The proximate–ultimate distinction and evolutionary developmental biology: causal irrelevance versus explanatory abstraction

Raphael Scholl · Massimo Pigliucci

Abstract Mayr's proximate–ultimate distinction has received renewed interest in recent years. Here we discuss its role in arguments about the relevance of developmental to evolutionary biology. We show that two recent critiques of the proximate–ultimate distinction fail to explain why developmental processes in particular should be of interest to evolutionary biologists. We trace these failures to a common problem: both critiques take the proximate–ultimate distinction to neglect specific causal interactions in nature. We argue that this is implausible, and that the distinction should instead be understood in the context of explanatory abstractions in complete causal models of evolutionary change. Once the debate is reframed in this way, the proximate–ultimate distinction's role in arguments against the theoretical significance of evo-devo is seen to rely on a generally implicit premise: that the variation produced by development is abundant, small and undirected. We show that a "lean version" of the proximate–ultimate distinction can be maintained even when this isotropy assumption does not hold. Finally, we connect these considerations to biological practice. We show that the investigation of developmental constraints in evolutionary transitions has long relied on a methodology which foregrounds the explanatory role of developmental processes. It is, however, entirely compatible with the lean version of the proximate–ultimate distinction.

Keywords Proximate–ultimate distinction · Ernst Mayr · Evolutionary developmental biology · Niche construction · Plasticity · Abstraction · Methodology · Pere Alberch

R. Scholl (✉)
History and Philosophy of Science, Institute of Philosophy, University of Bern, Sidlerstr. 5, CH-3012 Bern, Switzerland
e-mail: raphael.scholl@gmail.com

M. Pigliucci
Philosophy Program, The Graduate Center, City University of New York, New York City, NY, USA
e-mail: massimo.pigliucci@icloud.com

🖄 Springer

Introduction

Ernst Mayr's proximate–ultimate-distinction (Mayr 1961) has recently received renewed interest in the literature, especially in this journal. Biologists and philosophers of biology have been debating how the distinction (here abbreviated as PUD) should be understood, what conceptual work it does, and how it can illuminate debates about issues such as niche construction and evolutionary developmental biology (Laland et al. 2011, 2012; Haig 2013; Dickins and Barton 2012; Calcott 2013; Gardner 2013). In the context of evolutionary developmental biology, the PUD has long been especially controversial because of its use—by Mayr himself—as an argument against the relevance of development to evolution (Mayr 1984; Amundson 2005, especially chapters 10 and 11).

In the present paper, we argue that in order to fully articulate the PUD and to handle its problems, it is necessary to distinguish between causal irrelevance and explanatory abstraction. We will begin by introducing the PUD and some recent criticisms of it ("The proximate–ultimate distinction and its discontents" section). Where previous authors have claimed that the PUD is limited because it omits certain types of biological causation, we argue that a reframing of the debate is required ("Towards a better conception of the proximate–ultimate distinction: abstraction in complete causal models of evolutionary change" section). The PUD should be understood not in terms of the omission of causes, but in terms of appropriate *explanatory abstractions* in complete causal models of evolutionary change. We will show that this perspective allows us both to recover a useful "lean version" of the PUD and to embrace recent critiques. Finally, we will show how our account connects to actual scientific practice by discussing two studies of developmental constraints in evolutionary processes ("Case studies: Digital reduction in amphibians and the origin of pigment patterns in Drosophila" section): one from recent, molecular evo-devo and one from an earlier, morphological approach.

The proximate–ultimate distinction and its discontents

The PUD and evolutionary developmental biology

Mayr (1961) argued that biology, despite its seeming unity, is made up of at least two fields that differ in their choice of, and approach to, research problems. On the one hand, Mayr recognizes a discipline of *functional* biology, of which he writes:

> The functional biologist is vitally concerned with the operation and interaction of structural elements, from molecules up to organs and whole individuals. His ever-repeated question is 'How?' (p. 1502)

He contrasts this with the interests of the *evolutionary* biologist:

His basic question is 'Why?'. […] To find the causes for the existing characteristics, and particularly adaptations, of organisms is the main preoccupation of the evolutionary biologist. (p. 1502)

Mayr could be read as distinguishing between "how" and "why" questions; between causes acting in the past and causes acting in the present; between different *types* of causes; or between scientific disciplines.

The distinction between past and present causes appears to be favored by several authors (Amundson 2005; Haig 2013; Hochman 2012). However, it seems to us that Mayr's original text does not support this interpretation. Mayr begins by recognizing two distinct subdisciplines: Functional biology is concerned with "the operation and interaction of structural elements" (what we might call biological mechanisms in the sense of Machamer et al. 2000), whereas evolutionary biology studies "*the causes* for the existing characteristics", "*the reasons* for this diversity [of the organic world] as well as the pathway by which it has been achieved" or "*the forces* that bring about changes in faunas and floras" (all of these quotes appear in quick succession on p. 1502, the emphases are ours). Mayr then assigns the label "proximate" to the former and "ultimate" to the latter. He specifies the ultimate causes further:

These are causes that have a history and that have been incorporated into the system through many thousands of generations of natural selection. (p. 1503)

Together with the references to "causes", "reasons" and "forces" of change or existing characteristics, this quotation indicates that Mayr is distinguishing not between past and present causes, but between different *types* of causes or processes. We can read ultimate causes as evolutionary processes, among which Mayr regards natural selection as the most important. Mayr's chosen illustration is the seasonal migration of birds (pp. 1502–1503), where we can ask either what *biological mechanisms* regulate bird migration or what *evolutionary processes* explains that these mechanisms exist.

This interpretation is also compatible with a further reformulation of the distinction to which Mayr kept returning throughout his career: An informational metaphor according to which functional biology deals with "all aspects of the decoding of the programmed information contained in the DNA code of the fertilized zygote" while evolutionary biology is interested in "the history of these codes of information and in the laws that control changes of these codes from generation to generation" (p. 1502). A schematic representation of the "informational version" of the PUD is given in Fig. 1.

Adopting this reading of Mayr's original paper gives typical answers to the other readings as well, and this explains in part why it has been difficult to assign an unambiguous original intention to Mayr. Evolutionary explanations of organismic traits *typically* answer questions that can be naturally phrased as "why"-questions (but of course we can ask: "*how* did the present mechanisms of bird migration evolve?"); they *typically* refer to selection acting in the past (although selection does, of course, act in the present); and they are *typically* investigated by evolutionary biology.
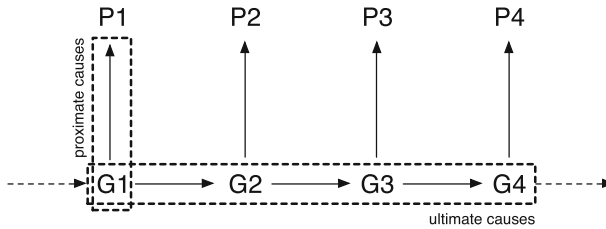
Fig. 1 A schematic view of the "informational" version of Mayr's proximate–ultimate-distinction. *G1–G4* indicate population gene pools at time intervals 1–4 (arbitrary units). *P1–P4* indicate population phenotypes at the same time intervals. Mayr's proximate causes (*vertical*) mediate between genotypes and phenotypes; in contrast, ultimate causes (*horizontal*) are responsible for change from one population gene pool to the next over time

In the context of 1961, Mayr used to PUD to argue for the continued relevance of organismic biology in the face of molecular approaches (Beatty 1994). It is only in the 1980s that Mayr began to deploy the distinction against evolutionary developmental biology:

> The suggestion that it is the task of the Darwinians to explain development […] makes it evident that Ho and Saunders [critics of the Modern Synthesis] are unaware of the important difference between proximate and ultimate causations […]. Expressed in modern terminology, ultimate causations (largely natural selection) are those involved in the assembling of new genetic programmes, and proximate causations those that deal with the decoding of the genetic programme during ontogeny and subsequent life. (Mayr 1984, p. 1262)

The PUD is claimed to give rise to an explanatory asymmetry: evolutionary causes (in the horizontal arrows) explain how the genetic programs underlying functional causes (in the vertical arrows) have changed across generations. But there is no comparable reverse relationship—functional causes are of no relevance to evolutionary causes. This view is echoed by the geneticist Bruce Wallace (a former student of Thedosius Dobzhansky), who wrote forcefully:

> Evolutionary geneticists have the responsibility for explaining the origins and subsequent fates of the genetic programs which determine developmental programs; embryologists, on the contrary, need not explain how somatic development might affect the evolution of these developmental programs. Except for achieving success in reproduction, they do not. (Wallace 1986, p. 150)

The PUD thus provides the basis for an argument for the irrelevance of development to evolutionary explanations.

The PUD's critics

There exists an ongoing debate about whether a distinction between proximate and ultimate causes should be drawn, and if so, how to draw it. Ariew (2003) has argued

that the proper distinction is one between individual-level causal and population-level statistical explanations. Haig (2013) has argued that the distinctions between past and present causes and between mechanisms and adaptive rationale get conflated, and that the PUD should be abandoned in favor of a distinction between "how come?" and "what for?" questions. Similarly, Gardner (2013) holds that the proper "ultimate" question concerns the adaptive rationale of organism design. Calcott (2013) has suggested the more pluralistic view that Mayr's PUD captures only one contrast among many which can be used to ask interesting biological questions.

In the context of evolutionary developmental biology, the most salient critiques are by Amundson (2005) and Laland et al. (2011, 2012).

Since natural selection invariably operates on developmental processes, Amundson (2005) argues that any account of an evolutionary transition is *causally incomplete* so long as it does not include the relevant developmental processes. Following Horder, he proposes a "causal completeness principle" or CCP:

> In order to achieve a modification in adult form, evolution must modify the embryological processes responsible for that form. Therefore an understanding of evolution requires an understanding of development. (p. 176)

Gene pools of one generation only become gene pools of the next generation via phenotypes. The PUD implies a historically continuous lineage of gene pools with phenotypic offshoots (as shown in Fig. 1) when, in fact, gene pools and phenotypes must be seen as occurring in series. On this view, causes cannot be either proximate or ultimate; they can only be more or less proximate and more or less ultimate (and this holds both on the "past/present" and the "types of causes" readings of the PUD). Amundson's CCP critique is also accepted by Alan Love (2006, p. 322), who argues that a version of the CCP constitutes part of the evo-devo synthesis's criteria for explanatory adequacy. Figure 2 summarizes the "causal completeness" critique.

Problematically, however, Amundson's critique fails to offer a reason why developmental processes in particular should be of interest to evolutionary biology. On the criterion of causal completeness, we might also ask for a biochemical-evolutionary synthesis (since biochemistry is both continuously operating in organisms and changed by selection) or indeed a physiological-evolutionary synthesis. Some might argue that these are indeed needed, but we think it is possible to give good reasons why development *in particular* is of interest to evolutionists.

Moreover, Amundson's critique is implausibly uncharitable towards Mayr's views. Mayr argued vehemently that selection acts on phenotypes and not on genotypes (Mayr 1997). That development was thus part of the complete causal story of any evolutionary explanation could not plausibly have been doubted by him or his contemporaries. Even in Mayr's chosen illustration of the PUD—bird migration—the role of development in a causally complete account is obvious. Thus, we find it implausible to think that Mayr did not believe development to be part of the complete causal story. It is more promising to argue with Love (2006, cited above) that Mayr adhered to different criteria of explanatory adequacy which did not include causal completeness. But this also seems to us an extraordinary assumption which would require extraordinary documentation.
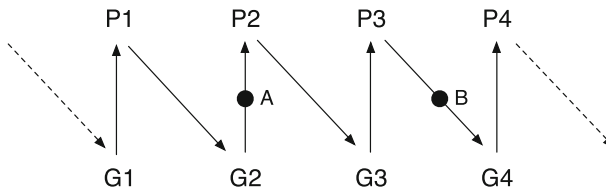
Fig. 2 The "causal completeness principle" after Amundson (2005) as a critique of the claim that the proximate–ultimate distinction justifies the exclusion of developmental causes from evolutionary explanations. The CCP stresses that genotypes of one generation (such as *G1*) become genotypes of the next generation (*G2*) only via phenotypes (*P1*). Therefore, evolutionary causes are not more "ultimate" than developmental causes on a "past/present" reading of the proximate–ultimate distinction: a cause occurring in the ontogeny of the second generation (*G2–P2, A*) is more "ultimate" in a temporal sense than one of Mayr's ultimate causes (e.g., selection) acting on the population between generations 3 and 4 (*P3–G4, B*). Moreover, the phenotypic variation on which selection acts is produced by development, and so development cannot be exclusively assigned to the "proximate" domain on a "types of causes" reading of the proximate–ultimate distinction. See text for details

More promising at first sight is the critique by Laland et al. (2011, 2012) who argue that the PUD ignores certain prevalent types of reciprocal causation:

> In reciprocal processes, ultimate explanations must include an account of the sources of selection (as these are modified by the evolutionary process) as well as the causes of the phenotypes subject to selection. (2011, p. 1512)

The authors' paradigmatic case is niche construction (Odling-Smee et al. 2003), where selection creates organisms which alter their environment and thus, in turn, alter selection pressures. Laland et al. regard Mayr's view as one of linear causation, where a series of genotypes are successively adapted to static selection pressures. In niche construction, however, there is a phenotype-mediated alteration of the selective environment, and thus reciprocal causation.

According to Laland et al., the same line of argument shows why the PUD is inapplicable in the case of evo-devo. They use the example of phenotypic plasticity permitting an adaptive response which is then only secondarily genetically stabilized (West-Eberhard 2003; Minelli 2003). The idea here is that plastic phenotypes can "bridge" evolutionary transitions by permitting phenotypic adaptations before any genetic adaptations take place. Hence, the plastic phenotype effectively changes the selection pressures and makes it possible for phenotype-stabilizing genes to be selected. In the absence of the phenotypic adaptation, these same genes might never have been selectively advantageous. Again, then, we have the reciprocity of causation which Laland et al. diagnose as lacking in the classical conception of the PUD: selection creates phenotype, phenotype changes selection pressures.

However, the criterion of reciprocal causation fails to pick out all the cases of interest, and only those. On the one hand, reciprocal causation is again too inclusive since it is ever present. Most phenotypes will be plastic to at least some degree, and so if we take the criterion of reciprocal causation seriously, we would conclude that the PUD never applies. On the other hand, reciprocal causation is not sufficiently inclusive: There are phenomena that are of great theoretical interest to

developmental evolutionists but that are not neatly captured as instances of reciprocal causation. Take for example *developmental drive* (Arthur 2001), where the availability of developmental variants co-determines the direction of evolution. It is not at all obvious how this is a case of reciprocal causation in the sense of Laland et al., but it is definitely a phenomenon of theoretical interest to evo-devo, at least potentially necessitating the inclusion of developmental causes in evolutionary explanations. On the criterion of reciprocal causation, however, cases of developmental drive would slip through our net.

Towards a better conception of the proximate–ultimate distinction: abstraction in complete causal models of evolutionary change

Our thesis, which we will defend below, is that it is a mistake to think about the PUD in terms of allegedly omitted biological causes. It is much more fruitful to think in terms of abstraction in causal explanations: The issue is not whether certain types of causation (e.g., between genotype and phenotype, or between phenotype and selective environment) *exist*, but whether these causal paths *carry much weight* in the explanations we give. We should not ask: Is there causation between genotype and phenotype? Of course there is. Or: Is there causation between phenotype and environment? Again, of course there is. We should ask instead what motivates the foregrounding or backgrounding of some parts of a complete causal account of any given evolutionary transition. The key concern is not causal relevance but explanatory salience.

Two types of abstractions from a causally complete account are possible in principle. First, we may de-emphasize "horizontal" factors, that is, parts of a cluster of causes which are jointly sufficient for bringing about an effect. Thus, when we explain why a shed burned down, we will place more emphasis on the arsonist's match than on the presence of oxygen in the air, even though both factors are part of a causally complete account. Second, we may de-emphasize "vertical" factors, that is, parts of a causal chain leading from cause to effect. Consider the explanation of scurvy as the effect of vitamin C deficiency: Depending on circumstances, the explanations we give will emphasize that vitamin C deficiency causes scurvy while neglecting a precise account of the intervening mechanism by which vitamin C is involved in collagen synthesis. Neither horizontal nor vertical abstraction is intended as a statement about causal irrelevance: only explanatory insignificance is implied.

Both horizontal and vertical abstraction can be motivated by objective or pragmatic reasons. That the arsonist's match has more explanatory salience than oxygen is partly due to the fact that oxygen is present around all our sheds, and so only the arsonist's match answers the contrastive question of why one shed rather than another burned down. In a legal context, by contrast, we may place pragmatic emphasis on the arsonist's match because it involves human agency.

How does the perspective of abstractions in complete causal models of evolutionary change help in the debate about the proximate–ultimate distinction? We have seen that Mayr's "informational" version of the PUD has been taken to

imply an explanatory asymmetry: evolutionary processes explain developmental mechanisms, but developmental mechanisms do not explain evolutionary processes. However, In the context of the Amundson critique, we have also seen that Mayr cannot plausibly have meant to imply that developmental *causes* are absent in evolutionary processes—after all, Mayr insisted that the phenotype and not the genotype is the object of selection. Thus, simply to ask for causal completeness would be to talk at cross-purposes with Mayr.

The solution is to recognize that Mayr considered natural selection to be the main force in evolutionary *explanations*. Thus, he foregrounded selection and abstracted development, although both are uncontroversially part of a complete causal account of evolutionary transitions. This foregrounding of selection is in line with Mayr's known theoretical commitments and the general "hardening" of the modern synthesis at the time (Gould 1983). The foregrounding of selection is also made explicit in Mayr's (1984) use of the PUD against evolutionary developmental biology, where he equates ultimate causes with natural selection clearly by speaking of "ultimate causations (largely natural selection)" (Mayr 1984, p. 1262).

The key question, then, is this: Under what circumstances can selection be taken to be the main explanatory force in evolutionary explanations? The conceptual answer to this is as old as the *Origin of Species* (1859), where Darwin writes:

> If selection consisted merely in separating some very distinct variety, and breeding from it, the principle would be so obvious as hardly to be worth notice; but its importance consists in the great effect produced by the accumulation in one direction, during successive generations, of differences absolutely inappreciable by an uneducated eye […]. (p. 32)

Darwin recognizes that the explanatory force of natural selection depends on the character of the variation upon which it acts. Gould (2002, pp. 144–146) made this more precise, noting that selection is maximally powerful if it acts on abundant, small and undirected variation. He called this "isotropic" variation. The underlying insight is that if variation is "fine grained" in the sense discussed, then selection explains every step in the cumulative change of the system. Even though the variation still has developmental *causes*, these *do not explain much* since only those variations that confer a fitness advantage will survive the selective sieve. The more "non-isotropic" or coarse grained variation is, by contrast, the more will we need to ask which developmental processes impart structure to it.

The idea that development carries no force in evolutionary explanations rests on the assumption of variational isotropy. In arguing for an explanatory asymmetry between development and evolution, Mayr, Wallace and others implicitly commit to this assumption.

Thus, two issues have become conflated. The first issue is the PUD, which does not necessarily imply an explanatory asymmetry between evolution and development. We think a "lean version" of the PUD should be maintained, since it separates research agendas (Mayr's *Fragestellungen*) that are indeed concerned with different aspects of causation in nature: Proximate questions about biological mechanisms are different from ultimate questions about the evolutionary processes that have produced these same mechanisms. Proximate and ultimate causes answer

different contrastive questions. The proximate questions asks why *this bird* flies south in contrast to another, otherwise identical bird that lacks the same neural mechanism. In contrast, the ultimate question asks why *these birds* fly south in contrast to another population of birds with a different history of natural selection (Knell and Weber 2009, p. 92ff; Ariew 2003).[1]

The second issue is the claim that developmental mechanisms carry no force in evolutionary explanations. This is only true in cases where variation is isotropic sensu Gould, and this empirical assumption is independent of the lean version of the PUD. On the assumption of isotropic variation, developmental mechanisms are indeed nothing but gory mechanistic detail in evolutionary explanations. When variation is not isotropic, however, developmental causes do carry explanatory force in evolutionary transitions—and our lean version of the PUD is entirely compatible with this possibility. The task of proponents of the evo-devo approach has always been to argue not against the PUD, but against the generality of the isotropy assumption—to argue that some evolutionary explanations require us to foreground developmental mechanisms.[2] It is not a new idea that a shift of explanatory focus is key to understanding the evo-devo synthesis: Günter Wagner has argued along these lines (Wagner 2000, 2001; Wagner and Larsson 2003), as has Ingo Brigandt (2010, p. 302).

With this conceptual background in place, we can now return to the criticisms and examples discussed so far. Consider Laland et al.'s example of developmental plasticity allowing evolutionary transitions. West-Eberhard's instructive example of this is a two-legged goat (West-Eberhard 2005). The goat was born with paralyzed front legs and had to use its hind legs to hop around. After its death, a dissection revealed a number of phenotypic accommodations to two-legged gait—accommodations which in evolutionary studies we might be tempted to assume to be produced by the cumulation of small, individually advantageous variations. In light of the foregoing discussion, however, we can recognize this clearly as a case where variation is non-isotropic: instead of abundant, small and undirected, the phenotypic accommodation is large and adaptive. Had this been the first stage of an evolutionary transition toward two-legged gait in goats (or in any other vertebrate), an adequate evolutionary explanation would have needed to foreground the developmental processes which permitted this accommodation.

Now consider Laland et al.'s example of niche construction. It is certainly not the case that in niche construction organisms causally interact with their environment whereas in other cases of evolution they do not. Reciprocity of causation between selection and the environment always exists. Rather, the argument is that in some cases the organism will change its environment such that its further evolutionary

---

[1] It has recently been suggested by several authors that ultimate questions should be restricted to issues of adaptive rationale or "what for?" (Haig 2013; Gardner 2013). To some extent, this is a dispute about preferred linguistic usage. However, in the spirit of Gould and Lewontin (1979) we prefer to include all evolutionary processes in the ultimate category. This includes selection, drift, developmental constraints, and more. Which of these processes are the most relevant should then be investigated as an empirical matter on a case by case basis.

[2] Similar arguments could be made for other processes that have an effect on variational isotropy, such as variation in mutation rates (Hodgkinson and Eyre-Walker 2011).

trajectory is markedly different from what it would have been if the environment had stayed static. In these cases, then, the impact of the organism on the environment has explanatory force, which motivates the foregrounding of that part of the complete causal story. However, in many cases of niche construction it may be perfectly appropriate to background *developmental* causes, if the variation available to selection is isotropic.

Finally, consider developmental drive, where the availability of variations co-determines the trajectory of evolutionary change. On the criterion of reciprocal causation, we would not recognize that such cases are of particular interest to developmental evolutionists. However, the difficulty disappears if we recognize the role of the isotropy assumption. In cases of developmental drive, variation is clearly not isotropic, and thus the developmental mechanisms responsible for the creation of variation need to be foregrounded if we are to understand the trajectory of evolution. Meanwhile, however, cases of developmental drive will generally allow us to background the interaction of the phenotype with the environment, unless significant niche construction happens in the same evolutionary episode.

In summary, we argue that the critiques of "causal completeness" and "reciprocal causation" fail because they do not reliably pick out those aspects of individual evolutionary episodes which are of interest to developmental evolutionists or experts in niche construction. Instead, the proper framing of the question concerns explanatory abstractions in causal models: Under what circumstances is it appropriate for explanatory purposes to foreground or background certain aspects of the complete causal account of any given evolutionary transition? Ernst Mayr abstracted from developmental causes because he assumed isotropy of variation. In cases where this assumption is not warranted, however, developmental causes carry explanatory force and must be included. In other cases, interactions between the phenotype and the environment may carry particular explanatory force. This approach permits us a charitable reading of Mayr, whose focus on development-less natural selection can be seen as a special case where variation is isotropic. In the half-century since Mayr's discussion of the proximate–ultimate distinction, it has become apparent that the isotropy assumption does not hold generally. Thus, developmental causes can carry explanatory force in both the proximate and the ultimate domain. Nevertheless, we have argued that we should maintain a lean version of the proximate–ultimate distinction: between different types of contrastive questions.

Case studies: digital reduction in amphibians and the origin of pigment patterns in Drosophila

In the preceding section, we accepted a lean version of the PUD while suggesting a framework in which we can ask which causal factors carry particular explanatory weight in different evolutionary transitions. We take it as given that development, plasticity and niche construction are causally relevant to a greater or lesser degree in practically all evolutionary transitions. However, when any of these processes carry particular explanatory force, they need to be foregrounded in evolutionary

explanations. Otherwise, they can be safely abstracted. This allows us to speak of some evolutionary transitions as if development or plasticity did not play a role, because variation is isotropic; or as if niche construction did not play a role, because the environment remains stable in relevant respects.

We will now connect this philosophical approach with actual scientific practice by considering two case studies where scientists argue by foregrounding parts of a complete causal account of evolutionary transitions. Both cases concern developmental constraints in evolutionary transitions, one of the classical theoretical concerns of evo-devo (Maynard-Smith et al. 1985).[3]

In the philosophical literature, the opinion is prevalent that the investigation of constraints or drive requires optimality analyses of the trait under selection. Elliott Sober in his "Six Sayings About Adaptationism" (Sober 1998), for example, discusses the following saying:

> Adaptationist thinking is an indispensable research tool. The only way to find out whether an organism is imperfectly adapted is to describe what it would be like if it were perfectly adapted. (p. 83)

Sober adds that this "is exactly right". However, we will argue that Sober's view does not do justice to the scientific practice of developmental evolutionists. In order to show that a trait was not formed by the cumulation of isotropic variation (abundant, small and undirected), it is often sufficient to show that the fine structure of the actual trait reflects developmental mechanisms. In other words, the developmental processes of the complete causal model of an evolutionary transition get foregrounded to show their role in determining actual structures. This does not, of course, exclude the possibility that one of the limited number of variants available to selection *just so happens* also to be functionally optimal *as if* it had been cumulatively selected from isotropic variation. However, it would then be up to the adaptationist to show that this coincidence has actually occurred, and so the burden of proof is shifted (Pigliucci and Boudry 2013).

Our first example is from work carried out by Pere Alberch, who wrote on both conceptual and methodological aspects of the problem of integrating evolution and development in the 1980s (theory in Alberch 1982, empirical results in Alberch and Gale 1985). These studies have become part of textbook canon within the field of evolutionary biology (Futuyma 1998, p. 672). They have also received attention from philosophers of evo-devo (Amundson 1994), but there has been a dearth of consideration of the methodology by which Alberch argues for developmental constraints.

---

[3] Gould (1989) argued valiantly that "constraint" should be understood not only in the negative sense of limiting the power of selection, but also in the positive sense of channeling evolutionary transitions in particular adaptive directions. It seems that this usage has not been widely adopted; the more recent term "developmental drive" (Arthur 2001) appears to have been received more favorably for the positive case. Both constraints and drive, however, have the same conceptual foundation: If the variation available to selection is strongly structured by developmental mechanisms, then evolutionary transitions will need to be explained by asking both which variants were selectively advantageous and which variants were produced in the first place.

Alberch's original papers reveal that he clearly appreciated the contrast between causal completeness and explanatory adequacy, even though he did not use these exact terms. Alberch and Gale begin their major 1985 paper by pointing out that the causal role of development in evolutionary processes is generally undisputed:

> That development, as the link between the genotype (level of variation) and the phenotype (level of selection), plays a role in structuring evolutionary patterns is not a point of contention. Instead, the basic issue is to show what additional insight could be gained by incorporating development into the current evolutionary scheme. (p. 8)

The language employed in this passage is somewhat infelicitous. Writing of development's "role in structuring evolutionary patterns" makes it appear as though development had an accepted role to play in actually directing the course of evolution (perhaps along with natural selection). The context indicates, however, that the uncontested fact expressed by the phrase "the link between the genotype … and the phenotype" is merely meant to be that development causally connects the genotype (which varies) and the phenotype (which is selected upon). This interpretation of the passage is in line with Alberch's earlier conceptual work (Alberch 1982).

To rephrase, Alberch and Gale point out that development is uncontroversially among the causal links between genotype and phenotype, but that it is not clear whether development has any role to play in setting the path of evolutionary change, and if it does, how to demonstrate such a role. Alberch and Gale regard their paper mainly as a methodological contribution showing how this problem could be approached. To put it in our terminology, they set out to construct a methodology for demonstrating that developmental factors need to be foregrounded in specific explanations of evolutionary change.
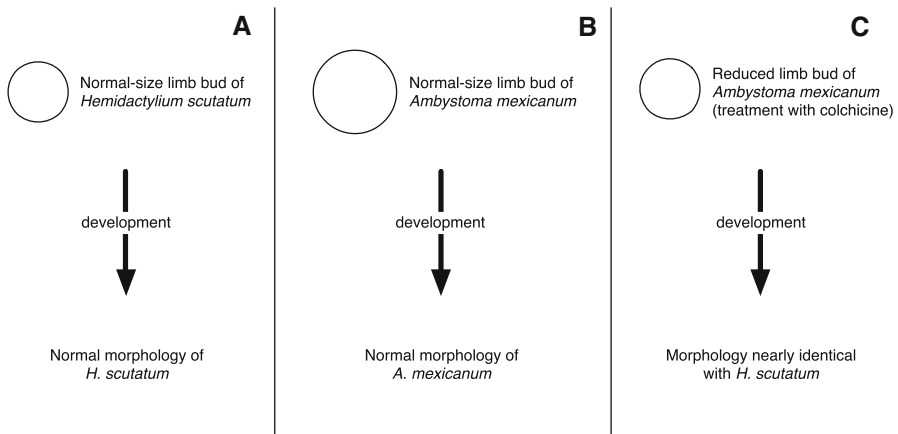
The empirical study presented by Alberch and Gale consists of comparisons within and among two orders of amphibians: plethodontid salamanders and anuran frogs. The trait of interest is loss of phalanges in the extremities. In short, the authors are able to show that the pattern of phalangeal loss is similar in species belonging to the same order, but different among orders. They further show that different morphologies within the same order can be reproduced experimentally by varying a single developmental parameter: size (= number of cells) in the limb bud, which can be manipulated by the (reversible) application of colchicine to the developing limb bud. For example, treatment of the limb bud of the salamander *Ambystoma mexicanum* with colchicine results in the loss of various phalanges in such a way as to mimic the (normal) morphology of the related species *Hemidactylium scutatum* (see Fig. 3).

Alberch presented a similar argument in another publication that same year, in which he asked why St. Bernards often have an extra digit on their hind limb (polydactyly) and poodles never do (1985). In the past, selective hypotheses had been proposed, such as the notion that an extra digit might impart some type of locomotory benefit in deep snow, where St. Bernards iconically bring Cognac to avalanche victims. Alberch found this and other such functional hypotheses implausible. He proposed, instead, that limb bud size is again the relevant factor,

and that large dogs, having large limb buds, will sometimes find themselves above a threshold limb bud size which results in the production of an additional digit, while poodles simply never approach that threshold (see Fig. 4).

Both examples illustrate a general methodology for demonstrating the need to foreground developmental processes in evolutionary models. What is *not* produced is evidence related to optimality—either evidence that a given feature is non-optimal or that some other structure would be optimal. What *is* produced is evidence that the feature under investigation (adult limb morphology) is a function of a developmental factor (limb bud size), and that the existing adult limb morphology can thus be interpreted as a side-effect of selection on body size alone (see Fig. 5). Such causal claims can be supported by experimental data which demonstrate the relevant developmental relationships. Any additional explanations of the foot morphology by natural selection are not so much "disproved" as made redundant.

Alberch is clearly aware of the fact that he is interested in developmental causes not simply because they occur between genotype and phenotype (causal complete-ness), but because in this case they influence the direction of evolutionary transitions. In Alberch (1985) he makes this explicit by arguing that some "intermediate causes" between genotype and phenotype are explanatorily relevant and therefore must be included in an adequate causal explanation, while others can be safely abstracted, since they add (in our terminology) only to causal completeness. Alberch writes:

A   Normal-size limb bud of *Hemidactylium scutatum*

development

Normal morphology of *H. scutatum*

B   Normal-size limb bud of *Ambystoma mexicanum*

development

Normal morphology of *A. mexicanum*

C   Reduced limb bud of *Ambystoma mexicanum* (treatment with colchicine)

development

Morphology nearly identical with *H. scutatum*

Fig. 3 Experiments performed by Alberch and Gale (1985) to support a developmental explanation of differences in foot morphology between the salamander species *H. scutatum* and *A. mexicanum*. Embroys were allowed to grow until their skeletal foot morphology could be determined by radiological techniques. *H. scutatum* and *A. mexicanum* developed clearly distinct morphologies when allowed to grow normally (**a**, **b**; note that *A. mexicanum* is the larger of the two species in terms of body size). When the size of the developing limb bud of *A. mexicanum* was reduced by treatment with the cellular toxin colchicine, which inhibits mitosis, *A. mexicanum* developed a skeletal morphology nearly identical to that of *H. scutatum* (**c**). The thereby established causal relationship supports the explanation of differences in limb morphologies as incidental consequences of differences in body size, rather than the cumulative product of natural selection acting on small heritable variations (the differences in body size themselves may have been produced by selection, of course; see also Fig. 5)
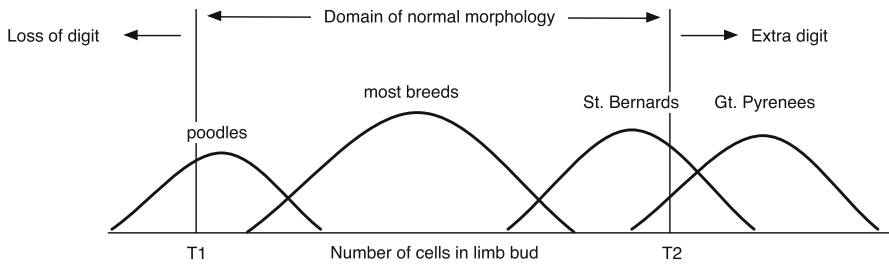
Fig. 4 Alberch's hypothetical explanation of the fact that large dog breeds sometimes have additional digits, while small breeds do not. It is assumed that the number of digits produced is a function of the size of the developing limb bud, with digits being gained or lost above or below certain threshold sizes (*T1* and *T2*). Only the size distributions of large breeds cross T2. If this causal relationship between body size (and indirectly limb size) and morphology can be demonstrated, then selectionist explanations of the appearance of additional digits in large breeds of dogs are redundant. This argument is identical to that presented for salamander foot morphology in Fig. 5, but it is not backed up by experimental causal inferences as shown in Fig. 3. Redrawn from Alberch (1985)

> Note that this discussion has not dealt with the genetic basis of the character. … It does not matter if the expression of the polydactylous morphology is controlled by an additive polygenic system […] or by a discrete Mendelian gene […]. The need to add a developmental component to the analysis is obvious, since a purely genetic study could not explain why mutations resulting in polydactyly appear only in large breeds and not in small ones. (p. 432)

We share Alberch's view that molecular detail is unlikely to add to the explanatory power of this particular analysis, since a description purely at the supra-cellular level already supplies all the causal factors and regularities that we need to understand the evolutionary trajectory in question. An analysis of molecular mechanisms might reveal why it is the case that limb buds, depending on their size, give rise to different specific morphologies. In other words, we would begin to understand why the relevant causal relationships between limb bud size and foot morphology *themselves* obtain (we would thus be decreasing vertical abstraction, or increasing mechanistic detail). Yet while this might be of interest to developmental biology proper, it is not at all clear how it would significantly deepen our insight into why evolution proceeded along the path that it did.

However, evo-devo is not necessarily more organismic than molecular. A second illustrative case, which exemplifies similar argument patterns, is from recent molecular work in Sean B. Carroll's laboratory. It concerns the evolution of wing pigmentation patterns in *Drosophila biarmipes* (Prud'homme, Gompel and Carroll 2007; Gompel et al. 2005).[4] Prud'homme et al. show that wing spots originated through the creation of novel *cis*-regulatory elements driving the expression (or repression) of the gene *yellow*. Importantly, the *trans*-acting factors that bind to this novel *cis*-regulatory sequence are expressed in the same pattern as in *D. biarmipes*

---

[4] We thank Günter Wagner for suggesting this example.
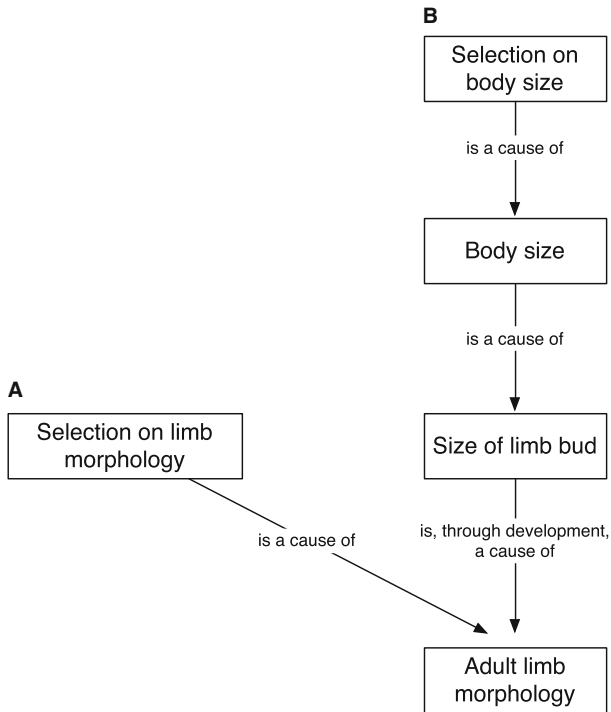
Fig. 5 A schematic illustration of the argument by which the experiments presented in figure 3 were taken to show the preferability of a developmental explanation to a selectionist interpretation. It would be difficult or impossible to demonstrate by experiment or modelling that differences in skeletal morphology cannot be due to natural selection acting under appropriate circumstances (the hypothetical causal pathway starting at *A*). However, it was possible to show by experiment that the size of the developing limb bud was an adequate causal explanation of the specifics of the skeletal morphology in the adult limb (the crucial part of the pathway starting at *B*). This positive demonstration of a causal relationship renders alternative, selectionist explanations superfluous

in the wings of unspotted relatives, and were therefore, presumably, also expressed in this pattern in the wings of unspotted ancestors. Thus, the distribution of *trans*-regulatory factors observed in *D. biarmipes* (what Gompel et al. call the "regulatory landscape" of the wing) existed before wing spots actually evolved.

The study by Gompel et al. strongly suggests that the pre-existing regulatory landscape was co-opted for the production of wing spots once an appropriate (environmental or sexual) selective pressure existed. Thus, in order to understand wing spot morphology, we need to understand both selection and development. We need to know what conditions favored the evolution of some type of wing spots in some species of the *Drosophila* genus, that is, we need to understand why such spots were advantageous. However, this will tell us little about the details of wing spot morphology, which appear not to have been minutely sculpted by selection for their functional role. In order to understand why the wing spots have their present, specific shape, we also need to understand something about the developmental machinery that was co-opted to produce them, that is, about the available

distributions of *trans*-acting regulatory factors. While selection without a doubt pushed *B. biarmipes* toward the evolution of wing spots, it was the pre-existing developmental machinery that largely determined their specific morphology.

Note that the logic of the argument in the study by Gompel et al. is the same as in the study by Alberch and Gale. The question of whether our explanation of the wing patterns under investigation should or should not include developmental constraints does *not* hinge on optimality models. The core argument is that purely developmental factors—in this case, the co-option of a pre-existing regulatory landscape—are sufficient to explain the specific morphology of the wing pigmentation. This claim can be demonstrated by causal inference, in this case molecular laboratory work. It is this demonstration of a causal relationship which renders the alternative, selectionist explanation (for the specific shape, not the existence of wing patterns per se) redundant. It demonstrates that, in this instance, an adequate explanation of evolutionary change needs to foreground developmental causes.

Conclusion

We have argued that two recent critiques of Mayr's proximate–ultimate distinction—based on the notions of causal completeness and reciprocal causation—are unsatisfying. A more fruitful approach to the problem is to think about abstractions in complete causal models of evolutionary change: Depending on which parts of the complete causal model carry the most explanatory force, different processes will be foregrounded or backgrounded.

This perspective allows us to charitably understand Mayr's dismissal of developmental causes as a statement about explanatory salience. Mayr assumed that selection acts on variation which is "isotropic" in Gould's terminology: abundant, small and undirected. On this assumption, developmental processes carry little explanatory force in evolutionary transitions, and this—rather than the distinction between ultimate and proximate causes per se—motivated the claim that developmental processes were irrelevant to evolutionary explanations. A half-century after Mayr's original discussion, we can now recognize isotropy of variation as a special case which is violated in instances of developmental constraint or drive, or in cases of developmental plasticity.

A lean proximate–ultimate distinction—between biological mechanisms and evolutionary processes—should be maintained because proximate and ultimate causes answer different contrastive questions. It is entirely compatible with the view that developmental causes carry explanatory force in some evolutionary explanations: namely, when the assumption of isotropy of variation is not met.

To connect our views with biological practice, we have shown that understanding these issues in terms of abstraction in scientific models illuminates important methodological points. Pere Alberch's paradigmatic studies of developmental constraints in the 1980s distinguish between causal completeness and explanatory salience, and they can be understood as foregrounding the explanatory role of developmental factors in particular evolutionary transitions. Moreover, more recent

molecular studies in evo-devo can be shown to employ similar argument patterns. The usual assumption that developmental constraints or drive must be investigated by way of optimality analyses is thus mistaken: The foregrounding of developmental processes is possible with the laboratory resources familiar to developmental biologists.

References

Alberch P (1982) Developmental constraints in evolutionary processes. In: Bonner JT (ed) Evolution and development. Springer, Heidelberg, pp 313–332

Alberch P (1985) Developmental constraints: why St. Bernards often have an extra digit and poodles never do. Am Nat 126(3):430–433

Alberch P, Gale EA (1985) A developmental analysis of an evolutionary trend: digital reduction in amphibians. Evolution 39(1):8–23

Amundson R (1994) Two concepts of constraint: adaptationism and the challenge from developmental biology. Philos Sci 61(4):556–578. http://www.jstor.org/stable/188335

Amundson R (2005) The changing role of the embryo in evolutionary thought. Roots of Evo-Devo. Cambridge University Press, Cambridge

Ariew A (2003) Ernst Mayr's' ultimate/proximate' distinction reconsidered and reconstructed. Biol Philos 18(4):553–565

Arthur W (2001) Developmental drive: an important determinant of the direction of phenotypic evolution. Evol Dev 3(4):271–278

Beatty J (1994) The proximate/ultimate distinction in the multiple careers of Ernst Mayr. Biol Philos 9(3):333–356

Brigandt I (2010) Beyond reduction and pluralism: toward an epistemology of explanatory integration in biology. Erkenntnis 73(3):295–311

Calcott B (2013) Why how and why aren't enough: more problems with Mayr's proximate–ultimate distinction. Biol Philos 28:767–780

Darwin CR (1859) On the origin of species. John Murray, London

Dickins TE, Barton RA (2012) Reciprocal causation and the proximate–ultimate distinction. Biol Philos 28:747–756

Futuyma DJ (1998) Evolutionary biology. Sinauer Associates Incorporated, Sunderland

Gardner A (2013) Ultimate explanations concern the adaptive rationale for organism design. Biol Philos 28:787–791

Gompel N, Prud'homme B, Wittkopp PJ, Kassner VA, Carroll SB (2005) Chance caught on the wing: cis-regulatory evolution and the origin of pigment patterns in drosophila. Nature 433(7025):481–487

Gould SJ (1983) The hardening of the modern synthesis. In: Grene M (ed) Dimensions of Darwinism: themes and counter themes in twentieth-century evolutionary theory. Cambridge University Press, Cambridge

Gould SJ (1989) A developmental constraint in cerion, with comments of the definition and interpretation of constraint in evolution. Evolution 43(3):516–539

Gould SJ (2002) The structure of evolutionary theory. Harvard University Press, Harvard

Gould SJ, Lewontin RC (1979) The spandrels of san marco and the panglossian paradigm: a critique of the adaptationist programme. Proc Biol Sci 205(1161):581–598

Haig D (2013) Proximate and ultimate causes: how come? And what for? Biol Philos Biol Philos 28(5):781–786

Hochman A (2012) The phylogeny fallacy and the ontogeny fallacy. Biol Philos 28(4):593–612

Hodgkinson A, Eyre-Walker A (2011) Variation in the mutation rate across mammalian genomes. Nat Rev Genet 12(11):756–766

Knell S, Weber M (2009) Menschliches Leben. De Gruyter, Berlin

Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T (2011) Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a pitx1 enhancer. Science 334(6062):1512–1516

Laland KN, Odling-Smee J, Hoppitt W, Uller T (2012) More on how and why: cause and effect in biology revisited. Biol Philos 28(5):1–27

Love A (2006) Evolutionary morphology and evo-devo: hierarchy and novelty. Theory Biosci 124(3–4):317–333

Machamer P, Darden L, Craver CF (2000) Thinking about mechanisms. Philos Sci 67(1):1–25

Maynard-Smith J, Burian R, Kauffman S, Alberch P (1985) Developmental constraints and evolution: a perspective from the mountain lake conference on development and evolution. Q Rev Biol 60(3):265–287

Mayr E (1961) Cause and effect in biology. Science 134(3489):1501–1506. http://www.jstor.org/stable/1707986

Mayr E (1984) The triumph of evolutionary synthesis. Times Lit Suppl 2:1261–1262

Mayr E (1997) The objects of selection. Proc Natl Acad Sci U S A 94(6):2091–2094

Minelli A (2003) The development of animal form. Ontogeny, morphology, and evolution. Cambridge University Press, Cambridge

Odling-Smee FJ, Laland KN, Feldman MW (2003) Niche construction: the neglected process in evolution. Princeton University Press, Princeton

Pigliucci M, Boudry M (2013) Prove it! The burden of proof game in science vs. pseudoscience disputes. Philosophia. doi:10.1007/s11406-013-9500-z

Prud'homme B, Gompel N, Carroll SB (2007) Emerging principles of regulatory evolution. Proc Natl Acad Sci 104(S1):8605–8612

Sober E (1998) Six sayings about adaptationism. In: Hull DL, Ruse M (eds) The philosophy of biology. Oxford University Press, Oxford

Wagner GP (2000) What is the promise of developmental evolution? Part I: why is developmental biology necessary to explain evolutionary innovations? J Exp Zool 288(2):95–98

Wagner GP (2001) What is the promise of developmental evolution? Part II: a causal explanation of evolutionary innovations may be impossible. J Exp Zool 291(4):305–309

Wagner GP, Larsson HCE (2003) What is the promise of developmental evolution? III. The crucible of developmental evolution. J Exp Zool Part B Mol Dev Evol 300(1):1–4

Wallace B (1986) Can embryologists contribute to an understanding of evolutionary mechanisms? In: Bechtel W (ed) Integrating scientific disciplines. Kluwer, Dordrecht

West-Eberhard MJ (2003) Developmental plasticity and evolution. Oxford University Press, Oxford

West-Eberhard MJ (2005) Phenotypic accommodation: adaptive innovation due to developmental plasticity. J Exp Zool Part B Mol Dev Evol 304B(6):610–618