1 Cell-type formation in evolution and development. Biological cell types are recognized as avatars, transient or terminal cell states deployed in a continuum of states by the developmental programme of the omnipotent cell on a landscape of permissible cell states. Peaks correspond to stable cell states (distinct cell types deployed by time-steady regulation modules; numbered and coloured) and the valleys—to intermediate ones, which the cell needs to cross to reach other stable states (transient cell types deployed by regulation modules in a process of their rewiring during cell differentiation, reprogramming or direct conversion). Type–type remodelling events correspond to trajectories (grey curved arrows) that the developmental programme is able to travel between the peaks to bring the cell from one state to another during differentiation in ontogenesis (blue horizontal axis). Cell types can share parts of the same regulation modules for deployment (Type 4 is deployed by omitting one regulator [red triangle] from the core module deploying Type 1; all coloured arrows denote the modules' involvement in cell-type deployment in ontogenesis). In evolution (red vertical axis), cell types can emerge via module duplication events (in Species B, novel Type 3 is realized by a duplicated [paralogous; pattern-filled] core module of Type 2). Chimeric origin of cell types implicates co-option of a regulator from other modules within the same organism (in Species B, Type 3 recruits the green regulator from Type 1 into its active module; unlike the duplicated core module, this regulator remains single-copy in the genome, thus serving as a potential driver of concerted evolution in Types 1 and 3; Type 3 is absent in Species A and pattern-filled to denote its chimeric nature). Epigenetic regulation is implicated in deploying Type 3 (dark red star and distinct arrow; this regulator is encoded distantly in the genome). Types 1 and 4, 2 and 3, respectively, are homologous by the related nature of their underlying modules, whereas Type 3 is related to both Types 1 and 2 (producing a network region in the cell-type phylogeny for Species B; red tree on the right panel). Notably, in a cell lineage, Types 1–4 are deployed in a sequential order (as visited on the cell-state landscape), which produces the basic incongruence between the developmental and evolutionary trees of cell types (blue and red trees, right panel).

to bring the cell from one state to another during the process of differentiation. The stable states remain transitive with respect to each other, so that the cell can visit them in any direction at a certain depth of a differentiation trajectory until no further regulation module rewiring is feasible (e.g., due to epigenetic silencing of enhancers), which produces a terminally committed cell fate and requires a special resetting for the programme to proceed.

An avatar concept allows for high evolutionary flexibility of cell types and does not predefine their number in ontogenesis, including embryonic development and adult regeneration. As the developmental programme ultimately translates into the genomic order and sequence of genes and their *cis*-regulatory elements, it is the evolvability of the enhancers that primarily affects the regulation segment and module composition via the combinatorics of TF-binding motifs and their specificity for epigenetic control. According to recent evidence, enhancers preserve their motif repertoire and a special, TF-interpretable regulatory syntax for hundreds of millions of years. Once established, however, enhancers are maintained as the basis of conserved regulation modules and can evolve through the expansion, loss and integration of new TF-binding motifs (E. S. Wong et al., 2020), thus fuelling the evolution of cell types. Notably, the integration of new motifs provides a mechanism for co-opting segments between modules, that is, for generating an experimental diversity of novel, chimeric types by module fusion. If favoured by natural selection, the invented modules eventually stabilize and become imprinted in the genome as characters inheritable via the omnipotent cell of the germline. The module origin and evolution are well described (albeit not explicitly implicating the epigenetic component) by a CoRC-based model, which considers the CoRC module an equivalent of cell type identity. However, biological cells may be driven by more than one module and the modules themselves may change dynamically in situ (see below). Therefore, this model explains the stem principle of cell-type origin and evolution via regulatory isolation but may not suffice to describe actual cell types, the living cells, as those are more versatile entities that may combine or even invent identities.

The number of cell states permissible for deployment in a cell lineage is only limited by the robustness of the currently operating regulatory network to adding or removing regulators, besides the survival of resulting avataric phenotypes. For instance, an encounter with a new, ectopic regulator (such as a protein factor, cofactor or small molecule agonist) may dynamically rewire the network and alter gene transcription pattern towards a new state and phenotype, which are not encoded in the genome as part of the developmental programme. This situation is feasible, as many enhancers remain hypomethylated and responsive even after decommissioning, thus forming a gene memory that can be activated once the affine TFs appear, for example, due to inductive signalling (Hon et al., 2013; Jadhav et al., 2019). Essentially, this allows for high flexibility of a cell lineage to deploy ad hoc, 'nonphysiological' cell states that never existed in evolution or normal development. Such cases are particularly envisaged during regeneration, when normally separated tissues come in close contact upon injury and start to receive ectopic molecular signalling (see Rajagopal & Stanger, 2016). The cell capacity to engender ad hoc avatars obviously bears critical adaptive advantages to a multicelled colony in a range of natural settings of exposure to extrinsic and unexpected signals.

An avatar concept essentially blurs the boundaries between cell types per se by recognizing them as time-steady cell states deployed in a continuum of states according to an instantaneous equilibrium of the cell regulatory network capable of rewiring in response to stimuli. This thinking eliminates the internal contradiction that emerges from the empirical single-cell evidence of graded transcriptomic, as well as functional, heterogeneity of cell types exhibiting transient, activity-dependent expression variation in a range of genes, including cell-type-defining TFs. These varying manifestations of the 'same' cell types hamper rigorous CoRC-based interpretation of the real cell-type data and stimulated attempts to discriminate between cell types versus cell states on the basis of formal or functional criteria (Cembrowski & Menon, 2018; Poulin et al., 2016; Tasic, 2018), as well as inspired views to consider transient types as a source of evolutionary novelty (Erkenbrack et al., 2018). Understanding the limits of transient within-cell-type variation has been also identified as a key challenge for future comparative single-cell studies (Arendt et al., 2019). Meanwhile, an avatar concept naturally explains these phenomena by recognizing that cell types in their strict definition simply never occurred and that biological cells have a 'relativistic' identity depending on the instantaneous co-activity of distinct regulation modules, each potentially having a distinct evolutionary history.

7 | IMPLICATIONS OF CELL TYPE THEORY IN EVO-DEVO STUDIES

The apprehension of principles underlying the origin and evolution of cell types, probably, stands among the major generalizations of biology since the embracing of natural selection. The discovery of key molecular signatures in the active gene kit of single cells provided for the first objective and testable criterion to capture backbone homologies, which permeate the living cells irrespective of their structure, form, function or developmental origin. Within the emerging theoretical framework of inherently regulation-defined cell identity, what are the implications for comparative anatomy and embryology in interpreting the evolution of cells, tissues, organs and, ultimately, the body plans? What are the prospects to reconstruct the appearance and function of ancestral cells and multicelled structures?

The notion that molecular regulatory mechanisms govern cell phenotype and differentiation is grounded in robust experimental evidence and is currently indisputable (e.g., see Ladewig et al. [2013] for a review on experimental, factor-induced cell conversion). Many cell functions, such as neuronal or contractile activity, are thoroughly characterized for their underlying regulatory pathways and specific modulators. Provided with the currently well elaborated theory and toolkit of molecular evolution, as well as an extensive genomic sampling of model and nonmodel organisms, it is possible to reliably establish common descent for most of the elements in these pathways and, thereby, elicit innate homology of their regulated cell processes, many of which are obviously conserved in evolution. It is this knowledge from where came major insights into the incredible transformation of ancestral functions that, being united under common regulation, supplied the cell with new instruments.

Among the most studied examples of a novelty by assembly are neuronal cells. Current evidence attests that major components of the synapse—a defining feature of neurons—evolved through the integration of cell exocytosis (secretion function), sensory and cytoskeleton control, the receptor machinery of septate- and adherens-like cell junctions (adhesion function), as well as voltage-gated ion channels (flagellar motility-related function) for electric neurons, while the first neurotransmitters initially were likely involved in mediating behaviour via cell response to injury (Arendt, 2020; Göhde et al., 2021; Moroz et al., 2021). Most, if not all, of these components had emerged before the origin of neurons and many key proteins—before the origin of animals, which makes it rather difficult to name a specifically 'neuronal' molecular gain for synapses, because the ancestral functions of these proteins were likely non-neural. Although sharing major molecular components that confer specific subfunctions, neurons and their synapses have been 'assembled' in a variety of ways across the 'synaptic' Metazoa, to the extent of their virtually independent, convergent acquisition in ctenophores (refer to Burkhardt & Jékely [2021] for a recent finding of an exceptional cell type, which may provide a link with the ancestral, eumetazoan neuronal diversity). Meanwhile, primordial pre- and postsynaptic modules have already been coregulated in 'nonsynaptic' sponges (see above; Musser et al., 2021) and may be considered a bright example of apomeric module evolution.

Likewise and as mentioned above, cnidarian lineage-specific nematocytes—peculiar 'explosive' cells with a dischargeable stinging organelle—began to coregulate an ancient sensory function, a large repertoire of organelle-specific proteins and a co-opted, voltage-gated calcium channel engaged in filtering extrinsic cues to control stinging response (Sebé-Pedrós, Saudemont, et al., 2018; Weir et al.,

2020), the channels being an ancient invention of unicelled eukaryotes who still use them for flagellar-mediated motility and gliding (e.g., Burkhardt & Jékely, 2021). Due to the presence of the commonly 'neuronal' genetic components, nematocytes were algorithmically clustered within the neural family of cell types in Tarashansky et al. (2021).

On a deeper, subcellular level where structures have always been regarded as maximally conserved, the cilium reveals an unexpected heterogeneity of its central apparatus' protein composition: two recent proteomic studies of flagella in the *Chlamydomonas* green alga identified ~40 novel candidate proteins, of which only 15 have human orthologues, whereas the central apparatus' functionality and ultrastructural morphology remain universally stable (see Samsel et al. [2021] for a review). Certainly, this evidence does not question the homology of eukaryotic cilium, although it does elicit that even critically important end-function genes are prone to significant evolutionary change, which may not affect morphology or primary function. As for cell-type evolution, comparative single-cell transcriptomics suggests that paralogous substitution of functional genes is a common feature of diversifying regulation modules in animals (Tarashansky et al., 2021).

Overall, it is becoming clear that the only cellular entities relatively robust against evolutionary change are regulation modules and segments of the cell developmental programme that execute small, tailored cell functions. This means that it is the homology of these 'elementary' function-linked units (below referred to as elementals), not of the cell types, that is eventually established in single-cell transcriptomic comparisons. In this sense, it is the evolution of the developmental programme that shapes cell diversity and stands behind the intuitive evolution of conventional cell types. Each integration/loss of an elemental into/from a higher-rank regulation module marks, by theoretical definition, the origin of a cell type by genetic individuation. However, the evolutionary history of such events will not be captured in a comparison of contemporary cell-type transcriptomes, unless there have survived all historical sister cell types preserving intermediate ancestral states, which means that it will return a flat current snapshot of actual cell-type phylogeny. A lucky exception may appear as neuronal cells, which may exhibit a clearly nested hierarchy of cell-type transcriptomes, thus suggesting a succession of divergence events in an ancestral regulation module (Arendt et al., 2019; Tasic, 2018; Zeng & Sanes, 2017). However, this scenario cannot be formally distinguished from the flattening of multiple independent regulation modification events having occurred in the individual histories of cell types, which produced similar modern states but due to convergency.

When the elementals become recruited between regulation modules, thus entailing cell-type novelty by fusion, a scenario becomes seriously more complex to form a phylogenetic network of cell types (Figure 1). That is, the emergence of neurons themselves would certainly not be described by hierarchic transcription. In fact, this situation is the most common in cell-type comparisons, with epi- and endothelia being an extreme example, where the types exhibit derived and interspersed patterns in cell-type alignments and clustering (Cazet et al., 2022; Tarashansky et al., 2021) or fall widely across the clades of a proxy cell type phylogeny, both within- and between-species (Wang et al., 2021). The deeper in time, the more entangled this exchange becomes, eventually forming a dense network of relationships, which means that the earliest history of cell types will remain obscure.

The important notion ensuing is that an evolutionary comparison, that is, any homology statement, is only valid between cell types at an equivalent stage of differentiation when potentially homologous regulation modules have already been deployed by the developmental programme. This means that stem cells or other intermediates are not evolutionarily comparable with terminally committed cells, whereas the intermediates themselves are to be compared at their equivalent stages. As heterochrony of gene transcription is, obviously, common during cell differentiation, as well as because the living cells may dynamically change or combine identities, any evolutionary study of cell types becomes a delicate task requiring special precautions, and, probably, this comparability condition will not be fully satisfied in most cases due to the natural absence of a fully comparable synchronic deployment window even between terminally committed cells. This notion is likely to explain the affinities inherently uniting various stem-like and progenitor types across cellular and phylogenetic lineages, as well as many of the puzzling type–type similarities captured in embryonic and cross-stage ontogenetic comparisons (Cazet et al., 2022; Meyer et al., 2022; Tarashansky et al., 2021; Wang et al., 2021).

Based on the foregoing, strict homology can only be inferred in practice for the genetic and lower-rank regulation components of the omnipotent cell's developmental programme. In principle, this also applies to cases when certain regulators in a module become substituted for same-function analogues (e.g., TF substitution by same TF family members), provided that the module's architecture remains unchanged (e.g., see Schlosser [2021] for a discussion on module stability). Thereby, homologous will appear many subcellular structures executing small functions and driven by elemental regulations. By contrast, homology statements are largely inapplicable to cell types, because the history of genetic individuations (a cell-type phylogeny) cannot be deduced from extant cell states without having an explicit model of cell type phylogenetic evolution (e.g., see Tanay & Sebé-Pedrós [2021] for the prospects in model development), which currently prohibits homology separation from convergence. However, once developed, cell-type phylogenetics is likely to recover rich evolutionary networks whose complexity increases dramatically towards the root of the animal tree, with one-to-one homologies appearing rather exceptional and mostly restricted to shallower and intraspecific levels. Likewise, the relatedness of cell types cannot be inferred from relative age content of shared signature genes, because a living cell is deployed by a heterogeneous collection of elemental regulations with potentially distinct evolutionary histories and it is only the degree of antiquity of the cell-type functions that becomes evident from phylostratigraphic data.

Although the above circumstances are to be considered for correct interpretations of cell homology, the similarities revealed from 'flat' transcriptomic comparisons certainly can be used to define

cell-type identities, as well as to capture evolutionary relatedness at shallower phylogenetic levels of the origin of extant cell types.

Phenotypic convergence, in turn, appears to be universally common at the cell level. A major source of convergency in evolutionarily related cell types is concerted evolution (Figure 1). It occurs when changes impact the CoRC elements shared by multiple cell types within an organism, which alters transcription of the same end-function genes across these cell types (Musser & Wagner, 2015). A good illustration of concerted evolution may serve neurons, where pan-neuronal genes are expressed jointly by most neuronal cell types, while being regulated by a handful of homeobox TFs (Leyva-Díaz & Hobert, 2022). Earlier studies have demonstrated that partial deletion of specific enhancers affects pan-neuronal gene expression across neuronal types in the nematode *C. elegans* (Stefanakis et al., 2015). Another well-documented case revealing the potential for concerted evolution are contractile genes shared across different muscle cell types within Bilateria (Brunet et al., 2016; Steinmetz et al., 2012).

Other forces shape convergent phenotypes in transcriptionally distinct cells, such as cryptic morphotypes within morphologically homogeneous cell populations, and those remain largely unexplored. As a possibility, a handful of signature end-function genes shared between the distant types may suffice to produce similar cell morphologies. However, a more realistic, and, at the same time, the least elaborated explanation would be that the wealth of animal cell morphology is driven by the complexity of genetic and epigenetic regulatory mechanisms, including posttranslational protein modification. Indeed, a phylogenomic analysis of the evolutionary dynamics of all gene families sequenced across the animal tree of life systematically demonstrated that evolutionary innovation in animals cannot be explained by gradual enrichment of the genomic repertoire (Fernández & Gabaldón, 2020). Instead, massive gain of novel animal genes had only occurred in the root of the Metazoa and was followed by rampant gene loss accompanied by local lineage-specific duplications or gains. Among the top 'genome-depleted' are all deuterostomes (and so all vertebrates), which, in spite, exhibit the greatest amount of novelty at the cell and tissue levels. Therefore, the structural and functional complexity of metazoans must have been shaped concurrently by other evolutionary processes, particularly, the expansion of distal *cis*-regulators (e.g., Sebé-Pedrós et al., 2016; Sebé-Pedrós, Chomsky, et al., 2018; Sebé-Pedrós, Saudemont, et al., 2018) and gain/change of cell function via the mechanisms of creating multidomain proteins, including exon shuffling (e.g., Patthy, 2021). It is the matter of future discovery to understand how cell phenotype stability is supported by the underlying gene expression, if and how regulatory factor specificity or the end function of regulated genes can be affected by engaging new specific distal enhancers (Gaiti et al., 2017; Schwaiger et al., 2014; Sebé-Pedrós et al., 2016), promoter types (Lenhard et al., 2012), regulatory RNAs and their binding sites (Gaiti et al., 2018) or other emerging sources of genetic regulation and enrichment, such as transposable elements (Mukherjee & Moroz, 2022).

Embracing the spectacular plasticity of the animal cell, shall we abandon morphotype as a source of comparative evolutionary information? Certainly not, since many traditionally related cell morphotypes also exhibit transcriptional similarity, thus appearing the legitimate relatives that share common descent. Sadly though, we may probably never be able to reconstruct the outward appearance of an urmetazoan ancestor, at least, until we learn how to predict certain morphological 'templates' from particular genome–regulation contexts. The exceptions may include certain ultrastructural features which are usually driven by small, elemental cellular functions and have conserved extant phenotypes (e.g., secretion vesicular trafficking or flagellar motility). Meanwhile, the incredible capacity of the cell to assemble and transform small functions beyond recognition (e.g., as with the synapse) should always be kept in mind.

Paradoxically, although being incredibly labile itself, the cell produces more stable structures at a supracellular level—tissues, organs and body plans, where common descent and homology statements become more relevant. This stabilization is achieved via the cell capacity to control stability of its differentiated states via response to extrinsic stimuli. Conceptually, inductive cell–cell signalling that appears in an incipient colony rewires certain cell-type deploying regulation modules of the developmental programme, which, as more cells appear, become progressively dependent on the previously deployed ones toward the establishment of a ratchet-like execution network, where some modules get irreversibly 'entrenched' and enter specific, time-steady states in a spatially controlled manner (see Libby & Ratcliff [2014] for a discussion on the ratcheting principle with respect to cell function and transition to multicellularity). It is this entrenchment of a complex and dynamic regulatory network that stabilizes the formation of supracellular-level structures and imparts evolutionary stability to organismal morphological units. The reader is referred at this point to theoretical studies on hierarchical regulatory character identity networks, which provide a mechanistic framework for the developmental control of character identities towards the proposal of a dynamic (vs. a 'phylum'-fixed) concept of body plan evolution via stepwise accretion of complexity as new interactions entail new dependencies and constraints (DiFrisco & Wagner, 2022; G. P. Wagner, 2007, 2015).

In 1949, Alexey Zakhvatkin published his Synzoospore hypothesis of the origin of Metazoa via transition from temporal to spatial cell differentiation (Sachwatkin, 1956; Zakhvatkin, 1949). Grounding solely in the evidence from protistology and embryology, this theory appeared decades ahead of its time, but even today, some in the scientific community hardly adopt this thinking due to a literal interpretation of this transition as an assembly of genetically heterogeneous, differentiated life cycle stages in a protistan ancestor into an integrated, multicelled colony representing the first animal. Instead, this phase transition was fuelled by the ability of sole animal cell to differentiate into all states required to form the colony. Heterochronic module deployment is a salient and inherent feature of the cell developmental programme, whereby a variation in the tempo of state–state trajectories does not break its correct sequential execution. Thus, trunk patterning by HOX genes has been recently shown to have shifted from premetamorphic stages to early embryogenesis in the evolution of annelids, with such a

heterochronic event being associated with noticeable changes in transcriptomic and chromatin remodelling in primitive vs. derived species (Martín-Zamora et al., 2023). This case study well illustrates the capacity of the developmental programme to 'condense,' in much the same manner as it had once occurred when all the life-cycle cell states once imprinted in the genome of a unicelled ancestor became executed by the first clonally multicellular holozoan.

8 | CONCLUSIONS

To briefly conclude, our understanding of the laws and mechanisms that shape the cell in evolution and drive it along the path of differentiation is largely in its infancy. Theoretical and experimental progress is imperative in the near future for closing several major gaps, primarily, in obtaining a formal definition of the cell type as a biological and computational unit; developing a sound phylogenetic framework for modelling cell type evolution, which is critically required to infer cell type trees and networks from transcriptomic data; and elucidating the genetic and regulatory bases of cell phenotype stability both in evolution and cell lineage. In just a decade, there has been a transformative change in perceptions that have been shaped for over two centuries, showing that the cell is the most flexible, omnipotent unit of living matter. As in the quantum world, where the laws of classical mechanics do not apply, many approaches of classical biology need to be reconsidered to understand evolution and diversity at the cell level. The emerging generalistic theory of cell types is already beginning to yield insights, suggesting the existence of rather simple rules behind the complexity of cell transformations, and that we should seek further discovery on the way that Edmund B. Wilson envisaged: 'The key to every biological problem must finally be sought in the cell, for every living organism is, or at some time has been, a cell' (Wilson, 1925).

ACKNOWLEDGMENTS

The author is grateful for two anonymous reviewers for their thoughtful comments and suggestions for additional levels of discussion. The work was supported by FASIE (4436GS1/72572).

CONFLICT OF INTEREST STATEMENT

The author declares no conflict of interest.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article, as no new data were created or analysed in this study.

ORCID

Leonid Y. Rusin http://orcid.org/0000-0003-2424-298X

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1002/jmor.21569.

REFERENCES

Achim, K., & Arendt, D. (2014). Structural evolution of cell types by stepwise assembly of cellular modules. *Current Opinion in Genetics & Development*, *27*, 102–108.

Achim, K., Eling, N., Vergara, H. M., Bertucci, P. Y., Musser, J., Vopalensky, P., Brunet, T., Collier, P., Benes, V., Marioni, J. C., & Arendt, D. (2018). Whole-body single-cell sequencing reveals transcriptional domains in the annelid larval body. *Molecular Biology and Evolution*, *35*, 1047–1062.

Alberts, B. (1998). The cell as a collection of protein machines: Preparing the next generation of molecular biologists. *Cell*, *92*, 291–294.

Alié, A., Hayashi, T., Sugimura, I., Manuel, M., Sugano, W., Mano, A., Satoh, N., Agata, K., & Funayama, N. (2015). The ancestral gene repertoire of animal stem cells. *Proceedings of the National Academy of Sciences*, *112*. https://doi.org/10.1073/pnas.1514789112

Alié, A., & Manuel, M. (2010). The backbone of the post-synaptic density originated in a unicellular ancestor of choanoflagellates and metazoans. *BMC Evolutionary Biology*, *10*, 34.

Allen, A. M., Neville, M. C., Birtles, S., Croset, V., Treiber, C. D., Waddell, S., & Goodwin, S. F. (2020). A single-cell transcriptomic atlas of the adult *Drosophila* ventral nerve cord. *eLife*, *9*, e54074. https://doi.org/10.7554/eLife.54074

Amadei, G., Handford, C. E., Qiu, C., de Jonghe, J., Greenfeld, H., Tran, M., Martin, B. K., Chen, D. Y., Aguilera-Castrejon, A., Hanna, J. H., Elowitz, M. B., Hollfelder, F., Shendure, J., Glover, D. M., & Zernicka-Goetz, M. (2022). Embryo model completes gastrulation to neurulation and organogenesis. *Nature*, *610*, 143–153.

Arendt, D. (2005). Genes and homology in nervous system evolution: Comparing gene functions, expression patterns, and cell type molecular fingerprints. *Theory in Biosciences*, *124*, 185–197.

Arendt, D. (2008). The evolution of cell types in animals: Emerging principles from molecular studies. *Nature Reviews Genetics*, *9*, 868–882.

Arendt, D. (2020). The evolutionary assembly of neuronal machinery. *Current Biology*, *30*, R603–R616.

Arendt, D. (2021). Animal evolution: Of flame and collar cells. *Current Biology*, *31*, R1003–R1006.

Arendt, D., Benito-Gutierrez, E., Brunet, T., & Marlow, H. (2015). Gastric pouches and the mucociliary sole: Setting the stage for nervous system evolution. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *370*, 20150286.

Arendt, D., Bertucci, P. Y., Achim, K., & Musser, J. M. (2019). Evolution of neuronal types and families. *Current Opinion in Neurobiology*, *56*, 144–152.

Arendt, D., Musser, J. M., Baker, C. V. H., Bergman, A., Cepko, C., Erwin, D. H., Pavlicev, M., Schlosser, G., Widder, S., Laubichler, M. D., & Wagner, G. P. (2016). The origin and evolution of cell types. *Nature Reviews Genetics*, *17*, 744–757.

Arendt, D., Tessmar-Raible, K., Snyman, H., Dorresteijn, A. W., & Wittbrodt, J. (2004). Ciliary photoreceptors with a vertebrate-type opsin in an invertebrate brain. *Science*, *306*, 869–871.

Arthur, W. (1997). *The origin of animal body plans: A study in evolutionary developmental biology*. Cambridge University Press.

Arthur-Farraj, P. J., Latouche, M., Wilton, D. K., Quintes, S., Chabrol, E., Banerjee, A., Woodhoo, A., Jenkins, B., Rahman, M., Turmaine, M., Wicher, G. K., Mitter, R., Greensmith, L., Behrens, A., Raivich, G., Mirsky, R., & Jessen, K. R. (2012). c-Jun reprograms Schwann cells of injured nerves to generate a repair cell essential for regeneration. *Neuron*, *75*, 633–647.

Ax, P. (1984). *Das phylogenetische System: Systematisierung der lebenden Natur aufgrund ihrer Phylogenese*. Gustav Fischer.

Bayraktar, O. A., & Doe, C. Q. (2013). Combinatorial temporal patterning in progenitors expands neural diversity. *Nature*, *498*, 449–455.

de Beer, G. R. (1954). Embryos and ancestors *(Revised ed)*. Oxford University Press.

Bird, A. (2002). DNA methylation patterns and epigenetic memory. *Genes & Development*, *16*, 6–21.

Bogdanović, O., Smits, A. H., de la Calle Mustienes, E., Tena, J. J., Ford, E., Williams, R., Senanayake, U., Schultz, M. D., Hontelez, S., van Kruijsbergen, I., Rayon, T., Gnerlich, F., Carell, T., Veenstra, G. J. C., Manzanares, M., Sauka-Spengler, T., Ecker, J. R., Vermeulen, M., Gómez-Skarmeta, J. L., & Lister, R. (2016). Active DNA demethylation at enhancers during the vertebrate phylotypic period. *Nature Genetics*, *48*, 417–426.

Brawley, C., & Matunis, E. (2004). Regeneration of male germline stem cells by spermatogonial dedifferentiation *in vivo*. *Science*, *304*, 1331–1334.

Briggs, J. A., Weinreb, C., Wagner, D. E., Megason, S., Peshkin, L., Kirschner, M. W., & Klein, A. M. (2018). The dynamics of gene expression in vertebrate embryogenesis at single-cell resolution. *Science*, *360*. https://doi.org/10.1126/science.aar5780

Brückner, A., Badroos, J. M., Learsch, R. W., Yousefelahiyeh, M., Kitchen, S. A., & Parker, J. (2021). Evolutionary assembly of cooperating cell types in an animal chemical defense system. *Cell*, *184*, 6138–6156.

Brunet, T., Fischer, A. H., Steinmetz, P. R., Lauri, A., Bertucci, P., & Arendt, D. (2016). The evolutionary origin of bilaterian smooth and striated myocytes. *eLife*, *5*, 1–24.

Brunet, T., & King, N. (2017). The origin of animal multicellularity and cell differentiation. *Developmental Cell*, *43*, 124–140.

Brunet, T., & King, N. (2022). The single-celled ancestors of animals. A history of hypotheses. In M. D. Herron, P. L. Conlin, & W. C. Ratcliff (Eds.), *The evolution of multicellularity* (1st Ed., pp. 251–278). CRC Press.

Budd, G. E., & Jensen, S. (2000). A critical reappraisal of the fossil record of the bilaterian phyla. *Biological Reviews of the Cambridge Philosophical Society*, *75*, 253–295.

Bürglin, T. R., & Affolter, M. (2016). Homeodomain proteins: An update. *Chromosoma*, *125*, 497–521.

Burkhardt, P., & Jékely, G. (2021). Evolution of synapses and neuro-transmitter systems: The divide-and-conquer model for early neural cell-type evolution. *Current Opinion in Neurobiology*, *71*, 127–138.

Burki, F., Roger, A. J., Brown, M. W., & Simpson, A. (2020). The new tree of eukaryotes. *Trends in Ecology & Evolution*, *35*, 43–55.

Buzgariu, W., al Haddad, S., Tomczyk, S., Wenger, Y., & Galliot, B. (2015). Multi-functionality and plasticity characterize epithelial cells in *Hydra*. *Tissue Barriers*, *3*, e1068908.

Cao, C., Lemaire, L. A., Wang, W., Yoon, P. H., Choi, Y. A., Parsons, L. R., Matese, J. C., Wang, W., Levine, M., & Chen, K. (2019). Comprehensive single-cell transcriptome lineages of a proto-vertebrate. *Nature*, *571*, 349–354.

Cao, J., O'Day, D. R., Pliner, H. A., Kingsley, P. D., Deng, M., Daza, R. M., Zager, M. A., Aldinger, K. A., Blecher-Gonen, R., Zhang, F., Spielmann, M., Palis, J., Doherty, D., Steemers, F. J., Glass, I. A., Trapnell, C., & Shendure, J. (2020). A human cell atlas of fetal gene expression. *Science*, *370*, eaba7721. https://doi.org/10.1126/science.aba7721

Cao, J., Packer, J. S., Ramani, V., Cusanovich, D. A., Huynh, C., Daza, R., Qiu, X., Lee, C., Furlan, S. N., Steemers, F. J., Adey, A., Waterston, R. H., Trapnell, C., & Shendure, J. (2017). Comprehensive single-cell transcriptional profiling of a multicellular organism. *Science*, *357*, 661–667.

Cao, J., Spielmann, M., Qiu, X., Huang, X., Ibrahim, D. M., Hill, A. J., Zhang, F., Mundlos, S., Christiansen, L., Steemers, F. J., Trapnell, C., & Shendure, J. (2019). The single-cell transcriptional landscape of mammalian organogenesis. *Nature*, *566*, 496–502.

Cazet, J. F., Siebert, S., Little, H. M., Bertemes, P., Primack, A. S., Ladurner, P., & Juliano, C. E. (2022). New *Hydra* genomes reveal conserved principles of hydrozoan transcriptionan regulation. *BioRxiv*. https://doi.org/10.1101/2022.06.21.496857

Cedar, H., & Bergman, Y. (2012). Programming of DNA methylation patterns. *Annual Review of Biochemistry*, *81*, 97–117.

Cembrowski, M. S., & Menon, V. (2018). Continuous variation within cell types of the nervous system. *Trends in Neurosciences*, *41*, 337–348.

Chari, T., Weissbourd, B., Gehring, J., Ferraioli, A., Leclère, L., Herl, M., Gao, F., Chevalier, S., Copley, R. R., Houliston, E., Anderson, D. J., & Pachter, L. (2021). Whole-animal multiplexed single-cell RNA-seq reveals transcriptional shifts across *Clytia medusa* cell types. *Science Advances*, *7*, eabh1683. https://doi.org/10.1126/sciadv.abh1683

Chen, C., Gao, Y., Liu, W., & Gao, S. (2022). Epigenetic regulation of cell fate transition: Learning from early embryo development and somatic cell reprogramming. *Biology of Reproduction*, *107*, 183–195.

Cote, L. E., Simental, E., & Reddien, P. W. (2019). Muscle functions as a connective tissue and source of extracellular matrix in planarians. *Nature Communications*, *10*, 1592.

Crispatzu, G., Rehimi, R., Pachano, T., Bleckwehl, T., Cruz-Molina, S., Xiao, C., Mahabir, E., Bazzi, H., & Rada-Iglesias, A. (2021). The chromatin, topological and regulatory properties of pluripotency-associated poised enhancers are conserved *in vivo*. *Nature Communications*, *12*, 4344.

Croset, V., Treiber, C. D., & Waddell, S. (2018). Cellular diversity in the *Drosophila* midbrain revealed by single-cell transcriptomics. *eLife*, *7*. https://doi.org/10.7554/eLife.34550

Davie, K., Janssens, J., Koldere, D., de Waegeneer, M., Pech, U., Kreft, Ł., Aibar, S., Makhzami, S., Christiaens, V., Bravo González-Blas, C., Poovathingal, S., Hulselmans, G., Spanier, K. I., Moerman, T., Vanspauwen, B., Geurs, S., Voet, T., Lammertyn, J., Thienpont, B., … Aerts, S. (2018). A single-cell transcriptome atlas of the aging *Drosophila* brain. *Cell*, *174*, 982–998.e20.

Davis, R. L., Weintraub, H., & Lassar, A. B. (1987). Expression of a single transfected cDNA converts fibroblasts to myoblasts. *Cell*, *51*, 987–1000.

Deneris, E. S., & Hobert, O. (2014). Maintenance of postmitotic neuronal cell identity. *Nature Neuroscience*, *17*, 899–907.

DiFrisco, J., & Wagner, G. P. (2022). Body plan identity: A mechanistic model. *Evolutionary Biology*, *49*, 123–141.

Dillon, N., Cocanougher, B., Sood, C., Yuan, X., Kohn, A. B., Moroz, L. L., Siegrist, S. E., Zlatic, M., & Doe, C. Q. (2022). Single cell RNA-seq analysis reveals temporally-regulated and quiescence-regulated gene expression in *Drosophila* larval neuroblasts. *Neural Development*, *17*, 7.

Dohrmann, M., & Wörheide, G. (2017). Dating early animal evolution using phylogenomic data. *Scientific Reports*, *7*, 3599.

Domazet-Lošo, T., Brajković, J., & Tautz, D. (2007). A phylostratigraphy approach to uncover the genomic history of major adaptations in metazoan lineages. *Trends in Genetics*, *23*, 533–539.

Domazet-Lošo, T., & Tautz, D. (2010). A phylogenetically based transcriptome age index mirrors ontogenetic divergence patterns. *Nature*, *468*, 815–818.

Drost, H.-G., Janitza, P., Grosse, I., & Quint, M. (2017). Cross-kingdom comparison of the developmental hourglass. *Current Opinion in Genetics & Development*, *45*, 69–75.

Du, Z., Zhang, K., & Xie, W. (2021). Epigenetic reprogramming in early animal development. *Cold Spring Harbor Perspectives in Biology*, *14*, a039677.

Dubaissi, E., Rousseau, K., Lea, R., Soto, X., Nardeosingh, S., Schweickert, A., Amaya, E., Thornton, D. J., & Papalopulu, N. (2014). A secretory cell type develops alongside multiciliated cells, ionocytes and goblet cells, and provides a protective, anti-infective function in the frog embryonic mucociliary epidermis. *Development*, *141*, 1514–1525.

Duboule, D. (1994). Temporal colinearity and the phylotypic progression: A basis for the stability of a vertebrate Bauplan and the evolution of morphologies through heterochrony. *Development*, *1994*, 135–142.

DuBuc, T. Q., Schnitzler, C. E., Chrysostomou, E., McMahon, E. T., Febrimarsa, I., Gahan, J. M., Buggie, T., Gornik, S. G., Hanley, S.,

Barreira, S. N., Gonzalez, P., Baxevanis, A. D., & Frank, U. (2020). Transcription factor AP2 controls cnidarian germ cell induction. *Science*, *367*, 757–762.

Duruz, J., Kaltenrieder, C., Ladurner, P., Bruggmann, R., Martìnez, P., & Sprecher, S. G. (2021). Acoel single-cell transcriptomics: Cell type analysis of a deep branching bilaterian. *Molecular Biology and Evolution*, *38*, 1888–1904.

Emerson, M. M., Surzenko, N., Goetz, J. J., Trimarchi, J., & Cepko, C. L. (2013). Otx2 and Onecut1 promote the fates of cone photoreceptors and horizontal cells and repress rod photoreceptors. *Developmental Cell*, *26*, 59–72.

Engström, P. G., Ho Sui, S. J., Drivenes, Ø., Becker, T. S., & Lenhard, B. (2007). Genomic regulatory blocks underlie extensive microsynteny conservation in insects. *Genome Research*, *17*, 1898–1908.

Erkenbrack, E. M., Maziarz, J. D., Griffith, O. W., Liang, C., Chavan, A. R., Nnamani, M. C., & Wagner, G. P. (2018). The mammalian decidual cell evolved from a cellular stress response. *PLoS Biology*, *16*, e2005594.

Erwin, D. H., & Davidson, E. H. (2009). The evolution of hierarchical gene regulatory networks. *Nature Reviews Genetics*, *10*, 141–148.

Fairclough, S. R., Chen, Z., Kramer, E., Zeng, Q., Young, S., Robertson, H. M., Begovic, E., Richter, D. J., Russ, C., Westbrook, M. J., Manning, G., Lang, B. F., Haas, B., Nusbaum, C., & King, N. (2013). Premetazoan genome evolution and the regulation of cell differentiation in the choanoflagellate *Salpingoeca rosetta*. *Genome Biology*, *14*, R15.

Farrell, J. A., Wang, Y., Riesenfeld, S. J., Shekhar, K., Regev, A., & Schier, A. F. (2018). Single-cell reconstruction of developmental trajectories during zebrafish embryogenesis. *Science*, *360*, eaar3131. https://doi.org/10.1126/science.aar3131

Fernández, R., & Gabaldón, T. (2020). Gene gain and loss across the metazoan tree of life. *Nature Ecology & Evolution*, *4*, 524–533.

Fincher, C. T., Wurtzel, O., de Hoog, T., Kravarik, K. M., & Reddien, P. W. (2018). Cell type transcriptome atlas for the planarian *Schmidtea mediterranea*. *Science*, *360*, eaaq1736. https://doi.org/10.1126/science.aaq1736

Fortunato, S. A., Leininger, S., & Adamska, M. (2014). Evolution of the Pax-Six-Eya-Dach network: The calcisponge case study. *EvoDevo*, *5*, 23.

Foster, S., Oulhen, N., & Wessel, G. (2020). A single cell RNA sequencing resource for early sea urchin development. *Development*, *147*, dev191528. https://doi.org/10.1242/dev.191528

Gaiti, F., Calcino, A. D., Tanurdžić, M., & Degnan, B. M. (2017). Origin and evolution of the metazoan non-coding regulatory genome. *Developmental Biology*, *427*, 193–202.

Gaiti, F., Degnan, B. M., & Tanurdžić, M. (2018). Long non-coding regulatory RNAs in sponges and insights into the origin of animal multicellularity. *RNA Biology*, *15*, 1–7.

Göhde, R., Naumann, B., Laundon, D., Imig, C., McDonald, K., Cooper, B. H., Varoqueaux, F., Fasshauer, D., & Burkhardt, P. (2021). Choanoflagellates and the ancestry of neurosecretory vesicles. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *376*, 20190759.

Graf, T., & Enver, T. (2009). Forcing cells to change lineages. *Nature*, *462*, 587–594.

Grau-Bové, X., Torruella, G., Donachie, S., Suga, H., Leonard, G., Richards, T. A., & Ruiz-Trillo, I. (2017). Dynamics of genomic innovation in the unicellular ancestry of animals. *eLife*, *6*, e26036. https://doi.org/10.7554/eLife.26036

Gąsiorowski, L., Andrikou, C., Janssen, R., Bump, P., Budd, G. E., Lowe, C. J., & Hejnol, A. (2021). Molecular evidence for a single origin of ultrafiltration-based excretory organs. *Current Biology*, *31*, 3629–3638.

Han, X., Zhou, Z., Fei, L., Sun, H., Wang, R., Chen, Y., Chen, H., Wang, J., Tang, H., Ge, W., Zhou, Y., Ye, F., Jiang, M., Wu, J., Xiao, Y., Jia, X., Zhang, T., Ma, X., Zhang, Q., … Guo, G. (2020). Construction of a human cell landscape at single-cell level. *Nature*, *581*, 303–309.

Harrison, F. W., & Ruppert, E. E. (Eds.). (1991). *Microscopic anatomy of invertebrates: Echinodermata* (Vol. 14). Wiley-Liss.

Hartenstein, V., & Martinez, P. (2019). Phagocytosis in cellular defense and nutrition: a food-centered approach to the evolution of macrophages. *Cell and Tissue Research*, *377*, 527–547.

Hartwell, L. H., Hopfield, J. J., Leibler, S., & Murray, A. W. (1999). From molecular to modular cell biology. *Nature*, *402*, C47–C52.

Hennig, W. (1965). Phylogenetic systematics. *Annual Review of Entomology*, *10*, 97–116.

Hirano, M., Guo, P., McCurley, N., Schorpp, M., Das, S., Boehm, T., & Cooper, M. D. (2013). Evolutionary implications of a third lymphocyte lineage in lampreys. *Nature*, *501*, 435–438.

Hobert, O. (2008). Regulatory logic of neuronal diversity: Terminal selector genes and selector motifs. *Proceedings of the National Academy of Sciences*, *105*, 20067–20071.

Hodge, R. D., Bakken, T. E., Miller, J. A., Smith, K. A., Barkan, E. R., Graybuck, L. T., Close, J. L., Long, B., Johansen, N., Penn, O., Yao, Z., Eggermont, J., Höllt, T., Levi, B. P., Shehata, S. I., Aevermann, B., Beller, A., Bertagnolli, D., Brouner, K., … Lein, E. S. (2019). Conserved cell types with divergent features in human versus mouse cortex. *Nature*, *573*, 61–68.

Hon, G. C., Rajagopal, N., Shen, Y., McCleary, D. F., Yue, F., Dang, M. D., & Ren, B. (2013). Epigenetic memory at embryonic enhancers identified in DNA methylation maps from adult mouse tissues. *Nature Genetics*, *45*, 1198–1206.

Horie, R., Hazbun, A., Chen, K., Cao, C., Levine, M., & Horie, T. (2018). Shared evolutionary origin of vertebrate neural crest and cranial placodes. *Nature*, *560*, 228–232.

Hu, M., Zheng, X., Fan, C.-M., & Zheng, Y. (2020). Lineage dynamics of the endosymbiotic cell type in the soft coral. *Nature*, *582*, 534–538.

Hulett, R. E., Kimura, J. O., Bolaños, D. M., Luo, Y.-J., Ricci, L., & Srivastava, M. (2022). Acoel single-cell atlas reveals expression dynamics and heterogeneity of a pluripotent stem cell population. *BioRxiv*. https://doi.org/10.1101/2022.02.10.479464

Hung, R.-J., Hu, Y., Kirchner, R., Liu, Y., Xu, C., Comjean, A., Tattikota, S. G., Li, F., Song, W., Sui, S. H., & Perrimon, N. (2020). A cell atlas of the adult *Drosophila* midgut. *Proceedings of the National Academy of Sciences*, *117*, 1514–1523.

Irimia, M., Tena, J. J., Alexis, M. S., Fernandez-Miñan, A., Maeso, I., Bogdanović, O., De la calle-Mustienes, E., Roy, S. W., Gómez-Skarmeta, J. L., & Fraser, H. B. (2012). Extensive conservation of ancient microsynteny across metazoans due to *cis*-regulatory constraints. *Genome Research*, *22*, 2356–2367.

Jadhav, U., Cavazza, A., Banerjee, K. K., Xie, H., O'Neill, N. K., Saenz-Vash, V., Herbert, Z., Madha, S., Orkin, S. H., Zhai, H., & Shivdasani, R. A. (2019). Extensive recovery of embryonic enhancer and gene memory stored in hypomethylated enhancer DNA. *Molecular Cell*, *74*, 542–554.e5.

Jaitin, D. A., Kenigsberg, E., Keren-Shaul, H., Elefant, N., Paul, F., Zaretsky, I., Mildner, A., Cohen, N., Jung, S., Tanay, A., & Amit, I. (2014). Massively parallel single-cell RNA-Seq for marker-free decomposition of tissues into cell types. *Science*, *343*, 776–779.

Jopling, C., Sleep, E., Raya, M., Martí, M., Raya, A., & Belmonte, J. C. I. (2010). Zebrafish heart regeneration occurs by cardiomyocyte dedifferentiation and proliferation. *Nature*, *464*, 606–609.

Juliano, C. E., Swartz, S. Z., & Wessel, G. M. (2010). A conserved germline multipotency program. *Development*, *137*, 4113–4126.

Karaiskos, N., Wahle, P., Alles, J., Boltengagen, A., Ayoub, S., Kipar, C., Kocks, C., Rajewsky, N., & Zinzen, R. P. (2017). The *Drosophila* embryo at single-cell transcriptome resolution. *Science*, *358*, 194–199.

Keller, T. E., Han, P., & Yi, S. V. (2016). Evolutionary transition of promoter and gene body DNA methylation across invertebrate–vertebrate boundary. *Molecular Biology and Evolution*, *33*, 1019–1028.

Kishi, Y., & Parker, J. (2021). Cell type innovation at the tips of the animal tree. *Current Opinion in Genetics & Development*, *69*, 112–121.

Klein, A. M., Mazutis, L., Akartuna, I., Tallapragada, N., Veres, A., Li, V., Peshkin, L., Weitz, D. A., & Kirschner, M. W. (2015). Droplet barcoding for single-cell transcriptomics applied to embryonic stem cells. *Cell*, *161*, 1187–1201.

Knoll, A. H. (2011). The multiple origins of complex multicellularity. *Annual Review of Earth and Planetary Sciences*, *39*, 217–239.

Kolodziejczyk, A. A., Kim, J. K., Svensson, V., Marioni, J. C., & Teichmann, S. A. (2015). The technology and biology of single-cell RNA sequencing. *Molecular Cell*, *58*, 610–620.

Kowalevsky, A. O. (1866). Entwickelungsgeschichte der einfachen Ascidien, *Mémoires l'Académie Impériale des Sciences de St.-Pétersbourg* (Vol. 10, pp. 1–19). Kaiserlichen Akademie der Wissenschaften.

Kusaba, T., Lalli, M., Kramann, R., Kobayashi, A., & Humphreys, B. D. (2014). Differentiated kidney epithelial cells repair injured proximal tubule. *Proceedings of the National Academy of Sciences*, *111*, 1527–1532.

Ladewig, J., Koch, P., & Brüstle, O. (2013). Leveling Waddington: The emergence of direct programming and the loss of cell fate hierarchies. *Nature Reviews Molecular Cell Biology*, *14*, 225–236.

Lamb, T. D. (2013). Evolution of phototransduction, vertebrate photoreceptors and retina. *Progress in Retinal and Eye Research*, *36*, 52–119.

Lau, K. Y. C., Rubinstein, H., Gantner, C. W., Amadei, G., Stelzer, Y., & Zernicka-Goetz, M. (2022). Mouse-embryo model derived exclusively from embryonic stem cells undergo neurulation and heart development. *Cell Stem Cell*, *29*, 1445–1458.

Lenhard, B., Sandelin, A., & Carninci, P. (2012). Metazoan promoters: Emerging characteristics and insights into transcriptional regulation. *Nature Reviews Genetics*, *13*, 233–245.

Leon, A., Subirana, L., Magre, K., Cases, I., Tena, J. J., Irimia, M., Gomez-Skarmeta, J. L., Escriva, H., & Bertrand, S. (2022). Gene regulatory networks of epidermal and neural fate choice in a chordate. *Molecular Biology and Evolution*, *39*, msac055.

Levin, M., Anavy, L., Cole, A. G., Winter, E., Mostov, N., Khair, S., Senderovich, N., Kovalev, E., Silver, D. H., Feder, M., Fernandez-Valverde, S. L., Nakanishi, N., Simmons, D., Simakov, O., Larsson, T., Liu, S. Y., Jerafi-Vider, A., Yaniv, K., Ryan, J. F., … Yanai, I. (2016). The mid-developmental transition and the evolution of animal body plans. *Nature*, *531*, 637–641.

Levy, S., Elek, A., Grau-Bové, X., Menéndez-Bravo, S., Iglesias, M., Tanay, A., Mass, T., & Sebé-Pedrós, A. (2021). A stony coral cell atlas illuminates the molecular and cellular basis of coral symbiosis, calcification, and immunity. *Cell*, *184*, 2973–2987.

Leyva-Díaz, E., & Hobert, O. (2022). Robust regulatory architecture of pan-neuronal gene expression. *Current Biology*, *32*, 1715–1727.

Li, P., Nanes Sarfati, D., Xue, Y., Yu, X., Tarashansky, A. J., Quake, S. R., & Wang, B. (2021). Single-cell analysis of *Schistosoma mansoni* identifies a conserved genetic program controlling germline stem cell fate. *Nature Communications*, *12*, 485.

Liang, C., Forrest, A. R. R., & Wagner, G. P. (2015). The statistical geometry of transcriptome divergence in cell-type evolution and cancer. *Nature Communications*, *6*, 6066.

Libby, E., & Ratcliff, W. C. (2014). Ratcheting the evolution of multicellularity. *Science*, *346*, 426–427.

López-Escardó, D., Grau-Bové, X., Guillaumet-Adkins, A., Gut, M., Sieracki, M. E., & Ruiz-Trillo, I. (2019). Reconstruction of protein domain evolution using single-cell amplified genomes of uncultured choanoflagellates sheds light on the origin of animals. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *374*, 20190088.

Mah, J. L., Christensen-Dalsgaard, K. K., & Leys, S. P. (2014). Choanoflagellate and choanocyte collar-flagellar systems and the assumption of homology. *Evolution & Development*, *16*, 25–37.

Mariño-Ramírez, L., Kann, M. G., Shoemaker, B. A., & Landsman, D. (2005). Histone structure and nucleosome stability. *Expert Review of Proteomics*, *2*, 719–729.

Martin-Duran, J. M., & Marletaz, F. (2020). Unravelling spiral cleavage. *Development (Cambridge)*, *147*, dev181081. https://doi.org/10.1242/dev.181081

Martín-Zamora, F. M., Liang, Y., Guynes Carrillo-Baltodano, A. M., Davies, B. E., Donnellan, R. D., Tan, Y., Moggioli, G., Seudre, O., Tran, M., Mortimer, K., Luscombe, N. M., Hejnol, A., Marlétaz, F., & Martín-Durán, J. M. (2023). Annelid functional genomics reveal the origins of bilaterian life cycles. *Nature*. https://doi.org/10.1038/s41586-022-05636-7

Massri, A. J., Greenstreet, L., Afanassiev, A., Berrio, A., Wray, G. A., Schiebinger, G., & McClay, D. R. (2021). Developmental single-cell transcriptomics in the *Lytechinus variegatus* sea urchin embryo. *Development*, *148*, dev198614. https://doi.org/10.1242/dev.198614

de Mendoza, A., & Sebé-Pedrós, A. (2019). Origin and evolution of eukaryotic transcription factors. *Current Opinion in Genetics & Development*, *58–59*, 25–32.

Metchnikoff, E. (1901). *L'immunité dans les maladies infectieuses*. Masson.

Metschnikoff, E. (1884). Memoirs: Researches on the intracellular digestion of invertebrates. *Journal of Cell Science*, *s2-24*, 89–111.

Metschnikoff, E. (1886). *Zur Verwandschaft und Ableitung der niederen Metazoen*. A. Hölder.

Meyer, A., Ku, C., Hatleberg, W., Telmer, C. A., & Hinman, V. (2022). New hypotheses of cell type diversity and novelty from comparative single cell and nuclei transcriptomics in echinoderms. *BioRxiv*. https://doi.org/10.1101/2022.05.06.490935

Mikhailov, K. V., Konstantinova, A. V., Nikitin, M. A., Troshin, P. V., Rusin, L. Y., Lyubetsky, V. A., Panchin, Y. V., Mylnikov, A. P., Moroz, L. L., Kumar, S., & Aleoshin, V. V. (2009). The origin of metazoa: A transition from temporal to spatial cell differentiation. *BioEssays*, *31*, 758–768.

Minnoye, L., Taskiran, I. I., Mauduit, D., Fazio, M., van Aerschot, L., Hulselmans, G., Christiaens, V., Makhzami, S., Seltenhammer, M., Karras, P., Primot, A., Cadieu, E., van Rooijen, E., Marine, J. C., Egidy, G., Ghanem, G. E., Zon, L., Wouters, J., & Aerts, S. (2020). Cross-species analysis of enhancer logic using deep learning. *Genome Research*, *30*, 1815–1834.

Moroz, L. L., Romanova, D. Y., & Kohn, A. B. (2021). Neural versus alternative integrative systems: Molecular insights into origins of neurotransmitters. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *376*, 20190762. https://doi.org/10.1098/RSTB.2019.0762

Mukherjee, K., & Moroz, L. L. (2022). Transposon-derived transcription factors across metazoans. *BioRxiv*. https://doi.org/10.1101/2022.12.18.520930

Musser, J. M., Schippers, K. J., Nickel, M., Mizzon, G., Kohn, A. B., Pape, C., Ronchi, P., Papadopoulos, N., Tarashansky, A. J., Hammel, J. U., Wolf, F., Liang, C., Hernández-Plaza, A., Cantalapiedra, C. P., Achim, K., Schieber, N. L., Pan, L., Ruperti, F., Francis, W. R., … Arendt, D. (2021). Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. *Science*, *374*, 717–723.

Musser, J. M., & Wagner, G. P. (2015). Character trees from transcriptome data: Origin and individuation of morphological characters and the so-called "species signal.". *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *324*, 588–604.

Nakanishi, N., Sogabe, S., & Degnan, B. M. (2014). Evolutionary origin of gastrulation: insights from sponge development. *BMC Biology*, *12*, 26.

Nedelcu, A. M. (2019). Independent evolution of complex development in animals and plants: Deep homology and lateral gene transfer. *Development Genes and Evolution*, *229*, 25–34.

Nielsen, C. (2008). Six major steps in animal evolution: Are we derived sponge larvae? *Evolution & Development*, *10*, 241–257.

Nikitin, M. (2015). Bioinformatic prediction of *Trichoplax adhaerens* regulatory peptides. *General and Comparative Endocrinology*, *212*, 145–155.

Niklas, K. J., & Newman, S. A. (2013). The origins of multicellular organisms. *Evolution & Development*, *15*, 41–52.

Oakley, T. H. (2017). Furcation and fusion: The phylogenetics of evolutionary novelty. *Developmental Biology*, *431*, 69–76.

Owen, R. (1848). *The archetype and homologies of the vertebrate skeleton*. John van Voorst.

Packer, J. S., Zhu, Q., Huynh, C., Sivaramakrishnan, P., Preston, E., Dueck, H., Stefanik, D., Tan, K., Trapnell, C., Kim, J., Waterston, R. H., & Murray, J. I. (2019). A lineage-resolved molecular atlas of *C. elegans* embryogenesis at single-cell resolution. *Science*, *365*, eaax1971. https://doi.org/10.1126/science.aax1971

Paganos, P., Voronov, D., Musser, J. M., Arendt, D., & Arnone, M. I. (2021). Single-cell RNA sequencing of the *Strongylocentrotus purpuratus* larva reveals the blueprint of major cell types and nervous system of a non-chordate deuterostome. *eLife*, *10*, e70416. https://doi.org/10.7554/eLife.70416

Patthy, L. (2021). Exon shuffling played a decisive role in the evolution of the genetic toolkit for the multicellular body plan of metazoa. *Genes*, *12*, 382.

Pereira-Leal, J. B., Levy, E. D., Kamp, C., & Teichmann, S. A. (2007). Evolution of protein complexes by duplication of homomeric interactions. *Genome Biology*, *8*, R51.

Pereira-Leal, J. B., Levy, E. D., & Teichmann, S. A. (2006). The origins and evolution of functional modules: Lessons from protein complexes. *Philosophical Transactions of the Royal Society, B: Biological Sciences*, *361*, 507–517.

Peter, I. S., & Davidson, E. H. (2015). *Genomic control process: development and evolution*. Academic Press.

Pijuan-Sala, B., Griffiths, J. A., Guibentif, C., Hiscock, T. W., Jawaid, W., Calero-Nieto, F. J., Mulas, C., Ibarra-Soria, X., Tyser, R. C. V., Ho, D. L. L., Reik, W., Srinivas, S., Simons, B. D., Nichols, J., Marioni, J. C., & Göttgens, B. (2019). A single-cell molecular map of mouse gastrulation and early organogenesis. *Nature*, *566*, 490–495.

Plass, M., Solana, J., Wolf, F. A., Ayoub, S., Misios, A., Glažar, P., Obermayer, B., Theis, F. J., Kocks, C., & Rajewsky, N. (2018). Cell type atlas and lineage tree of a whole complex animal by single-cell transcriptomics. *Science*, *360*, eaaq1723. https://doi.org/10.1126/science.aaq1723

Poulin, J.-F., Tasic, B., Hjerling-Leffler, J., Trimarchi, J. M., & Awatramani, R. (2016). Disentangling neural cell diversity using single-cell transcriptomics. *Nature Neuroscience*, *19*, 1131–1141.

Putnam, N. H., Butts, T., Ferrier, D. E. K., Furlong, R. F., Hellsten, U., Kawashima, T., Robinson-Rechavi, M., Shoguchi, E., Terry, A., Yu, J. K., Benito-Gutiérrez, E., Dubchak, I., Garcia-Fernàndez, J., Gibson-Brown, J. J., Grigoriev, I. V., Horton, A. C., de Jong, P. J., Jurka, J., Kapitonov, V. V., … Rokhsar, D. S. (2008). The *Amphioxus* genome and the evolution of the chordate karyotype. *Nature*, *453*, 1064–1071.

Putnam, N. H., Srivastava, M., Hellsten, U., Dirks, B., Chapman, J., Salamov, A., Terry, A., Shapiro, H., Lindquist, E., Kapitonov, V. V., Jurka, J., Genikhovich, G., Grigoriev, I. V., Lucas, S. M., Steele, R. E., Finnerty, J. R., Technau, U., Martindale, M. Q., & Rokhsar, D. S. (2007). Sea anemone genome reveals ancestral eumetazoan gene repertoire and genomic organization. *Science*, *317*, 86–94.

Raff, R. A. (1996). *The shape of life: genes, development, and the evolution of animal form*. University of Chicago Press.

Rajagopal, J., & Stanger, B. Z. (2016). Plasticity in the adult: How should the Waddington diagram be applied to regenerating tissues? *Developmental Cell*, *36*, 133–137.

Rao, A., Barkley, D., França, G. S., & Yanai, I. (2021). Exploring tissue architecture using spatial transcriptomics. *Nature*, *596*, 211–220.

Regev, A., Teichmann, S. A., Lander, E. S., Amit, I., Benoist, C., Birney, E., Bodenmiller, B., Campbell, P., Carninci, P., Clatworthy, M., Clevers, H., Deplancke, B., Dunham, I., Eberwine, J., Eils, R., Enard, W., Farmer, A., Fugger, L., Göttgens, B., … Yosef, N. (2017). The human cell atlas. *eLife*, *6*, e27041. https://doi.org/10.7554/eLife.27041

Remane, A. (1963). The evolution of the Metazoa from colonial flagellates vs. plasmodial ciliates. In C. E. Dougherty (Ed.), *The lower Metazoa. Comparative biology and phylogeny* (pp. 23–32). University of California Press.

Richardson, M. K., & Keuck, G. (2002). Haeckel's ABC of evolution and development. *Biological Reviews of the Cambridge Philosophical Society*, *77*, 495–528.

Richter, D. J., Fozouni, P., Eisen, M. B., & King, N. (2018). Gene family innovation, conservation and loss on the animal stem lineage. *eLife*, *7*, e34226. https://doi.org/10.7554/eLife.34226

Riesgo, A., Farrar, N., Windsor, P. J., Giribet, G., & Leys, S. P. (2014). The analysis of eight transcriptomes from all poriferan classes reveals surprising genetic complexity in sponges. *Molecular Biology and Evolution*, *31*, 1102–1120.

Rompolas, P., Mesa, K. R., & Greco, V. (2013). Spatial organization within a niche as a determinant of stem-cell fate. *Nature*, *502*, 513–518.

Ros-Rocher, N., Pérez-Posada, A., Leger, M. M., & Ruiz-Trillo, I. (2021). The origin of animals: An ancestral reconstruction of the unicellular-to-multicellular transition. *Open Biology*, *11*, 200359.

Royuela, M., Fraile, B., Arenas, M. I., & Paniagua, R. (2000). Characterization of several invertebrate muscle cell types: A comparison with vertebrate muscles. *Microscopy Research and Technique*, *48*, 107–115.

Ruppert, E. E., Fox, R. S., & Barnes, R. D. (2004). *Invertebrate zoology: A functional evolutionary approach* (7th ed.). Brooks/Cole.

Rust, K., Byrnes, L. E., Yu, K. S., Park, J. S., Sneddon, J. B., Tward, A. D., & Nystul, T. G. (2020). A single-cell atlas and lineage analysis of the adult *Drosophila* ovary. *Nature Communications*, *11*, 5628.

Ryan, J. F., Pang, K., Schnitzler, C. E., Nguyen, A.-D., Moreland, R. T., Simmons, D. K., Koch, B. J., Francis, W. R., Havlak, P., Smith, S. A., Putnam, N. H., Haddock, S. H. D., Dunn, C. W., Wolfsberg, T. G., Mullikin, J. C., Martindale, M. Q., & Baxevanis, A. D. (2013). The genome of the ctenophore *Mnemiopsis leidyi* and its implications for cell type evolution. *Science*, *342*, 1242592. https://doi.org/10.1126/science.1242592

Sachwatkin, A. A. (1956). *Vergleichende Embryologie der niederen Wirbellosen: Ursprung und Gestaltungswege der individuellen Entwicklung der Vielzeller*. VEB Deutscher Verlag der Wissenschaften.

Salvini-Plawen, L. V. (2009). On the origin and evolution of the lower Metazoa. *Journal of Zoological Systematics and Evolutionary Research*, *16*, 40–87.

Samsel, Z., Sekretarska, J., Osinka, A., Wloga, D., & Joachimiak, E. (2021). Central apparatus, the molecular kickstarter of ciliary and flagellar nanomachines. *International Journal of Molecular Sciences*, *22*, 3013.

Schlosser, G. (2021). *Evolutionary origin of sensory and neurosecretory cell types* (1st ed.). CRC Press.

Scholtz, G. (2004). Baupläne versus ground patterns, phyla versus monophyla: aspects of patterns and processes in evolutionary developmental biology. In G. Scholtz (Ed.), *Evolutionary developmental biology of Crustacea. Crustacean Issues 15* (Vol. 15, pp. 3–16). A. A. Balkema Publishers.

Schwaiger, M., Schönauer, A., Rendeiro, A. F., Pribitzer, C., Schauer, A., Gilles, A. F., Schinko, J. B., Renfer, E., Fredman, D., & Technau, U. (2014). Evolutionary conservation of the eumetazoan gene regulatory landscape. *Genome Research*, *24*, 639–650.

Sebé-Pedrós, A., Ballaré, C., Parra-Acero, H., Chiva, C., Tena, J. J., Sabidó, E., Gómez-Skarmeta, J. L., Di Croce, L., & Ruiz-Trillo, I. (2016). The dynamic regulatory genome of *Capsaspora* and the origin of animal multicellularity. *Cell*, *165*, 1224–1237.

Sebé-Pedrós, A., Chomsky, E., Pang, K., Lara-Astiaso, D., Gaiti, F., Mukamel, Z., Amit, I., Hejnol, A., Degnan, B. M., & Tanay, A. (2018). Early metazoan cell type diversity and the evolution of multicellular gene regulation. *Nature Ecology & Evolution*, *2*, 1176–1188.

Sebé-Pedrós, A., Irimia, M., del Campo, J., Parra-Acero, H., Russ, C., Nusbaum, C., Blencowe, B. J., & Ruiz-Trillo, I. (2013). Regulated aggregative multicellularity in a close unicellular relative of Metazoa. *eLife*, *2*, e01287. https://doi.org/10.7554/eLife.01287

Sebé-Pedrós, A., & de Mendoza, A. (2015). Transcription factors and the origin of animal multicellularity. In I. Ruiz-Trillo, & A. Nedelcu (Eds.), *Evolutionary transitions to multicellular life. Advances in marine genomics* (Vol. 2, pp. 379–394). Springer.

Sebé-Pedrós, A., Saudemont, B., Chomsky, E., Plessier, F., Mailhé, M. P., Renno, J., Loe-Mie, Y., Lifshitz, A., Mukamel, Z., Schmutz, S., Novault, S., Steinmetz, P. R. H., Spitz, F., Tanay, A., & Marlow, H. (2018). Cnidarian cell type diversity and regulation revealed by whole-organism single-cell RNA-Seq. *Cell*, *173*, 1520–1534.

Seipel, K., & Schmid, V. (2005). Evolution of striated muscle: Jellyfish and the origin of triploblasty. *Developmental Biology*, *282*, 14–26.

Senatore, A., Reese, T. S., & Smith, C. L. (2017). Neuropeptidergic integration of behavior in *Trichoplax adhaerens*, an animal without synapses. *Journal of Experimental Biology*, *220*, 3381–3390.

Shafer, M., Sawh, A. N., & Schier, A. F. (2022). Gene family evolution underlies cell-type diversification in the hypothalamus of teleosts. *Nature Ecology & Evolution*, *6*, 63–76.

Shao, Y., Wang, X.-B., Zhang, J.-J., Li, M.-L., Wu, S.-S., Ma, X.-Y., Wang, X., Zhao, H. F., Li, Y., Zhu, H. H., Irwin, D. M., Wang, D. P., Zhang, G. J., Ruan, J., & Wu, D. D. (2020). Genome and single-cell RNA-sequencing of the earthworm *Eisenia andrei* identifies cellular mechanisms underlying regeneration. *Nature Communications*, *11*, 2656.

Shapiro, E., Biezuner, T., & Linnarsson, S. (2013). Single-cell sequencing-based technologies will revolutionize whole-organism science. *Nature Reviews Genetics*, *14*, 618–630.

Sheng, X. R., Brawley, C. M., & Matunis, E. L. (2009). Dedifferentiating spermatogonia outcompete somatic stem cells for niche occupancy in the *Drosophila* testis. *Cell Stem Cell*, *5*, 191–203.

Siebert, S., Farrell, J. A., Cazet, J. F., Abeykoon, Y., Primack, A. S., Schnitzler, C. E., & Juliano, C. E. (2019). Stem cell differentiation trajectories in *Hydra* resolved at single-cell resolution. *Science*, *365*, eaav9314. https://doi.org/10.1126/science.aav9314

de Silva, M. I., Stringer, B. W., & Bardy, C. (2022). Neuronal and tumourigenic boundaries of glioblastoma plasticity. *Trends in Cancer*, *S2405-8033*, 00233-3. https://doi.org/10.1016/j.trecan.2022.10.010

Simakov, O., Bredeson, J., Berkoff, K., Marletaz, F., Mitros, T., Schultz, D. T., O'Connell, B. L., Dear, P., Martinez, D. E., Steele, R. E., Green, R. E., David, C. N., & Rokhsar, D. S. (2022). Deeply conserved synteny and the evolution of metazoan chromosomes. *Science Advances*, *8*, 5884.

Simakov, O., Kawashima, T., Marlétaz, F., Jenkins, J., Koyanagi, R., Mitros, T., Hisata, K., Bredeson, J., Shoguchi, E., Gyoja, F., Yue, J. X., Chen, Y. C., Freeman, R. M., Sasaki, A., Hikosaka-Katayama, T., Sato, A., Fujie, M., Baughman, K. W., Levine, J., … Gerhart, J. (2015). Hemichordate genomes and deuterostome origins. *Nature*, *527*, 459–465.

Simakov, O., Marletaz, F., Cho, S.-J., Edsinger-Gonzales, E., Havlak, P., Hellsten, U., Kuo, D. H., Larsson, T., Lv, J., Arendt, D., Savage, R., Osoegawa, K., de Jong, P., Grimwood, J., Chapman, J. A., Shapiro, H., Aerts, A., Otillar, R. P., Terry, A. Y., … Rokhsar, D. S. (2013). Insights into bilaterian evolution from three spiralian genomes. *Nature*, *493*, 526–531.

Slack, J. M. W., Holland, P. W. H., & Graham, C. F. (1993). The zootype and the phylotypic stage. *Nature*, *361*, 490–492.

Sladitschek, H. L., Fiuza, U.-M., Pavlinic, D., Benes, V., Hufnagel, L., & Neveu, P. A. (2020). MorphoSeq: Full single-cell transcriptome dynamics up to gastrulation in a chordate. *Cell*, *181*, 922–935.

Slaidina, M., Banisch, T. U., Gupta, S., & Lehmann, R. (2020). A single-cell atlas of the developing *Drosophila* ovary identifies follicle stem cell progenitors. *Genes & Development*, *34*, 239–249.

Smith, C. L., Varoqueaux, F., Kittelmann, M., Azzam, R. N., Cooper, B., Winters, C. A., Eitel, M., Fasshauer, D., & Reese, T. S. (2014). Novel cell types, neurosecretory cells, and body plan of the early-diverging metazoan *Trichoplax adhaerens*. *Current Biology*, *24*, 1565–1572.

Sogabe, S., Hatleberg, W. L., Kocot, K. M., Say, T. E., Stoupin, D., Roper, K. E., Fernandez-Valverde, S. L., Degnan, S. M., & Degnan, B. M. (2019). Pluripotency and the origin of animal multicellularity. *Nature*, *570*, 519–522.

Spitz, F., & Furlong, E. E. M. (2012). Transcription factors: From enhancer binding to developmental control. *Nature Reviews Genetics*, *13*, 613–626.

Srivastava, M., Begovic, E., Chapman, J., Putnam, N. H., Hellsten, U., Kawashima, T., Kuo, A., Mitros, T., Salamov, A., Carpenter, M. L., Signorovitch, A. Y., Moreno, M. A., Kamm, K., Grimwood, J., Schmutz, J., Shapiro, H., Grigoriev, I. V., Buss, L. W., Schierwater, B., … Rokhsar, D. S. (2008). The *Trichoplax* genome and the nature of placozoans. *Nature*, *454*, 955–960.

Srivastava, M., Simakov, O., Chapman, J., Fahey, B., Gauthier, M. E. A., Mitros, T., Richards, G. S., Conaco, C., Dacre, M., Hellsten, U., Larroux, C., Putnam, N. H., Stanke, M., Adamska, M., Darling, A., Degnan, S. M., Oakley, T. H., Plachetzki, D. C., Zhai, Y., … Rokhsar, D. S. (2010). The *Amphimedon queenslandica* genome and the evolution of animal complexity. *Nature*, *466*, 720–726.

Stefanakis, N., Carrera, I., & Hobert, O. (2015). Regulatory logic of pan-neuronal gene expression in *C. elegans*. *Neuron*, *87*, 733–750.

Stegle, O., Teichmann, S. A., & Marioni, J. C. (2015). Computational and analytical challenges in single-cell transcriptomics. *Nature Reviews Genetics*, *16*, 133–145.

Steinmetz, P. R. H., Kraus, J. E. M., Larroux, C., Hammel, J. U., Amon-Hassenzahl, A., Houliston, E., Wörheide, G., Nickel, M., Degnan, B. M., & Technau, U. (2012). Independent evolution of striated muscles in cnidarians and bilaterians. *Nature*, *487*, 231–234.

Sulston, J. E., Schierenberg, E., White, J. G., & Thomson, J. N. (1983). The embryonic cell lineage of the nematode *Caenorhabditis elegans*. *Developmental Biology*, *100*, 64–119.

Tam, P. P. L., & Gad, J. M. (2004). Gastrulation in the mouse embryo. In C. D. Stern (Ed.), *Gastrulation: From cell to embryo* (pp. 233–262). Cold Spring Harbor Laboratory Press.

Tanay, A., & Sebé-Pedrós, A. (2021). Evolutionary cell type mapping with single-cell genomics. *Trends in Genetics*, *37*, 919–932.

Tarashansky, A. J., Musser, J. M., Khariton, M., Li, P., Arendt, D., Quake, S. R., & Wang, B. (2021). Mapping single-cell atlases throughout Metazoa unravels cell type evolution. *eLife*, *10*, 1–24.

Tarlow, B. D., Pelz, C., Naugler, W. E., Wakefield, L., Wilson, E. M., Finegold, M. J., & Grompe, M. (2014). Bipotential adult liver progenitors are derived from chronically injured mature hepatocytes. *Cell Stem Cell*, *15*, 605–618.

Tasic, B. (2018). Single cell transcriptomics in neuroscience: Cell classification and beyond. *Current Opinion in Neurobiology*, *50*, 242–249.

Tata, P. R., Mou, H., Pardo-Saganta, A., Zhao, R., Prabhu, M., Law, B. M., Vinarsky, V., Cho, J. L., Breton, S., Sahay, A., Medoff, B. D., & Rajagopal, J. (2013). Dedifferentiation of committed epithelial cells into stem cells *in vivo*. *Nature*, *503*, 218–223.

Temereva, E. N., & Malakhov, V. v (2012). Embryogenesis in phoronids. *Invertebrate Zoology*, *9*, 1–39.

The Tabula Muris Consortium. (2018). Single-cell transcriptomics of 20 mouse organs creates a Tabula Muris. *Nature*, *562*, 367–372.

Thorel, F., Népote, V., Avril, I., Kohno, K., Desgraz, R., Chera, S., & Herrera, P. L. (2010). Conversion of adult pancreatic α-cells to β-cells after extreme β-cell loss. *Nature*, *464*, 1149–1154.

Tosches, M. A., Yamawaki, T. M., Naumann, R. K., Jacobi, A. A., Tushev, G., & Laurent, G. (2018). Evolution of pallium, hippocampus, and cortical

cell types revealed by single-cell transcriptomics in reptiles. *Science*, *360*, 881–888.

Trapnell, C. (2015). Defining cell types and states with single-cell genomics. *Genome Research*, *25*, 1491–1498.

Uesaka, M., Kuratani, S., & Irie, N. (2022). The developmental hourglass model and recapitulation: An attempt to integrate the two models. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, *338*, 76–86.

Valentine, J. W. (2003). Cell types, numbers, and body plan complexity. In B. K. Hall, & W. M. Olson (Eds.), *Keywords and concepts in evolutionary developmental biology* (pp. 35–43). Harvard University Press.

Vogt, G. (2022). Paradigm shifts in animal epigenetics: Research on non-model species leads to new insights into dependencies, functions and inheritance of DNA methylation. *BioEssays*, *44*, 2200040.

Waddington, C. H. (1957). *The strategy of the genes*. George Allen and Unwin.

Wagner, D. E., Weinreb, C., Collins, Z. M., Briggs, J. A., Megason, S. G., & Klein, A. M. (2018). Single-cell mapping of gene expression landscapes and lineage in the zebrafish embryo. *Science*, *360*, 981–987.

Wagner, G. P. (2007). The developmental genetics of homology. *Nature Reviews Genetics*, *8*, 473–479.

Wagner, G. P. (2014). *Homology, genes, and evolutionary innovation*. Princeton University Press.

Wagner, G. P. (2015). Homology in the age of developmental genomics, *Evolutionary developmental biology of invertebrates 1* (pp. 25–43). Springer.

Wang, J., Sun, H., Jiang, M., Li, J., Zhang, P., Chen, H., Mei, Y., Fei, L., Lai, S., Han, X., Song, X., Xu, S., Chen, M., Ouyang, H., Zhang, D., Yuan, G. C., & Guo, G. (2021). Tracing cell-type evolution by cross-species comparison of cell atlases. *Cell Reports*, *34*, 108803.

Wei, Z., Angerer, R. C., & Angerer, L. M. (2011). Direct development of neurons within foregut endoderm of sea urchin embryos. *Proceedings of the National Academy of Sciences*, *108*, 9143–9147.

Weir, K., Dupre, C., van Giesen, L., Lee, A. S.-Y., & Bellono, N. W. (2020). A molecular filter for the cnidarian stinging response. *eLife*, *9*, 1–20.

Wilson, E. B. (1925). *The cell in development and heredity (Third edition, revised and enlarged)*. The Macmillan Co.

Wolpert, L., Tickle, C., Martinez-Arias, A., & Locke, J. (2019). *Principles of development* (6th ed.). Oxford University Press.

Wong, E., Mölter, J., Anggono, V., Degnan, S. M., & Degnan, B. M. (2019). Co-expression of synaptic genes in the sponge *Amphimedon queenslandica* uncovers ancient neural submodules. *Scientific Reports*, *9*, 15781.

Wong, E. S., Zheng, D., Tan, S. Z., Bower, N. I., Garside, V., Vanwalleghem, G., Gaiti, F., Scott, E., Hogan, B. M., Kikuchi, K., McGlinn, E., Francois, M., & Degnan, B. M. (2020). Deep conservation of the enhancer regulatory code in animals. *Science*, *370*, eaax8137. https://doi.org/10.1126/science.aax8137

Wu, L., Ferger, K. E., & Lambert, J. D. (2019). Gene expression does not support the developmental hourglass model in three animals with spiralian development. *Molecular Biology and Evolution*, *36*, 1373–1383.

Wu, X., Zhang, H., Zhang, B., Zhang, Y., Wang, Q., Shen, W., Wu, X., Li, L., Xia, W., Nakamura, R., Liu, B., Liu, F., Takeda, H., Meng, A., & Xie, W. (2021). Methylome inheritance and enhancer dememorization reset an epigenetic gate safeguarding embryonic programs. *Science Advances*, *7*, eabl3858. https://doi.org/10.1126/sciadv.abl3858

Xu, Y., Zhao, J., Ren, Y., Wang, X., Lyu, Y., Xie, B., Sun, Y., Yuan, X., Liu, H., Yang, W., Fu, Y., Yu, Y., Liu, Y., Mu, R., Li, C., Xu, J., & Deng, H. (2022). Derivation of totipotent-like stem cells with blastocyst-like structure forming potential. *Cell Research*, *32*, 513–529.

Yanay, C., Morpurgo, N., & Linial, M. (2008). Evolution of insect proteomes: Insights into synapse organization and synaptic vesicle life cycle. *Genome Biology*, *9*, R27.

Yanger, K., Zong, Y., Maggs, L. R., Shapira, S. N., Maddipati, R., Aiello, N. M., Thung, S. N., Wells, R. G., Greenbaum, L. E., & Stanger, B. Z. (2013). Robust cellular reprogramming occurs spontaneously during liver regeneration. *Genes & Development*, *27*, 719–724.

Yeates, D. K. (1995). Groundplans and exemplars: Paths to the tree of life. *Cladistics*, *11*, 343–357.

Zakhvatkin, A. A. (1949). *The comparative embryology of the low invertebrates. Sources and method of the origin of metazoan development*. Soviet Science.

Zemach, A., McDaniel, I. E., Silva, P., & Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science*, *328*, 916–919.

Zeng, H., & Sanes, J. R. (2017). Neuronal cell-type classification: Challenges, opportunities and the path forward. *Nature Reviews Neuroscience*, *18*, 530–546.

Zheng, G. X. Y., Terry, J. M., Belgrader, P., Ryvkin, P., Bent, Z. W., Wilson, R., Ziraldo, S. B., Wheeler, T. D., McDermott, G. P., Zhu, J., Gregory, M. T., Shuga, J., Montesclaros, L., Underwood, J. G., Masquelier, D. A., Nishimura, S. Y., Schnall-Levin, M., Wyatt, P. W., Hindson, C. M., … Bielas, J. H. (2017). Massively parallel digital transcriptional profiling of single cells. *Nature Communications*, *8*, 14049.

Zimmermann, B., Robert, N. S. M., Technau, U., & Simakov, O. (2019). Ancient animal genome architecture reflects cell type identities. *Nature Ecology & Evolution*, *3*, 1289–1293.

How to cite this article: Rusin, L. Y. (2023). Evolution of homology: From archetype towards a holistic concept of cell type. *Journal of Morphology*, 284, e21569. https://doi.org/10.1002/jmor.21569