Concept Paper

Reading the Palimpsest of Cell Interactions: What Questions May We Ask of the Data?

Mihaela Pavlicev [*] and Günter Wagner

Concept Paper

Reading the Palimpsest of Cell Interactions: What Questions May We Ask of the Data?

Mihaela Pavlicev [1,2,*] and Günter Wagner [1,3]

[1]   Unit for Theoretical Biology, Department for Evolutionary Biology, University of Vienna, Austria
[2]   Complexity Science Hub, Vienna, Austria
[3]   Yale University, USA
[*]   Correspondence: Mihaela Pavlicev, Department for Evolutionary Biology, University of Vienna, Austria; mihaela.pavlicev@univie.ac.at

Abstract: Biological function depends on the composition as well as structure of the organism, that is the organization of interactions between its parts. In multicellular organisms, the complex communication between cells is crucial for the coordination of cellular functions (growth, differentiation, apoptosis), compartmentalization of cellular processes within the organism, and the integration of cells into functioning higher levels of structural organization. Here we focus on the interaction among cells, as the basic level of biological organization, capable of a remarkable degree of autonomy and self-maintenance. While a wealth of data on putative cell interactions has recently become available, its biological interpretation depends on expectations we have for the structure of these interaction networks. This contribution is an attempt at formulating basic questions we may ask about cell interaction data. We build on the understanding that cells fulfill two general functions: the integrity-maintaining and the organismal service function. We derive the expected patterns of cell interactions considering two intertwined aspects: the immediate functional and the evolutionary. Based on these, we propose several guidelines for analysis and interpretation of transcriptional cell-interactome data.

Keywords: cell communication; tissue structure; multicellularity; levels of organization; self-maintenance

Introduction

The importance of the interactions between parts in constituting the structure and function of organisms has been powerfully illustrated by the physiologist Paul Weiss (1898-1989): an intact fertilized chicken egg has the potential to develop into a chicken. If, as Paul Weiss did for his students, put into a blender (and switched on), the embryo turns into a mess which consists of exact same molecular parts as the embryo, but is not able to develop into a chicken. Structure, arising from past and current interactions between parts (cells, tissues and organs), matters to the proper functioning of an organism. We can apply this thinking, in principle, to any level of organization, but to turn the general insight into productive research requires appropriate data. With the development of single cell technology, we can detect potential signaling interactions at scale for many organisms and a myriad of tools is able to derive the cell interactomes from this data (e.g., 1, 2-5). In this paper, we want to propose a few guiding perspectives to analyze and interpret cell-cell interaction networks.

The focus on cells advocated here has its justification in the degree of autonomy that cells manifest. Cells can exist as organisms, and even where their degree of autonomy is reduced, as in multicellular organisms, cells manifest their autonomy in the fact that some are surprisingly easily grown in cell culture. Levels of organization other than the cell interactions have gained much attention: transcriptional regulatory networks and protein interactions have been measured and modeled extensively (6-12). Both levels, even though often without consideration of cellular or other

structural context, led to interesting systemic implications, which may become ever more relevant if put into the context of organismal structure.

Beyond Weiss' early argument cited above, it is self-evident that many of cellular functions depend on interactions with or signals from other cells, the ECM and the physico-chemical conditions of their tissue niche, rather than being inherent to the cells (although the potential to interact is dependent on the cell's competence based on expressing ligands/receptors, secondary messengers or otherwise sensing its environment). Overall, it can be stated that the characteristics of cells are not entirely inherent to the cells themselves but arise from cell-external factors as well as. Cell external factors consist of those emanating from the cell's environment, that is other cells, as well as the extra organismal environment. In fact, many extant mechanisms for intercellular communication in multicellular organisms have evolved from mechanisms for environment-sensing in unicellular organisms. For instance, hormone nuclear receptors such as steroid receptors evolved from xeno-sensors responsive to free fatty acids (13, 14).

Recent decades have seen an upsurge of technological innovations to study processes at single cell resolution. This resulted in numerous cell atlases and even cell interaction atlases available for many tissues, organs (15) and even, in some cases, whole animals (16-18). Interpreting the abundant data on [putative] cell interactions is facing many challenges. On the one hand there is the technical challenge of how to determine putative cell-cell interactions: especially in non-model species it is difficult to apply the often very promiscuously assembled databases of potential ligand-receptor pairs. Some of the most inclusive databases contain many false interactions, e.g., mistaking co-receptors as ligands. Assuming that the technical issues can be overcome, a complementary challenge is the question, how do we extract biologically meaningful information from the "hairball" of putative cell interactions?

We will not address the problem of quality of ligand receptor databases, beyond stating that critical manual curation is highly advisable when compiling putative cell interactomes; the cell-cell interactions databases may need to be limited to those experimentally confirmed (rather than collected by text mining), the interactions with ECM might need to be evaluated separately from secreted ligand-mediated interactions etc. The challenge we want to address here is the second one: the biological interpretation of cell interaction networks. By biological interpretation we mean the analysis of putative cell-cell interactions beyond statistical summaries of the number and pattern of interactions, by integrating the observed patterns into the biological background knowledge.

101 on cell interactions: not all interactions are alike.

Before continuing towards our aim, we will briefly summarize the basic cell biology of cell-cell signaling. While all points summarized below are textbook knowledge it is still important to articulate them here, as the summary statistics of "the number of interactions," more often than not, lump biologically heterogenous forms of interactions and thereby potentially confuse more than enlighten the biology at hand.

- Cells communicate with <u>two basic types of signals</u>. At the one hand there are secreted ligands that can reach cells beyond the sender cell's immediate neighborhood. On the other hand, there are ligands that are embedded in the cell membrane that need physical contact with the receiving cell to reach the cell surface receptors of their immediate neighbors (juxtracrine signaling). In the latter case inference of cell interactions is impossible without anatomical information about the spatial localization of the cell types. In addition, there are secreted but not diffusible molecules that also can interact with cell surface receptors like ECM components (see below).
- One talks of <u>allocrine signaling</u> when the signaling is reaching cells of a different type. <u>Autocrine signaling</u> is to cells of the same cell type (we will avoid the terms paracrine and endocrine as these relate to the distance over which signals are acting). Cells of the same cell type present the same spectrum of ligands and receptors, and therefore the signaling to the same cell type, not just that to an individual cell, is also considered "signaling to self."

- The ECM can act as a ligand (19). Can we say that the cell that produces ECM components is signaling to ECM attaching cells? Does this form of signaling play the same role as secreted and diffusible signals, like small molecule ligands, say prostaglandins? It is important to clarify this question before engaging in a statistical analysis of cell-cell interaction networks that consistently point to fibroblasts as hub of communication because they produce most of the ECM components. We propose to separate these interactions from those mediated by secreted and diffusible ligands.
- The ECM can act as a repository for secreted ligands. As a consequence, ECM degrading enzymes can cause the release of ligands even if the cell is not producing the ligand itself.
- For all we know, the secreted ligands are anonymous, i.e., they do not have an imprint of their origin. They are perceived by all cells that present the cognate receptor, including the secreting cell (autocrine effects) without recognizing the difference between self (molecules emitted from the same cell type) and non-self (the same ligand emitted from another cell type). This is most certainly the case for small molecule ligands, but peptide ligands are often post-transcriptionally modified and could, in principle, have a cell type of origin imprint ("return address"). We are not aware that this possibility is a biological reality.
- Modulation of signaling can occur by changing the expression (or synthesis) of signal and the receptor, by production or inhibition of decoy receptors or binding proteins, or by the degradation of the ligands or the receptors.

These basic molecular principles define the complexity of what is often simply summarized as "cell interactions". We propose that the statistical analysis of any data set should be stratified by the form and nature of potential interactions to gain biological validity.

Patterns of cell-interactions

No doubt there will be many highly specific expectations for every dataset that are based on the biology of a particular tissue or organ, in a particular species. These are not the kind of patterns that we wish to address here. Instead, we wish to address the theoretical expectations for cell-interactions that can be formulated at an abstract level and that may be applicable to cell interaction studies in general. These expectations arise from two intertwined sources –from functional and evolutionary considerations.

Functional Considerations

Cells either exist as individuals, i.e., as single celled organisms, or as parts of multicellular organisms. Multicellular organism come with various degrees of sub-organization of cells, such as tissues, organs and organ systems. Cnidarians and ctenophores are composed mostly of epithelia, but do not have distinct organs, while on the other extreme vertebrates are composed of many different tissue types organized in spatially segregated organs, like liver and kidneys. As these subsystems are themselves specialized for specific functions, this hierarchical organization suggests that cellular functions can be dedicated to two complementary roles. At the one hand there are activities that serve the functional role of the next higher organizational level, the tissue, the organ etc. These activities we want to call *service functions* (S-functions). Examples are the secretion of hormones that act at the systemic level. On the other hand, there are activities that are necessary for constituting and maintaining the cells themselves as well as *tissue or organ integrity* (I-functions for "integrity functions"). The fact that I-functions also ultimately benefit the organism because they enable the conditions necessary to execute S-functions is acknowledged but does not invalidate the distinction between S- and I-functions. The I-functions are necessary for the S-functions, but the opposite is not necessarily true. This would suggest stronger conservation of interactions involved in I-functions, relative to those involved in S-functions, as will be discussed later. The integration of I- and S-functions at the level of the subsystem might be expected to constrain the organization of cell interaction networks.

For the tissue level, a recent theory of tissue organization postulates that most tissues consist of a generic core module of four to five cell categories regardless of their specific functional role (Ruslan

Medzhitov, personal communication, discussed in 20). This theory will serve here to explicate the integration of S and I functions. Medzhitov model suggest that these cell categories include parenchymatic cells, i.e., a cell population that is specialized for an S-function, i.e., the functionally specialized cell type of the tissue, such as hepatocytes in the liver. The remaining cell categories are fibroblasts, macrophages and endothelial cells. Their crucial contributions are primarily I-functions: producing ECM (fibroblasts), monitor tissue integrity and remove cell debris (macrophages), and ensuring oxygen supply (endothelial cells). Other cells of the tissue are so-called ancillary cells, specialized for supporting the parenchymatic cells. An example of the latter are Schwann cells for peripheral nerve cells.

According to the Medzhitov model, the signaling relationships *among* these cell categories, including the parenchymatic cells, are expected to be mostly homeostatic, ensuring the adequate proportion of these cells in the tissue and thus an adequate supply of their specific functional contributions. A mutual regulatory relationship between fibroblasts and macrophages, mediated by PDGFa and CSF1, has experimentally been demonstrated (21), supporting the notion that cell-cell interactions *within a tissue module* are, to a large degree, homeostatic.

The broad strokes pattern summarized above is of course a simplification. Fibroblasts and macrophages come in tissue-specific varieties, which also perform service functions in addition to their I-functions. For instance, osteoclasts - specialized macrophages of bone tissue- play an important role in bone remodeling, and lung macrophages clear surplus surfactants from the alveoli, ensuring proper function of the lung. There are also other exceptions: cartilage, for example, consists of a single cell type, the chondrocyte. Nevertheless, the Medhzitov model raises interesting questions with respect to the expected patterns of cell-cell interactions.

The mutual homeostatic interactions between the cells of a tissue module could conceivably be mediated by a generic set of ligands even in tissues with different S-functions. For instance, PDGFa could be used in any tissue to stimulate the replication of fibroblasts, CSF1 for macrophages, VEGFA for endothelial cells and EGF for parenchymatic epithelial cells. As each tissue is spatially contiguous and locally exclusive with respect to other tissues that would suffice to ensure homeostatic maintenance of the tissue core module. To our knowledge, this question has not been addressed on a broad comparative basis.

Another question is whether the ligands that mediate the homeostatic interactions are always expressed or only when the equilibrium is disturbed. For instance, the expression of VEGFA would only be detectable if the tissue is in a hypoxic state. This means that in a single cell transcriptomic study of an intact (before cell separation) tissue ligand-receptor interactions dedicated to the homeostatic regulation of tissue composition might not be detectable, but still not contradict the existence of such a network.

Then there is the question whether the network of homeostatic cell interactions is fully connected, meaning that each cell can stimulate each other cell to either increase or decrease in number if the cellular composition of the tissue is in dis-equilibrium? Or is there a hierarchical structure to the intra-tissue communication maintaining tissue integrity? The former would allow each cell category to call in additional help if it is experiencing some strain, but not necessarily the whole tissue. The latter possibility might allow a more centralized decision-making allowing the integration of various streams of information.

In any case, this way of thinking about cell-cell interactions in tissues challenges researchers to identify the interactions dedicated to S-functions and those necessary for the I-functions in a particular tissue and to identify the network topology dedicated to either.

Expected patterns of cell interactions due to evolutionary considerations

A dimension that has been less explored so far, is how patterns of cell interactions are shaped in evolution and thus what is to be expected when we compare interaction networks between species. This is our aim in this section. We develop several hypotheses about the evolution of patterns of intercellular interaction.

Internal, in addition to external selection

As mentioned above, the service and integration functions of cell interactions are under different evolutionary constraints. S-functions are more guided and directed toward the environmental adaptive needs as well as the I-functions of the next higher unit. I-functions are shaped to ensure the integrity of the tissues and organs. S- and I-functions are thus roughly subject to external (adaptive) selection and internal selection respectively (22-26) . Given these broad patterns, the cellular interactions underlying S-functions are likely more variable among species and certainly among parts of the organism, i.e., are specific to tissue and organ types, because they define the functional role of the respective higher-level unit (tissue or organ, or body part). In contrast the I-functions and the underlying cellular interaction patterns are likely more conserved and generic as they ensure the existence and integrity of the higher order unit, independent of the specific S-function performed.

The double contingency constraint

A signaling interaction between two cell types, say X $\rightarrow$ Y, is the result of two evolutionarily contingent outcomes: the expression of a ligand in cell type X and the simultaneous expression of the cognate receptor in cell type Y. This obvious fact has important implication, namely that the expression of ligand and receptor, is contingent, i.e., there is no mechanistic necessity that either is expressed in a particular cell type. Because of this, the *de novo* evolution of a signaling relationship is unlikely, as the two scenarios below briefly illustrate.

The first scenario is that two independent mutations, one causing the expression of the ligand specifically in cell type X and the other causing the expression of the receptor specifically in cell type Y, occur simultaneously to produce a functional signaling relationship. Each mutation separately would be functionally irrelevant or even costly, as the cell would waste energy to produce a useless product. The only way natural selection could be able to pick up on such a *de novo* interaction is if these two mutations occur simultaneously in the same individual. If they occur in different individuals in the same population, the likelihood that they combine in the same genome of a diploid species is initially $(1/2N_e)^2$, with N_e being the effective population size. The likelihood of recombining into the same genome (and thus becoming functionally relevant) is thus very low. An alternative scenario would be for a single mutation to change the expression of the ligand in X and the receptor in Y simultaneously. We are not aware of a molecular mechanism that would make such a mutation plausible. Hence, the evolution of functionally relevant cell-cell interactions is more likely to occur through evolutionary steps with single mutations causing the expression or repression of either a ligand or a receptor in a particular cell type or cell type family. This principle severely constrains the plausible evolutionary scenarios for the evolution of a cell-cell communication network. One such scenario is the pruning of autocrine signals discussed in the next section. Pruning only requires the removal of one of the two components (ligand or receptor) to interrupt and thus re-shape the interaction network.

Pruning of autocrine interactions

We can envision two settings in which pruning of cell interaction would be a way to generate cell's interactome in evolution: the evolution of cell's autonomy, and the individualization against the external context.

<u>*Evolving cell type identity*</u> Cells are the building blocks of multicellular organisms, and correspondingly, organismal complexity is often estimated as the number of distinct cell types in an organism (27-29). A relevant evolutionary question in this context is how do novel cell types originate? A broad-based effort to address this question is underway by documenting the phylogenetic origin of cell types and the underlying mechanism in terms of the intrinsic gene regulatory logic of cell type identities (30). However, given that cell types play specific roles in the broader organismal context, an important aspect of their origination is how new cell types are integrated into the network of cell-cell interactions in a tissue or organ.

To address this question, one can think of the possible cell interaction networks as lying on a spectrum between two (theoretical) extremes: on the one side is an aggregate of cells of the same type, and on the other an organism with as many cell types as there are cells, possibly approximated by euthelic animals, i.e., small animals with constant cell numbers, like nematodes and rotifers, in which each cell performs a different essential function (31). In an aggregate of identical cell types, all cells of the same cell type are producing the same ligands and receptors, and therefore all intercellular signaling is autocrine and ubiquitous. In the other extreme, the cell types never express receptors for ligands they themselves produce. In this case all interactions would be allocrine.

From this perspective, the divergence between cell type identities would be reflected in the divergence between the spectrum of ligands emitted and signals that can be received. In this process, autocrine interactions are transformed into allocrine ones by pruning autocrine mediators, i.e., differentially shutting down the expression of ligands and/or receptors or expressing new signal mediators (Figure 1). By pruning, a set of cells that originally were of the same type, expressing the same ligands and receptors, are turned into two different cell types with directional allocrine signaling between sister cell types. To test this, one could expect that the pattern of ligand-receptor pairs in sister cell types reflects the full autocrine signals from the ancestral cell types. This would be detectable by comparing the ligand-receptor spectrum of sister cell types with that of the homologous single cell type in an outgroup species.

One proxy to express where on this spectrum of individualization a cell type is, is to determine the proportion of allocrine signaling compared to the total signaling a cell type engages in. That is, the proportion of receptors, out of all expressed receptors, for which the signals are not produced by the cell type itself, and the same for the ligands. These proportions can be expected to increase in the process of cell type divergence. One may thus distinguish individualized cell types that to a great extent rely on allocrine signaling, from less autonomous in which a great portion of cell signaling is potentially autocrine.
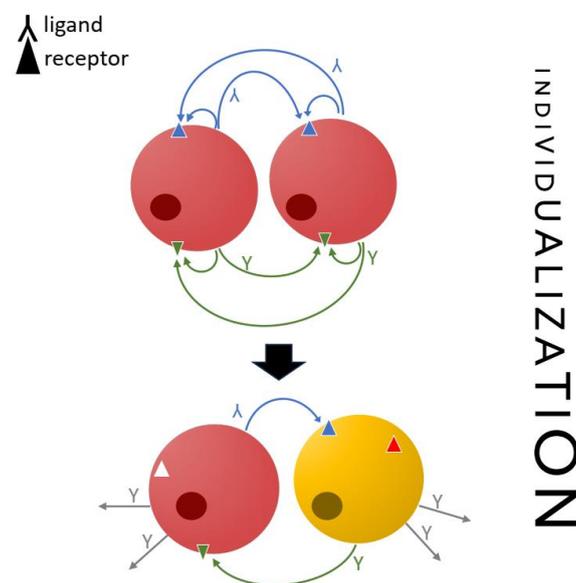

Figure 1

Novel spatial relationships generate novel interactions During embryonic development and during evolution, the spatial relationship and thus the potential for signaling between tissues, can change. In particular in early development, induction via signaling between tissues is common. For instance, during the development of the amphibian embryo the epidermis that eventually develops the lens of the eye interacts with a variety of cell populations, including pharyngeal and heart mesenchyme, in addition to the eyestalk.

In evolution, epidermis can contact different kinds of dermal tissue with implications for the epidermis differentiation. For instance, whether rodents form cheek pouches lined with mucosal tissue or fur depends on the location of the cheek pouch invagination point and thus what induction tissue the pouch is exposed to (32). Another dramatic novel tissue contact is the evolution of placentation, where extra-embryonic tissues of the embryo get into contact with the uterine endometrium, if/when the eggshell is lost. A question to ask is, how likely are new spatial relationships compatible with the cell-cell signaling relationships in each of the participating tissues?

The above model by Medzhitov suggests that the cell type categories (e.g., parenchyme, macrophage…) occur repeatedly across tissues. If so, can we expect that the new cell interactions will more readily be accommodated if involved cell categories have evolved in a respective tissue context, such as epithelial cells that have been in contact with stromal cells and become introduced to a different stromal cell type at a new interface?   A potential example in which to ask this question is the abovementioned origin of the maternal-fetal interface in mammalian viviparity. In mammals with hemo- or endotheliochorial placentation the uterine luminal epithelia are eroded during the placental development and replaced with placental (=epithelial) cells, resulting in the new maternal-fetal tissue unit.

Adapting to the stability of tissue context   The cell type-specific profiles of ligand and receptor expression may be expected to evolve dependent on the temporal stability of a cell's tissue context. Specifically, migratory cells, such as leukocytes, are exposed to a wide range of contexts over time, and therefore to a wide range of ligands. One would expect these cells to reduce the expression of receptors, limiting them to those involved in the cell type specific S-function and thus respond to only the "relevant" signals. Migratory cells are thus also not expected to engage in tissue maintenance as are tissue resident stromal cells. Exposure of the cell to variable environments may thus drive a reduction of the cell's receptor inventory to function-relevant signals. In contrast, the tissue-resident cells encounter a less variable signal environment and may therefore be able to express a less fine-tuned spectrum of receptors, overall, including some for which the ligand is simply not present, or that would be too promiscuous in a more stringent context. In addition, the role of these interactions may be more focused on tissue homeostasis and being more redundant rather than performing specific service functions for the organism. Redundancy among these interactions would also predict that these interactions are readily subject to developmental systems drift (33, 34), i.e., differences between species that do not affect the function and identity of the tissue.

Conclusion

The elementary considerations summarized above suggest a number of recommendations for the analysis of putative cell interactions networks deduced from single cell transcriptomic data:

- The analysis of putative interactions should be stratified by their molecular biology (paracrine, juxtacrine, ECM mediated, etc.). Different molecular types of interactions reflect different kinds of biology and shall not be lumped into summary statistics.
- Trying to distinguish between subnetworks for service (S-) and integrity (I-) functions might help focus attention to relevant parts of the network. For instance, degenerative diseases are likely due to failure of I-functions, while organ function diseases (renal failure) could be due to both, failures of I functions as well as S-functions, because I functions are necessary for performing S-functions.
- Scenarios for the evolution of cell interaction networks should respect the double contingency condition, i.e., the unlikely de-novo origination of the expression of a complete ligand-receptor pair in appropriate cell types. Scenarios based on single step deletion or expression of ligands or receptors are more likely, as for instance the pruning of an autocrine network of interactions into allocrine interactions, or duplication and divergence of ligand or receptor, or evolution of receptors for existing molecules not previously acting as ligands (ligand exploitation model; 35).

- Evolutionary changes in the spatial relationship between tissues can lead to novel interaction patterns, as is the case of the evolution of the placenta, where maternal and fetal tissues come into contact.
- The receptor inventory of a cell might reflect the stability of the cell's environment with more migratory cell being more restrictive in receptor expression than tissue resident cells.
- Finally, the complementarity of ligands and receptors in dataset may serve as a hint about data quality. If there is a strongly expressed ligand that has no local receptor, it is either a hormone, or the expression of the cognate receptor might have been missed.

References

1. Jin S, Guerrero-Juarez CF, Zhang L, Chang I, Ramos R, Kuan CH, et al. Inference and analysis of cell-cell communication using CellChat. Nat Commun. 2021;12(1):1088.
2. Hu Y, Peng T, Gao L, Tan K. CytoTalk: De novo construction of signal transduction networks using single-cell transcriptomic data. Sci Adv. 2021;7(16).
3. Efremova M, Vento-Tormo M, Teichmann SA, Vento-Tormo R. CellPhoneDB: inferring cell-cell communication from combined expression of multi-subunit ligand-receptor complexes. Nat Protoc. 2020;15(4):1484-506.
4. Browaeys R, Saelens W, Saeys Y. NicheNet: modeling intercellular communication by linking ligands to target genes. Nat Methods. 2020;17(2):159-62.
5. Armingol E, Baghdassarian HM, Martino C, Perez-Lopez A, Aamodt C, Knight R, et al. Context-aware deconvolution of cell-cell communication with Tensor-cell2cell. Nat Commun. 2022;13(1):3665.
6. Kauffman SA. Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol. 1969;22(3):437-67.
7. Vidal M, Cusick ME, Barabási AL. Interactome networks and human disease. Cell. 2011;144(6):986-98.
8. Vidal M. A unifying view of 21st century systems biology. FEBS Lett. 2009;583(24):3891-4.
9. Rolland T, Taşan M, Charloteaux B, Pevzner SJ, Zhong Q, Sahni N, et al. A proteome-scale map of the human interactome network. Cell. 2014;159(5):1212-26.
10. Goh KI, Cusick ME, Valle D, Childs B, Vidal M, Barabási AL. The human disease network. Proc Natl Acad Sci U S A. 2007;104(21):8685-90.
11. Ghiassian SD, Menche J, Chasman DI, Giulianini F, Wang R, Ricchiuto P, et al. Endophenotype Network Models: Common Core of Complex Diseases. Sci Rep. 2016;6:27414.
12. Davidson EH. Evolutionary bioscience as regulatory systems biology. Dev Biol. 2011;357(1):35-40.
13. Bridgham JT, Eick GN, Larroux C, Deshpande K, Harms MJ, Gauthier ME, et al. Protein evolution by molecular tinkering: diversification of the nuclear receptor superfamily from a ligand-dependent ancestor. PLoS Biol. 2010;8(10).
14. Markov GV, Laudet V. Origin and evolution of the ligand-binding ability of nuclear receptors. Mol Cell Endocrinol. 2011;334(1-2):21-30.
15. Pavlicev M, Wagner GP, Chavan AR, Owens K, Maziarz J, Dunn-Fletcher C, et al. Single-cell transcriptomics of the human placenta: inferring the cell communication network of the maternal-fetal interface. Genome Res. 2017;27(3):349-61.
16. Achim K, Eling N, Vergara HM, Bertucci PY, Musser J, Vopalensky P, et al. Whole-Body Single-Cell Sequencing Reveals Transcriptional Domains in the Annelid Larval Body. Mol Biol Evol. 2018;35(5):1047-62.
17. Musser JM, Schippers KJ, Nickel M, Mizzon G, Kohn AB, Pape C, et al. Profiling cellular diversity in sponges informs animal cell type and nervous system evolution. Science. 2021;374(6568):717-23.
18. Raredon MSB, Adams TS, Suhail Y, Schupp JC, Poli S, Neumark N, et al. Single-cell connectomic analysis of adult mammalian lungs. Sci Adv. 2019;5(12):eaaw3851.
19. Yue B. Biology of the extracellular matrix: an overview. J Glaucoma. 2014;23(8 Suppl 1):S20-3.
20. DiFrisco J, Love A, Wagner GP. Character identity mechanisms: a conceptual model for comparative-mechanistic biology. Biology & Philosophy. 2020(35):1-32.
21. Zhou X, Franklin RA, Adler M, Jacox JB, Bailis W, Shyer JA, et al. Circuit Design Features of a Stable Two-Cell System. Cell. 2018;172(4):744-57.e17.
22. Whyte LL. Internal factors of evolution. Acta Biotheoretica. 1963;16:33-48.

23. Wagner GP, Schwenk K. Evolutionarily Stable   Configurations: Functional Integration and the Evolution of Phenotypic Stability. Evolutionary Biology. 2000;31:155-217.
24. Schwenk K, Wagner PG. Function and the Evolution of Phenotypic Stability: Connecting Pattern to Process. American Zoologist. 2001;41:552-63.
25. Arthur W. Developmental drive: an important determinant of the direction of phenotypic evolution. Evol Dev. 2001;3(4):271-8.
26. Arthur W. The origin of animal body plans: a study in evolutionary developmental biology   . . Cambridge: Cambridge University Press; 1997.
27. Valentine JW, Collins AG, Meyer CP. Morphological complexity increase in metazoans.1994.
28. Vickaryous MK, Hall BK. Human cell type diversity, evolution, development, and classification with special reference to cells derived from the neural crest. Biol Rev Camb Philos Soc. 2006;81(3):425-55.
29. Arendt D, Musser JM, Baker CVH, Bergman A, Cepko C, Erwin DH, et al. The origin and evolution of cell types. Nat Rev Genet. 2016;17(12):744-57.
30. Wagner GP, Erkenbrack EM, Love AC. Stress-Induced Evolutionary Innovation: A Mechanism for the Origin of Cell Types. Bioessays. 2019;41(4):e1800188.
31. Sternberg PW, Horvitz HR. The genetic control of cell lineage during nematode development. Annu Rev Genet. 1984;18:489-524.
32. Brylski P, Hall BK. ONTOGENY OF A MACROEVOLUTIONARY PHENOTYPE: THE EXTERNAL CHEEK POUCHES OF GEOMYOID RODENTS. Evolution. 1988;42(2):391-5.
33. True JR, Haag ES. Developmental system drift and flexibility in evolutionary trajectories. Evol Dev. 2001;3(2):109-19.
34. Weiss KM, Fullerton SM. Phenogenetic drift and the evolution of genotype-phenotype relationships. Theor Popul Biol. 2000;57(3):187-95.
35. Thornton JW. Evolution of vertebrate steroid receptors from an ancestral estrogen receptor by ligand exploitation and serial genome expansions. Proc Natl Acad Sci U S A. 2001;98(10):5671-6.