Correlation or Coevolution? Investigating the Genetic Architecture of Male and Female Genitalia in the *Drosophila melanogaster* Subgroup

by

Eden W. McQueen

BA, Bryn Mawr College, 2009

Submitted to the Graduate Faculty of

the Dietrich School of Arts and Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2020

UNIVERSITY OF PITTSBURGH

DEITRICH SCHOOL OF ARTS AND SCIENCES

This dissertation was presented

by

Eden W. McQueen

It was defended on

September 25th 2020

and approved by

Mark Rebeiz, Department of Biological Sciences

Tia-Lynn Ashman, Department of Biological Sciences

Walter Carson, Department of Biological Sciences

Corinne Richards-Zawacki, Department of Biological Sciences

Nathan Clark, Department of Biological Sciences

Dissertation Director: Mark Rebeiz, Department of Biological Sciences

Correlation or Coevolution? Investigating the Genetic Architecture of Male and Female Genitalia in the *Drosophila melanogaster* Subgroup

Eden W. McQueen, PhD

University of Pittsburgh, 2020

To understand the evolution of organismal traits, one must consider that not all traits are fully independent at the developmental level. This lack of independence can result from several mechanisms, one of which is gene regulatory network (GRN) re-use. Novel developmental outcomes resulting from the redeployment of existing GRNs may be advantageous if these outcomes are adaptive. However, because redeployed developmental networks are shared by multiple traits, GRN reuse also act as a constraint if lacking trait independence is detrimental. In my doctoral work, I investigated trait integration and GRN reuse (network co-option), in two ways. In Chapter one, I explore of the phenomenon of network co-option from a theoretical standpoint. I discuss the various forms that network co-option can take, as well as the possible consequences of network co-option with regard to evolution. In Chapter two, I use novel morphological structures in the male and female genitalia of the *Drosophila melanogaster* subgroup as an empirical case study. The male structure, the posterior lobe, and the female structure, the oviscapt pouch, arose concomitantly in this lineage, and are correlated in size across species. Male and female genitalia develop from serially homologous segments and share some of their developmental programming, suggesting that the novel traits may also be developmentally linked. We performed a QTL analysis of the posterior lobe and oviscapt pouch using two species in this group (*D. simulans* and *D. mauritiana*), and found that many loci associated with the size divergence of these two traits occur in the same genomic intervals. While further experimental work will determine whether loci responsible for the male and female QTL in these regions are shared (pleiotropic) or linked, the results suggest a possible role of overlapping developmental programming in the across-species correlation. Moreover, I explore a known case of network co-option involved in the origination of the posterior lobe, and find that genes and regulatory regions of those genes from a co-opted network also show activity in the female genitalia. This is evidence that

network co-option may have played a role in the concomitant appearance of these structures and their apparent coordinated evolution.

Table of Contents

List of Tables

List of Figures

Preface

I want to delight in the smallest of small things, a bit of moss 2 cm in diameter on a little piece of rock, and I want to try here what I have been wishing for so long, namely to copy these tiniest bits of nothing as accurately as possible just to realize how great they are. I've already started that but is so dreadfully difficult. With your nose right on top of it, you see all of its beauty and all of its simplicity, but when you start drawing, only then do you realize how terribly complicated and shapeless that beauty really is.

- MC Escher, Letter to Jan van der Does de Willebois, Ravello, Spring 1923

1.0 On the Specificity of Gene Regulatory Networks: How Does Network Re-use Affect Subsequent Evolution?

The content presented in this chapter was published in the journal *Current Topics in Developmental Biology* [McQueen and Rebeiz, 2020].

The process of multicellular organismal development hinges upon the specificity of developmental programs: for different parts of the organism to form unique features, processes must exist to specify each part. This specificity is thought to be hardwired into gene regulatory networks, which activate cohorts of genes in particular tissues at particular times during development. However, the evolution of gene regulatory networks sometimes occurs by mechanisms that sacrifice specificity. One such mechanism is network co-option, in which existing gene networks are re-deployed in new developmental contexts. While network co-option may offer an efficient mechanism for generating novel phenotypes, losses of tissue specificity at redeployed network genes could restrict the ability of the affected traits to evolve independently. At present, there has not been a detailed discussion regarding how tissue-specificity of network genes might be altered due to gene network co-option at its initiation, as well as how trait independence can be retained or restored after network co-option. A lack of clarity about network co-option makes it more difficult to speculate on the long-term evolutionary implications of this mechanism. In this review, we will discuss the possible initial outcomes of network co-option, outline the mechanisms by which networks may retain or subsequently regain specificity after network co-option, and comment on some of the possible evolutionary consequences of network co-option. We place special emphasis on the need to consider selectively-neutral outcomes of network co-option to improve our understanding of the role of this mechanism in trait evolution.

1.1 HOW DOES ONE PART BECOME DIFFERENT FROM OTHER PARTS?

The biology of gene regulatory networks (GRNs) has played a key role in our understanding of how parts are differentiated during development [Davidson, 2010]. Each gene (or network "node") within a GRN is deployed through the action of transcription factors which bind specifically to *cis*-regulatory elements (CREs) to activate its tissue-specific expression [Levine, 2010, Farley et al., 2015]. Phenotypic changes can often be traced to changes in GRN structure that have tissue-specific effects [Wray, 2007, Carroll, 2008, Stern and Orgogozo, 2008], and thus understanding the mechanisms by which GRNs can be modified gives us insight into evolution. One mechanism that has emerged as a potential player in the evolution of GRNs is the phenomenon of "gene network co-option" (For definition, see **Box 1**), particularly in the origins of novel phenotypes [True and Carroll, 2002, Olson, 2006, Shubin et al., 2009, Monteiro, 2012, Peter and Davidson, 2015]. Changes to a single regulator (an "initiating *trans* change") in an existing GRN could recruit many terminal effectors in just one or a few steps to produce a novel phenotype, rather than a slow accumulation of the necessary mutations in the CRE of each effector (**Figure 1**).

While co-option is a mechanism for rapidly establishing a complex network in a tissue, because the CREs of a co-opted GRN have their function expanded, network co-option is predicted to cause an immediate loss of the tissue-specificity for the reused CREs [Duboule and Wilkins, 1998, Rebeiz et al., 2015, Rice and Rebeiz, 2019]. If a large number of co-opted CREs are causally linked to an increased number of phenotypes (i.e. increased pleiotropy for every co-opted CRE), this lack of specificity could be detrimental over evolutionary time, as it may preclude the independent movement of affected traits towards fitness maxima, at least via changes to those CREs [Fisher, 1930, Hansen, 2003]. That is, an excess of pleiotropic linkage between traits as a result of co-option could ultimately act as a relative constraint, impinging on the evolvability of traits. Since we generally do not observe such strong pleiotropic constraints [Wagner and Zhang, 2011], we must either assume that network co-option is quite rare, or explain why repeated occurrences of network co-option do not hamper evolvability. The growing number of studies that implicate co-option suggest that

Box 1

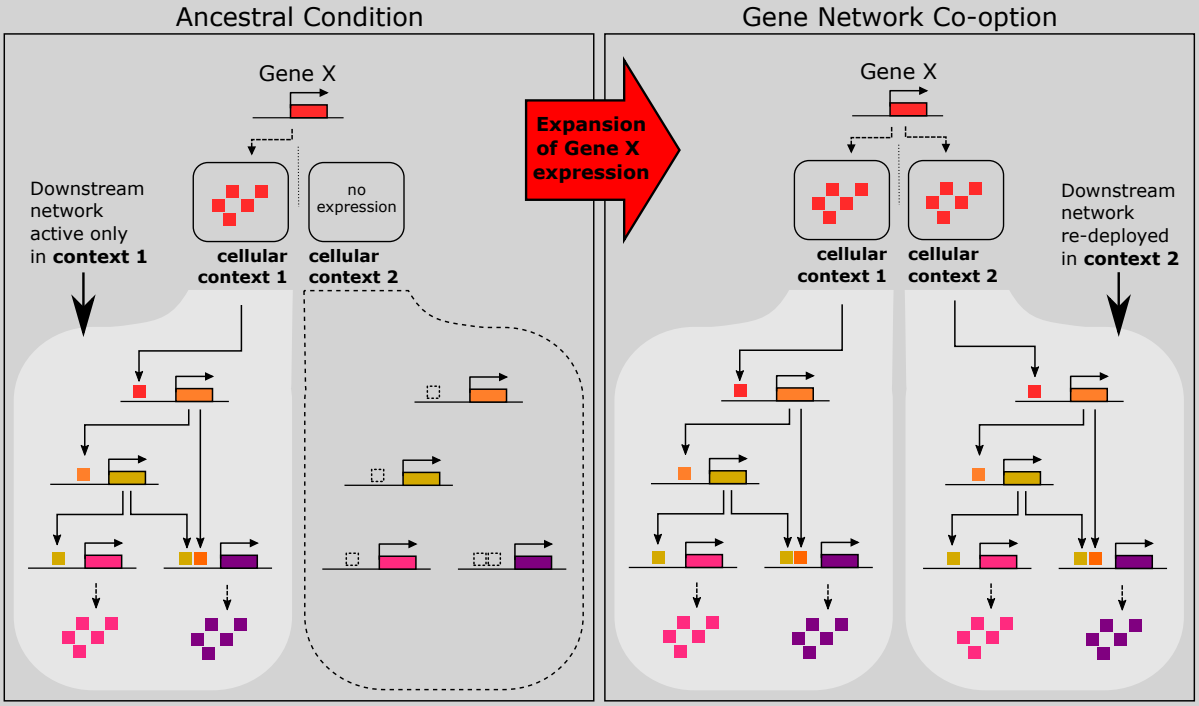

Figure 1: Redeployment of a gene regulatory network via gene network co-option. (Left) The ancestral condition reflects that gene X directly regulates downstream targets only in cellular context 1 and not in context 2 because it is not expressed there. (Right) Co-option occurs when there is an expansion of the expression domain of gene X, such that it is now also expressed in cellular context 2 in the derived condition. The novel expression of gene X in cellular context 2 results in the redeployment of the downstream targets of gene X in cellular context 2, employing its existing cis-regulatory elements (CREs).

Box 1, Cont... The term "co-option" has been applied to a range of phenomena. For instance, co-option is sometimes used to describe a case in which a gene network with a known ancestral function may evolve along a certain lineage to be employed instead for a novel derived function (e.g., [Hinman et al., 2007, Suryamohan et al., 2016]). Our usage here is somewhat more restricted. For our purposes, "gene network co-option" (or "network co-option") is a specific way of modifying the developmental "program". In network co-option, a regulatory factor is deployed in a new location or at a new time during development such that this factor interacts with already existing cis-regulatory elements (CREs) in the next developmental time step. These extant CREs were previously functional in the process of specifying some other trait, i.e., regulated nodes in an existing part (gene regulatory network or "GRN") of the developmental program. Thus, the activation of these CREs may initiate a second instantiation of some or all of the subsequent time steps of that preexisting program (**Figure 1**). The regulatory machinery that defines the GRN of this other trait is therefore being reused, recruited, or "co-opted" to a new location or at a novel point in time [Shubin et al., 2009, True and Carroll, 2002]. Our usage therefore defines co-option as a mechanism, not as an outcome *per se*. This distinction is important, as the deployment of an existing GRN in a novel location could occur by other mechanisms, such as de-novo construction of network connections or some combination of de-novo building and co-option. "Co-option" of a terminal effector gene via changes to that gene's locus is not conceptually distinct from what we describe here (e.g., [Gompel et al., 2005], but our focus is specifically co-option of multiple interconnected elements in networks simultaneously. Other closely related and interesting phenomena that we do not discuss here are co-option of host gene expression by pathogens (e.g., [Faust et al., 2017, Saeij et al., 2007]) and the alternate developmental trajectories induced in cancer cells via co-option of extant network architecture (e.g., [Minafra et al., 2014, Shah et al., 2013]). It would be interesting to connect these areas of research to the concepts discussed here in the future.

this mechanism is common enough to warrant a discussion of the latter. We presume that networks as a whole regain or maintain specificity after network co-option, as we observe some degree of modularity for GRNs that have thus far been characterized [Davidson and Erwin, 2006, Wagner et al., 2008, Sabarís et al., 2019]. What we would like to provide here is a thorough analysis of how the specificity of network nodes fluctuates over time after network co-option. We hope this will be informative for our understanding of co-option as a mechanism, and in particular, our understanding of how this mechanism might relate to evolvability. We will break down the phenomenon by first outlining the range of possible immediate outcomes for a given instance of network co-option. We will then describe the mechanisms by which network nodes may either retain or subsequently regain specificity of their *cis*-regulatory information. Finally, we will draw on our outline of this mechanism to discuss the potential role(s) of network co-option in the evolution of organismal parts.

1.2 IMMEDIATE OUTCOMES OF NETWORK CO-OPTION

Network co-option events can theoretically yield a wide range of outcomes. In most cases, there will exist many differences in the *trans*-regulatory landscape of the cells of the novel (differing in space or time) context and that of the context in which the network operated previously. Distinct regulatory information in the novel cellular context can intersect or interfere with the newly redeployed network at any point downstream of the initiating *trans* change, and thus individual cases of network co-option may differ in the number of network genes redeployed, as well as the identities of downstream targets. We can visualize the spectrum of possible outcomes at the initiation of network co-option by outlining four broad categories (**Figure 2**): Wholesale co-option, partial co-option, functionally divergent co-option, and aphenotypic co-option.

1.2.1 Wholesale co-option

One possibility is that the entire, or nearly the entire network downstream of the initiating *trans* change is redeployed in the novel tissue. We call such cases "wholesale co-option". The result of wholesale co-option is that the same set of terminal effectors is activated in the novel location, and there will be a recapitulation or near-recapitulation in the novel location of the trait generated by the network downstream of the initiating *trans* change in the ancestral location (**Figure 2C**). Gain-of-function homeotic transformations are an illustrative example of wholesale network re-use, where the initiating *trans* change may be the introduction of a *Hox* gene. For example, in *Drosophila melanogaster* the antennae can be transformed into legs by the overexpression of the homeobox gene *Antennapedia* [Schneuwly et al., 1987]. Here, the addition of a single upstream factor results in the deployment of an entire leg formation network in a different location, with the terminal result being easily recognizable as the trait for which the network is generally employed in wild type animals. Likewise, misexpression of the *eyeless* (*ey*) gene in *Drosophila melanogaster* is capable of generating ectopic eyes [Halder et al., 1995]. Similar kinds of homeotic transformations involving changes to single factors have also been observed in floral parts [Coen and Meyerowitz, 1991, Álvarez Buylla et al., 2010]. Wholesale co-option might also be common when repeated structures, such as neurons, muscles, epithelial appendages, and even serially-homologous body segments increase in number, as these networks have already been subject to recurrent reuse and therefore may possess nodes capable of "selector-like" [García-Bellido, 1975, Mann and Carroll, 2002] or "input-output" [Stern and Orgogozo, 2009] function (i.e. largely sufficient to produce the phenotype). For example, Marcellini and Simpson [2006] showed that the expanded domain of a single enhancer of the gene *scute* was sufficient to explain the derived condition in which the number or dorsocentral bristles increased from two to four in the Drosophilid species *Drosophila quadrilineata* [Marcellini and Simpson, 2006]. The *D. quadrilineata* enhancer was able to recapitulate the derived condition when used to drive *scute* expression in *D. melanogaster* (ordinarily possessing only two dorsocentral bristles), demonstrating that the existing downstream regulatory logic was used in the construction of the novel pair of bristles, consistent with wholesale co-option. We define wholesale co-option

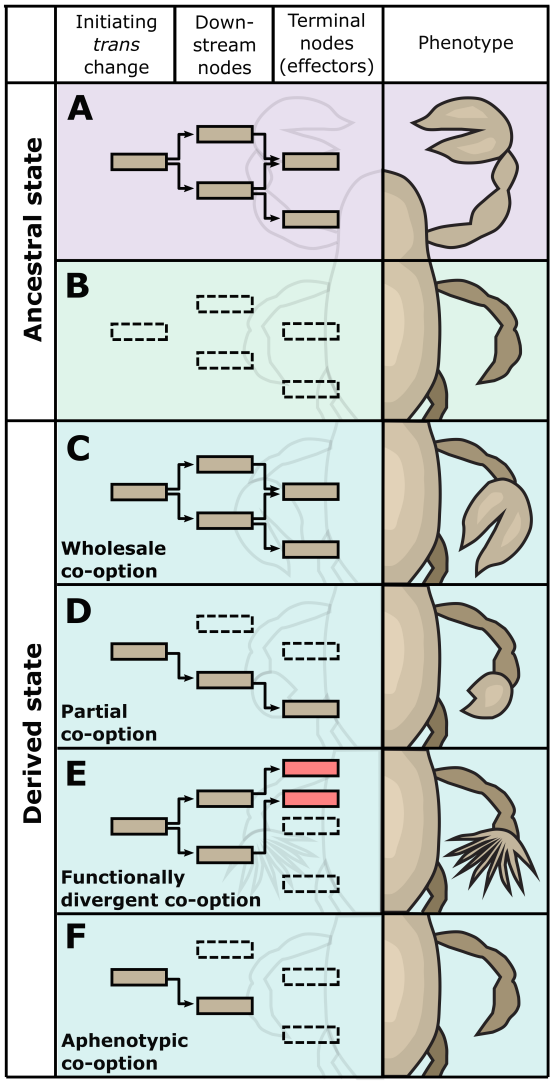

Figure 2: Network Co-option results in a range of possible outcomes. (A) Ancestral location and outcome of network deployment, showing the phenotype in that context. (B) Prior to network co-option, the network is inactive in the second location. (C–F) Activation of the upstream "initiating trans factor" results in redeployment of some or all of this network in the second location. (C) Wholesale co-option involves the redeployment of the entire network in the novel context, resulting in the recapitulation of the phenotype that appears in the ancestral context. (D) Partial co-option, in which some of the downstream transcription factors and terminal effectors are not redeployed in the novel context. The phenotype in the novel context may share some features with the phenotype in the ancestral context. (E) Functionally-divergent co-option is similar to D, except that in the novel context, some of the downstream targets of the redeployed network are distinct from the ancestral context. The phenotype is not necessarily recognizable as being associated with the phenotype in the ancestral context. (F) Aphenotypic co-option, in which no terminal effectors are activated, although there are changes to the upstream developmental program. No phenotype is observed, apart from changes of expression that can be detected experimentally.

as an instance of network co-option in which the direct downstream consequences of the redeployment of the initiating *trans* factor are identical (or nearly identical) in the novel cellular context to those in the ancestral cellular context. This definition does not specify what qualifies as a recapitulated "trait." A trait does not, for example, need be a discrete organ like a bristle, eye, or wing. It could instead be a characteristic such as pigmentation or a chemical signal, provided that the initiating *trans* change is sufficient to recapitulate the downstream effect. For instance, a possible example of wholesale co-option is in the exoskeletalization of the elytra (exoskeletalized forewings) of beetles. Experiments on *Tribolium castaneum* suggest that the derived state of elytral exoskeletalization may have evolved via redeployment of the entire exoskeletalization network of the body wall to the beetle forewing, involving novel upstream regulatory roles of existing wing-patterning elements [Tomoyasu et al., 2009]. The recapitulated "trait" in this case is the exoskeletal fate of cells in the forewing tissue.

1.2.2 Partial network co-option and functionally-divergent network co-option

Wholesale co-option represents the most comprehensive case of network re-use. The striking differences between structures that we believe were built via co-option and the ancestral traits from which the networks were co-opted suggests that simple transplantation of entire networks will be rare. Rather, we anticipate that in the majority of co-option cases, only a portion of the network downstream of the initiating *trans* change will be redeployed [Erwin, 2020]. We refer to cases wherein some subset, but not the entirety of a downstream network is re-deployed as "partial co-option" (**Figure 2D**). Network architecture can be highly context-dependent [Luscombe et al., 2004], and the fidelity of the network redeployment can range from substantial, in which case many features of the ancestral trait are identifiable, to quite minimal, such that the imported elements of the ancestral trait are unrecognizable, or nearly so, in the novel context. Many factors may come into play to prevent activation of some downstream nodes, including (but not be limited to): other tissue specific transcription factors, tissue-specific post-transcriptional modification (e.g. splicing, phosphorylation, protein cleavage), extrinsic signaling from adjacent tissues [Barolo and Posakony, 2002], and boundary conditions set up by developmental timing or mechanical constraints [Davidson,

2012, Green and Batterman, 2017, Womack et al., 2019]. A similar phenomenon occurs in a case of what we will call "functionally divergent" co-option (**Figure 2E**. In these cases, upstream network architecture is co-opted, but the terminal effectors activated by that upstream network differ in the novel context. The upstream nodes would be active in both ancestral and novel settings, but the CREs of the distinct downstream nodes would not be. We sketch out the distinctions between wholesale, partial, and functionally divergent co-option largely for theoretical purposes here, to clearly outline the full range of possible implications for this mechanism. Empirically, these outcomes would only be definitively distinguishable from each other at the time the co-option was first initiated, and would require detailed knowledge of the network CREs, such that the activity of those CREs and the downstream targets of network genes could be compared across contexts. Any subsequent changes to these networks would obscure the distinction between categories of co-option. Many well-known examples of network co-option referenced in the literature likely represent instances of partial co-option, functionally-divergent co-option, or some combination of both. For example, in contrast to the ectopic eyes generated by *Pax6* mis-expression, a fascinating example of a possible partial network co-option of an eye network was found in a study of extinct dipterans. Two species in the genus *Eohelea* possessed a structure on their wings that bore a remarkable resemblance to the compound eyes of individuals of those same species, leading the authors to conclude that this structure was likely built through network re-use. However, it appears that in the case of this novel wing structure, the novelty consisted only of the cuticular part of the eye, and was not an entire ectopic eye [Dinwiddie and Rachootin, 2011]. We note that in this case because these are not extant species, it is not possible to distinguish between a partial network co-option and a wholesale co-option of a single independent part of the eye network. Functionally divergent co-option may often result from the co-option of signaling pathways, which are utilized throughout development and quite commonly implicated in the formation of novel traits [Loredo et al., 2001, Harris et al., 2002, Cebra-Thomas et al., 2005, Harris et al., 2005, Wasik and Moczek, 2011, Nakamura et al., 2015]. For example, studies on butterfly eyespots suggest that the evolution of these novelties likely included functionally divergent co-option of a deeply conserved anterior-posterior boundary-forming network. The downstream consequences of this network in the novel con-

text provides pattern information for the color phenotypes manifested in scales [Carroll et al., 1994, Keys et al., 1999]. Thus, an important process was co-opted, that is, the formation of a particular transcription factor landscape pattern, yet the downstream targets differed in the novel location. We currently do not understand how these functionally divergent outcomes became connected to the anterior-posterior boundary network, and many hypotheses exist [Özsu and Monteiro., 2017]. A similar case has been suggested in plants, for which a abaxial-adaxial polarity gene network responsible for the flattening of organs such as leaves appears to have been co-opted to cause flattening of stamen filaments, a derived condition [Almeida et al., 2014]. A well-known example that appears to support partial co-option is the redeployment of the leg network in beetle head horns [Moczek and Nagy, 2005]. Not every gene of the canonical leg network is actually required for horn development, as evidenced by the lack of a phenotype when the gene *dachshund* was knocked down [Moczek and Rose, 2009]. There is also evidence for partial co-option in the case of the posterior lobe (a genital structure) of fruit flies in the *Drosophila melanogaster* clade. The evolution of this novel structure appears to have involved co-option of part, but not all, of a network responsible for the development of an ancestral larval structure [Glassford et al., 2015]. Additional examples include tree-hopper helmets, the evolution of which appears to have involved co-opted elements of the wing-patterning network [Prud'homme et al., 2011, Fisher et al., 2019], bilaterian appendages, which may have involved redeployment of an existing network for anteroposterior patterning[Lemons et al., 2010], and the use of *Hox* genes in the evolution of paired vertebrate limbs [Zakany and Duboule, 2007]. We imagine that these cases represent an amalgam of partial and functionally-divergent co-option, but we are still uncovering the full picture of how the ancestral networks were reused and rewired. Future work characterizing these networks more extensively will help us understand how and when nodes were lost and gained across contexts.

1.2.3 Co-option resulting in downstream regulatory expression only (aphenotypic co-option)

Finally, network co-option may involve the introduction of an upstream regulator that results in the deployment of some of the upstream network nodes in the novel context, but causes no phenotype (defined as measurable change in morphology/behavior); only transcription factor expression patterns are altered. Such cases of "aphenotypic" co-option could ensue if there is a total lack of activation, or inadequate activation, of required terminal effector genes (**Figure 2F**). Aphenotypic co-option is essentially an extreme case of partial network co-option. As the term "pleiotropy" is usually restricted to mean that changes to one locus can induce multiple phenotypes [Wagner et al., 2008], aphenotypic co-option generates no new pleiotropy, although changes to the developmental program have occurred. Examples of aphenotypic co-option are lacking, but there are reasons to believe they exist, or at least we have evidence of the possibility in that we often observe expression patterns for which we can offer no functional explanation. For instance, RNAi screens sometimes find that knockdowns of some transcription factors expressed in the tissue of interest have no phenotype (e.g. [Staller et al., 2013, Zattara et al., 2016]). Similarly, a comparison of the expression of 20 genes in imaginal tissues of four very closely related species in the melanogaster clade uncovered striking differences in expression across species, many of which are not connected to any known phenotype [Rebeiz et al., 2011]. Many CREs for genes exhibit expression patterns outside the focal tissue of a given study, these patterns having no known functional role (e.g. late anterior expression driven by the minimal *even-skipped* stripe 2 enhancer in *Drosophila* embryos [Janssens et al., 2006]). Besides the aforementioned examples, many more cases like these may suffer from the "file drawer problem" [Rosenthal, 1979]. Such results are usually ignored, or interpreted as evidence of robustness, but may sometimes represent cases of aphenotypic network redeployment or non-functional nodes of partially co-opted networks. Aphenotypic co-option as an idea has generally received sparse attention, although it has been mentioned as a possibility in the past [True and Carroll, 2002]. While considering such an outcome on its own may seem irrelevant, when considered in the light of long evolutionary periods, this phenomenon could nonetheless have some quite

interesting implications, as we will discuss in Sections 3 and 4.

1.3 HOW DO GRNS MAINTAIN OR RECOVER SPECIFICITY AFTER NETWORK CO-OPTION?

Restoration of at least partial regulatory specificity of co-opted CREs is almost certain to be a pervasive phenomenon, considering the number of morphological novelties we have discussed here that arose through likely network co-options but now apparently evolve in a largely independent manner (e.g. treehopper helmets [Prud'homme et al., 2011], beetle horns [Emlen et al., 2007, Moczek, 2009], butterfly wing spots [Brunetti et al., 2001, Oliver et al., 2012], and feathers [Prum, 1999, Prum and Brush, 2002, Prum, 2005]). The process of re-establishing CRE tissue specificity could happen in two ways: In *cis*, via changes to the co-opted CREs themselves, or in *trans*, via changes outside the network that introduce tissue-specific regulators of the co-opted CREs. We discuss these possibilities below. We note that although our discussion is centered on co-opted CREs, multiple studies have noted pleiotropy in enhancer sequences in the absence of network co-option events [Nagy et al., 2018, Rebeiz et al., 2011, Preger Ben-Noon et al., 2018], and the mechanisms we describe below apply broadly to the evolution of regulatory specificity.

1.3.1 Changes in *trans*

Many genes outside of the co-opted GRN will likely have pre-existing roles in the tissue that predated the network co-option event. These genes may be available for genetic tinkering that yields tissue-specific modifications, or could contribute to an immediate plastic response that could modulate pleiotropy, and be genetically modified later [West-Eberhard, 2005]. Novel expression domains of such genes beyond the co-opted GRN could also arise subsequently, after co-option has occurred, and be exploited to achieve tissue independence at that future time. A modification of this type can be inferred from the striking instance of wholesale co-option demonstrated for the embryonic skeleton of sea urchins (a derived trait)

which employs a GRN co-opted from adult skeletogenesis [Gao and Davidson, 2008]. In this case, while most of the genes in the co-opted network are shared across the two contexts, implicating wholesale co-option, the embryonic skeletogenesis network incorporates a small number of genes that are not part of the adult skeletogenesis network. One of these genes, *tbrain* (*tbr*), has known ancestral roles in endomesoderm specification in other echinoderms, leading the authors to suggest that the addition of this regulator was a modification to the embryonic skeletogenesis network that occurred after the initial co-option event. Direct regulators of the initiating *trans* factor could be targets for modification, if these differ between the novel and ancestral contexts. For instance, in the example discussed above concerning the exoskeletalization of elytra in beetles, the authors showed that the gene *apterous* (*ap*), which is part of the ancestral wing network, has gained a novel role in redeploying the exoskeleton network in the elytra, whereas *ap* is not a direct regulator of the exoskeleton network in the mesonotum (a cuticular part of a thoracic segment). Consequently, when RNAi was performed targeting *ap*, defects were seen in exoskeletalization of the forewing, but not the mesonotum [Tomoyasu et al., 2009]. This demonstrates that, in principle, exoskeletalization could be targeted independently in the elytra by modifying regulators further upstream of the co-opted network, even if the exoskeletalization network itself were pleiotropic.

1.3.2 Changes in *cis*

If tissue-specificity is to be regained via changes to CREs of nodes in the co-opted GRN, the first requirement is that the ancestral and novel tissues must have at least one qualitative or quantitative difference in cellular regulatory content (e.g. transcription factor identity, activity, or concentration) at the time the node in question is active. That is, there must exist some form of potentially exploitable tissue specificity, otherwise all changes to the CRE would necessarily affect both the ancestral and novel contexts. We suspect that this is usually the case, although in principle it is possible that the regulatory states of the ancestral and novel contexts would not differ at all after co-option (i.e. if the only difference between the two contexts prior to co-option was the presence/absence of the initiating *trans* factor itself). In such cases, multiple mutations would be required to regain tissue-specificity. There

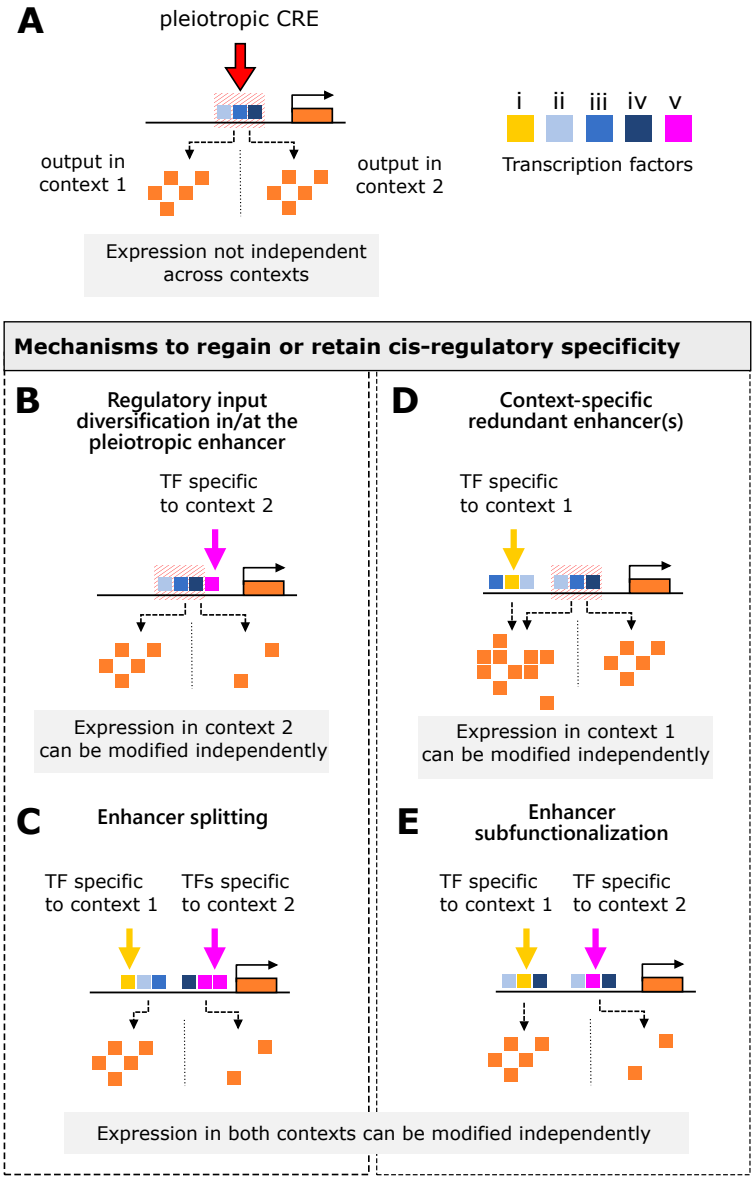

Figure 3: Mechanisms to retain or regain specificity of pleiotropic CREs. (A) A gene that is redeployed during network co-option possesses a pleiotropic cis-regulatory element that drives expression in contexts 1 and 2. The CRE is activated by the binding of transcription factors ii, iii, and iv, and the output of expression in both contexts is not independent. (B) Regulatory input diversification: evolution of a binding site for transcription factor v, which acts as a repressor, only affects expression in context 2, as v is not present in context 1. Further modification can occur to achieve greater or full independence via enhancer splitting (C), in which a single enhancer fragments into two enhancers employing context-specific activators. (D) Redundant enhancers: A second enhancer for the target gene affects expression in context 1 only, due to the fact that this redundant enhancer requires the binding of transcription factor i, which is not present in context 2. Further modification can occur via enhancer subfunctionalization (E), in which redundant enhancers that have partial or full overlap in their expression profiles gain or lose binding sites for context-specific factors to achieve complete independence.

are two primary mechanisms that can mitigate or eliminate the potential for pleiotropic constraint at a co-opted CRE: regulatory input diversification (**Figure 3B**) and context-specific redundant enhancers (**Figure 3D**). The result of these mechanisms would be either a facilitation of independent regulation of the node in the novel and ancestral contexts, or a modulation of the CRE in question in one context (for instance inactivating it in one context). A critical note is that while these features that restore CRE tissue-specificity may evolve subsequent to network co-option, they also may already be in place at the time of co-option. Indeed, the modifications to CREs we describe here would be causal explanations of why some nodes are not expressed in partial or aphenotypic co-option outcomes. Similarly, we note that although entirely novel mutations could be the causal changes in these mechanisms, novel genetic combinations of existing variants already segregating in the population could also alter pleiotropy or mask it through epistasis [Pavlicev et al., 2008]. Given the pervasive nature of epistasis in natural populations [Phillips, 2008, Mackay, 2014] this may be a frequently employed path to recapturing lost specificity. It is also important to mention that tracing the process of how tissue specificity was restored may be very difficult when comparing species that have diverged for long periods of time. Once sufficient time has passed, the footprints of this process will have been erased. As such, evidence for this mechanism will likely be found in cases where at least some of the ancestral pleiotropic and redundant enhancers are still detectable.

1.3.2.1 Regulatory input diversification

Diversification of regulatory inputs (**Figure 3B**) mitigates pleiotropy via binding sites at the pleiotropic enhancer that affect the regulatory outcome of the enhancer differentially across tissues. For example, this might be the gain of a binding site for a repressor that is only present in one tissue. Enhancer splitting (**Figure 3C**) is an extension of the process above, and is related to the idea of enhancer sprawl [Rice and Rebeiz, 2019], in which it is understood that enhancers sometimes expand and contract due to the addition and removal of binding sites via turnover. In this case, tissue-specific binding sites accumulate such that a single enhancer that has some tissue-specific and some pleiotropic binding sites may eventually split into two completely separate tissue-specific *cis*-regulatory elements in adjacent positions on the DNA.

1.3.2.2 Specificity conferred by redundant CREs Redundant CREs (also called "shadow enhancers") are defined here as at least two CREs driving expression of the same target gene in redundant or semi-redundant expression patterns [Hong et al., 2008, Barolo, 2012]. Redundant CREs are a route to at least partial recovery of modularity in cases of co-option because if the CREs employ a unique set of regulators, one or both of the redundant CREs may drive expression differently across the two tissues. We already have several empirical examples of redundant enhancer pairs that display different regulatory logic [Wunderlich et al., 2015, Vincent et al., 2018], lending credence to this potential route to specificity. With redundant CREs, there are two possible conditions. First, a redundant CRE could drive expression in only one of the tissues (**Figure 3D**). This could in fact provide an immediate mode of retaining specificity, as a redundant CRE of this type could already exist in the *cis*-regulatory region of a GRN node at the time of co-option. A redundant CRE of this kind could also evolve later and restore specificity [Rebeiz and Tsiantis, 2017]. However, in the two cases just described, the redundant CRE from the co-opted network is still pleiotropic. CRE sub-functionalization (**Figure 3E**), in which two redundant CREs of a single gene (i.e. a redundant CRE pair) each evolve independent roles specific to one of their initial developmental contexts, would be required to erase all pleiotropic linkages between ancestral and novel contexts (e.g. the " *cis*-regulatory element duplication, degeneration, and complementation" model [Monteiro and Gupta, 2016]. This may or may not be favorable, as robustness via redundant enhancers is also considered to be potentially beneficial [Perry et al., 2010, Frankel et al., 2010, Barolo, 2012].

1.4 THE ACTION OF SELECTION ON CO-OPTED NETWORKS OVER TIME

Co-option is often viewed as a potential mechanism for facilitating the origins of morphological novelties. In other words, the appeal of this mechanism rests in its possible explanatory value with regard to evolution. However, while network co-option is often invoked in this way [True and Carroll, 2002, Olson, 2006, Shubin et al., 2009, Monteiro, 2012,

Peter and Davidson, 2015], in order to appreciate the full evolutionary implications of this mechanism, we must carefully consider the manner in which we expect natural selection and neutral processes to operate on co-opted networks after they occur. As with any mutation, in a simple sense there are three potential fitness effects of an initiating *trans* change (**Figure 4, A**): beneficial, neutral, or deleterious (**Figure 4, B**). When thinking about network co-option and evolution, the added element to consider is the concomitant reduction of tissue-specificity, which may have long term consequences (**Figure 4, C**). Below we discuss the evolutionary implications of pleiotropy at co-opted network CREs, given each of the three possible fitness consequences of the initiating mutation.

1.4.1 Initiating *trans* change with positive fitness effects

If network co-option generates a novel phenotype and the net effect on fitness is beneficial, we expect that the initiating *trans* change, would be under positive selection (**Figure 4, B, i**). If the beneficial phenotype is in the novel deployment context, this situation exemplifies what is imagined to be the major upshot of network co-option as a mechanism for evolutionary change that we discussed in the introduction: a novel, beneficial, phenotype produced in one or just a few evolutionary steps. Still, while the overall phenotype may be beneficial, the effects of pleiotropy resulting from the network co-option may be detrimental, either due to the lack of modularity between the ancestral and novel traits that limits adaptation, or because some nodes have negative pleiotropic effects (**Figure 4, C, ii**). In such cases, selection should favor mutations that maintain the expression of beneficial co-opted nodes in both tissues but restore specificity, for example, inactivation of particular nodes that have negative pleiotropic effects. In spite of the popularity of this view of network co-option, we do not have empirical examples that explicitly demonstrate this sequence of events.

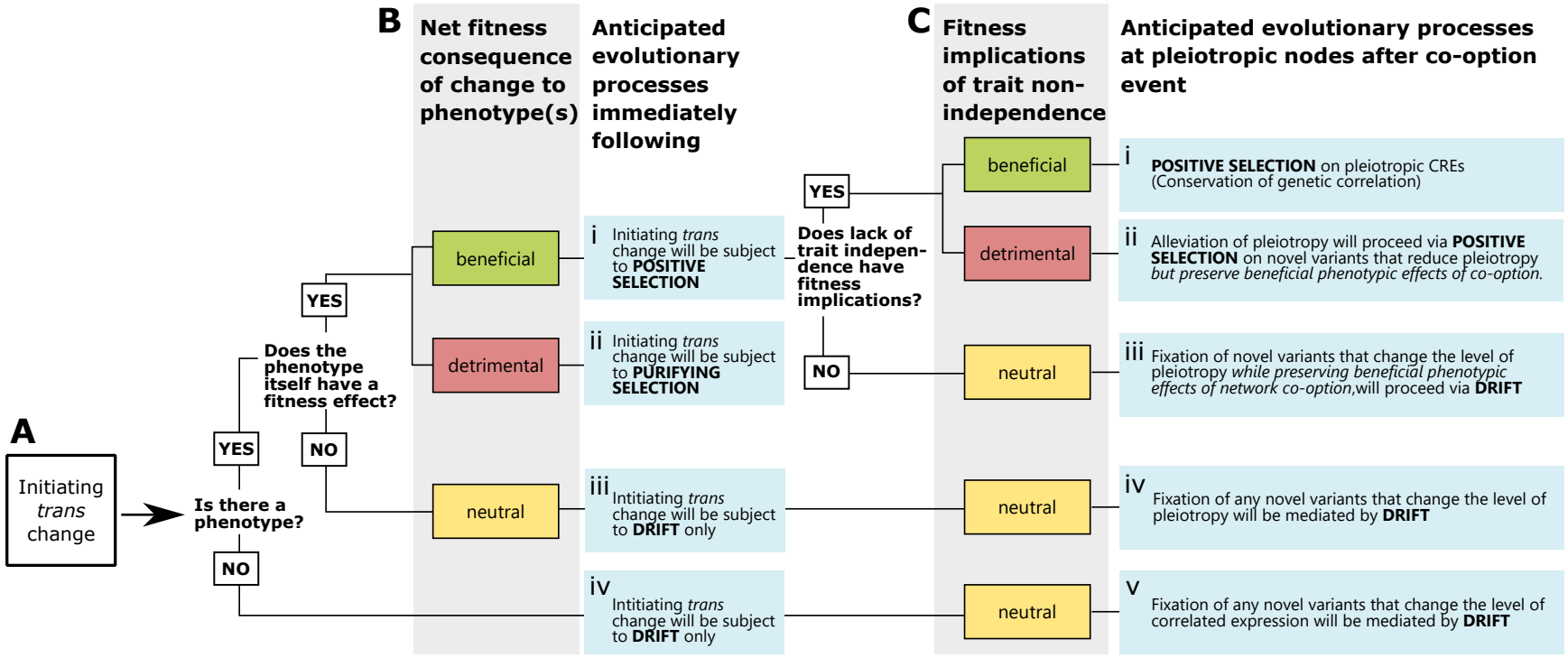

Figure 4: Fitness effects of network co-option evolution at pleiotropic nodes. Figure depicts a decision tree that partitions the evolutionary ramifications of GRN co-option and subsequent modification of pleiotropic nodes. While co-option events that have immediately advantageous effects are often discussed, here we outline a wider range of possibilities that could result from network co-option, including those with neutral outcomes. Starting on the left side of the diagram with the initiating trans change (A), we may first anticipate the evolutionary processes that would follow that mutation based on its phenotypic effects (or lack thereof) (B). After network co-option, the fitness effects of decreased tissue-specificity are considered as the network is subject to additional modifications (C). These would be the predictions for long term selection on the affected network, given the assumption that selection regimes do not change from the initial state. As discussed in the text, selection on these networks/phenotypes may indeed change over time and affect how co-opted networks evolve, adding further complexity to this picture.

In theory, it could be that when a co-option event occurs, the resulting pleiotropy is itself advantageous, in the sense that if the two traits usually experience selection together and in the same direction, the genetic correlation between them would allow selection to operate more efficiently. A simple example would be a trait that tracks environmental conditions, such as fur thickness or cryptic coloration, expanding via co-option to a new location on the body. If selection on the character in both contexts is uniform, the novel and ancestral traits would functionally amount to only one trait with respect to the co-opted network. In such cases, there could be selection against variants that broke up pleiotropy at CREs in the co-opted network, as trait independence would represent an unnecessary increase in complexity, possibly slowing the ability of the population to adapt overall [Orr, 2000, Welch and Waxman, 2003](**Figure 4, C, i**). To our knowledge this particular outcome of network co-option does not currently have direct empirical support, although the idea of selectively maintaining beneficial pleiotropy has been suggested more generally (e.g. between interacting parts such as integrated skeletomuscular traits [**?**]). A possible scenario of this sort could also occur in plants, where it is known that male and female floral parts (androecium and gynoecium) of some species share much of their developmental toolkit [Dornelas et al., 2011]. A correlation of male and female floral structures could be favorable in some cases if it were required for efficient pollination. Alternatively, the network co-option could confer a fitness benefit in the novel location, and the existence of pleiotropic roles of any given CRE could be neutral, or nearly neutral (**Figure 4, C, iii**). This could happen if there is only selection on the beneficial novel trait and the ancestral trait is either completely neutral (i.e. has no function), or if the majority of phenotypic changes to the ancestral trait via mutations at co-opted nodes would be neutral such that the pleiotropy is nearly neutral. Over time, evolution of the pleiotropic *cis*-regulatory regions could erode the genetic correlation by chance if changes arose to increase regulatory independence without negatively impacting the phenotype(s). Otherwise, because the CRE confers a functional benefit to the novel context, it may be conserved and the correlation could be maintained incidentally. A study of the genetic correlations among tetrapod limb developmental serial homologs suggests that covariance structures that result from reuse of networks (as is thought to be the case with hindlimbs and forelimbs [Sears et al., 2015]) can persist for long periods

[Young and Hallgrímsson, 2005]. The authors found that the correlation between the lengths of hind limbs and forelimbs is only broken in cases of extreme functional necessity, such as is observed in the extremely divergent limb and digit proportions that enabled flight in the lineage leading to bats. However, we do not know whether the covariance structure in this case was maintained actively or passively, and the authors conclude that stabilizing selection on such correlations may often be an important factor beyond genetic constraints [Young and Hallgrímsson, 2005, Hallgrímsson et al., 2009]. Another important possibility to consider is that the mutation confers a benefit in the ancestral context and initiates network co-option neutrally elsewhere as a byproduct. In such cases, the initiating mutation could be subject to selection irrespective of the co-option per se (this causal mechanism for selectively neutral traits is discussed in Lovejoy et al. [2002]). A modeling study showed that the addition of genes to a network generally improved the "fit" of the model to its target data, which suggests that recruitment of genes to already functioning networks could be common [Spirov et al., 2012], and might be a source of this type of "collateral" network co-option. This could be difficult to detect, as neutral phenotypes generated by network co-option could appear to be under selection if there is selection on the genetically correlated character [**?**]. We will discuss the implications of this potential outcome more in the section on neutral fitness outcomes below.

1.4.2 Initiating *trans* change with detrimental fitness effects

If network co-option is deleterious, the initiating *trans* change should be lost due to purifying selection (**Figure 4, B, ii**) unless it is fixed by drift, which is more likely in small populations and when the fitness consequence is mild [Fisher, 1930]. It is also possible that the phenotypic consequences of a given co-option event on the novel tissue were initially beneficial or neutral, and only later became detrimental (e.g. accompanying a change in environment that alters selective regime or developmental plasticity, epistatic changes that reveal larger effects on phenotype, etc.). In such cases, the upstream mutation may have already been fixed in the population. In either of the above cases, the detrimental effects of network co-option could either be eliminated by another change at the upstream *trans*

factor that reverts the co-option, or the effects could be reduced over time by evolving tissue specific repression of downstream network nodes individually. Any given case of this latter process would be indistinguishable from an initial state of partial or aphenotypic co-option, although in some cases a comparison across species or populations that diverged after the initiating *trans* change might reveal a history of modifications deactivating the co-opted network.

1.4.3 Initiating *trans* change with selectively neutral effects

One possible outcome of an initiating *trans* change that incurs network co-option is the generation of a phenotype which is completely neutral with respect to fitness (**Figure 4, B, iii**). Another neutral outcome, which we discussed in section one, is the possibility that no phenotype is generated in the novel tissue at all (e.g. aphenotypic co-option) (**Figure 4, B, iv**). In both of these cases, the genetic correlation generated between the tissues is also likely neutral, unless future mutations alter the neutrality of the phenotype or induce a new phenotype via the previously aphenotypic network. Otherwise, both the initiating *trans* change (**Figure 4, B, iii, iv**), and any future mutations that alter the genetic correlation generated by co-option would be fixed only by drift (**Figure 4, C, iv, v**). We have no reason at all to believe that fixation of a mutation of this type this would be more uncommon than the stochastic fixation of any other neutral mutation. This scenario is therefore especially important to consider in small populations that are more heavily influenced by drift. Modeling and analysis of changes to gene expression across species of *Heliconius* butterflies [Catalán et al., 2019], fish [Whitehead and Crawford, 2006], and primates [Khaitovich et al., 2005, Chaix et al., 2008] all showed that the majority of changes to gene expression across species were consistent with neutral evolution, lending credibility to this possibility. With respect to latent expression generated by co-opted networks specifically, we do not currently have examples. However, a study on the evolution of *Onthophagus* beetle horns suggested that exploitation of an existing expression pattern in the beetle anterodorsal head tissue was important to the evolution of the novel horn structures. A key member of this gene network, an ortholog of the *Drosophila* gene *orthodenticle* (*otd*), was also found to be expressed

ancestrally in the anterodorsal head tissue of an outgroup species (*Tribolium castaneum*), which lacks horns. Interestingly, knockdown of *otd* in *Tribolium* does not induce detectable defects in the head, suggesting that this pattern is not functional in *Tribolium* [Zattara et al., 2016].

1.5 NETWORK CO-OPTION AND THE ORIGIN AND DIVERSIFICATION OF TRAITS

Our breakdown of the phenomenon of network co-option in the previous sections now puts us in a position to offer a few discussion points on the relationship between network co-option and the evolvability of traits. This is in no way a comprehensive list. Our comments here will hopefully serve as a jumping-off point for further conversation. First, as we discussed in Section 1, it is important to recognize that the comprehensive case, wholesale co-option, is likely not the most common outcome. We anticipate that many more instances of network co-option will be only partial, and therefore a degree of trait independence may be retained even at the onset of network co-option. It has been suggested that intermediate levels of pleiotropy maximize evolvability [Hansen, 2003], and thus many cases of co-option may be well within the range of pleiotropic effects that do not cause serious problems for evolvability. Nevertheless, in such cases that the pleiotropy generated by network co-option acts as a constraint, many routes exist to modify CREs directly in *cis* or via their regulators to regain specificity, as we discussed in Section 2. Second, we must keep in mind that the effects of pleiotropy are not always detrimental. Not all phenotypes generated by co-option will initially, or ever, affect fitness, and not all co-option events will have a phenotype. These neutral outcomes would still alter modularity in the strict sense that the co-opted CREs would have decreased potential to confer tissue-specificity, however there would be no immediate visibility of these events in terms of selection, and thus the evolvability of ancestral traits would not be affected, at least initially. Models that allow for neutral pleiotropic effects of co-option would improve our understanding in this area. One model predicting the degree to which pleiotropy would act as a developmental constraint revealed that the

level of constraint was sensitive to changing the fitness effect of pleiotropy [Otto, 2004]. As has been pointed out previously in the case of gene pleiotropy [Stern and Orgogozo, 2008], concerns about pleiotropic constraint resultant from network co-option may be mitigated by a clearer understanding of the forms pleiotropy can take as a result of this mechanism. Beyond simply failing to obstruct the evolvability of traits, we should also keep in mind that neutral or nearly neutral outcomes of network co-option that are retained stochastically (phenotypes, expression patterns) could provide a reservoir of cryptic genetic variation. Such variation may have phenotypic and selective consequences later if subsequent mutations activate processes, such as additional network co-option events, downstream of these nodes. Initially aphenotypic outcomes of co-option might therefore contribute positively to trait evolvability (e.g. in cases of preadaptation). This possibility has been noted before [True and Carroll, 2002], but we currently lack empirical examples to support this conjecture. However, there is a growing interest in understanding how cellular and morphological phenotypes may evolve neutrally [Ruths and Nakhleh, 2013, Zhang, 2018, Wideman et al., 2019]. A recent study on cryptic genetic variation demonstrated that neutral mutations accumulated at the level of a single protein facilitated subsequent adaptation of that protein [Zheng et al., 2019]. This result might scale up to the level of networks. More examples such as these that examine multi-gene interactions, and especially comparative analyses of the network architecture of such cases will help us understand the role of network co-option in the generation of cryptic genetic variation. The implications of the observations above are magnified when we consider that the simple version of co-option that begins with GRN deployment in one tissue and expands to deployment in two tissues is probably not realistic. More extensive effects across many tissues are likely to be common. As networks evolve downstream of newly redeployed nodes after network co-option, a complex collage forms rather than a pre-made template which is simply "copy-pasted" to a new location. Indeed, this view is supported by a mathematical modelling study, which demonstrated that the construction of a novel expression domain is facilitated by reuse of multiple but distinct existing modules that contribute to that domain elsewhere [Espinosa-Soto and Wagner, 2010]. Such a scenario, wherein CRE pleiotropy is spread out over multiple ancestral GRNs, might lend more flexibility to circumventing developmental constraints for both the ancestral and

novel structures. To be sure, there are likely to be many cases where network co-option does result in constraint on some properties. For example, it has been suggested that limb outgrowth was constrained to have anterior-posterior polarity due to the fact that *Hox* genes were co-opted to initiate extension from the body wall [Tarchini et al., 2006]. We are still in the process of discovery in the area of network co-option, and there are many ways forward. With respect to modelling, it would be very enlightening to incorporate network co-option into dynamical models [Irons and Monk, 2007, Alexander et al., 2009, Verd et al., 2019], which take spatio-temporal information into account when designating modules. Models incorporating some of the neutral outcomes of co-option that we discussed here would also be very useful. Beyond model development, many more empirical examples of network co-option are needed. In particular, to gain insight into the evolution of co-opted networks, we need examples wherein the structure of known or suspected co-opted networks is compared across species. One study investigated the expression of genes in the network co-opted to generate eyespots across 21 species of *Nymphalid* butterflies [Oliver et al., 2012]. They showed that the expression of some network members was highly conserved, whereas repeated losses of others suggests that these nodes were more evolutionarily labile or possibly not necessary in the first place. More examples like this one would greatly improve our understanding of how co-opted networks are incorporated into existing networks and change over time. We hope the framework that we have outlined highlights the scope of possibilities that accompany network co-option and inspires a wide range of research questions into this intriguing mechanism of developmental evolution.

1.6 CONCLUDING REMARKS

In writing this review, we were reminded of the way that general thinking has progressed with regard to genes. What began by attributing strict functional identities to individual genes ("a gene for function x"), eventually became more nuanced in light of empirical data that was inconsistent with a one-to-one view [Duboule and Wilkins, 1998]. Considering network reuse as a mechanism of altering development similarly complicates our concept of

GRN identity. GRNs are not tidy, self-contained "programs" for specific traits [Nijhout, 1990, DiFrisco and Jaeger, 2019], but are instead highly context-dependent and may therefore yield different outcomes in different developmental circumstances. This suggests that we must caution ourselves against falling into a "GRN for function x" trap, and instead recognize that the GRN for any given trait will be a haphazard assembly of parts, often with a few spare odds and ends, drawn from existing GRNs over evolutionary time. Like all products of evolution, GRNs will be the result of evolutionary "tinkering" [Jacob, 1977]: functional, but messy.

2.0 Many Co-localizing QTL Underlie Species-level Differences in a Pair of Novel Male and Female Genital Structures

2.1 INTRODUCTION

It has been remarked over and over again how harmoniously the whole organism hangs together, and how throughout its fabric one part is related to another in strictly functional correlation. But this conception, though never denied, is sometimes apt to be forgotten in the course of that process of more and more minute analysis by which, for simplicity's sake, we seek to unravel the intricacies of a complex organism.

- D'Arcy Thompson, On Growth and Form, 1917, p.262

A longstanding and difficult question of evolutionary biology is how we are to understand the evolution of individual organismal features in light of the fundemental integration of those features into the whole organism. It can be tempting to take Darwin's picture of evolution by natural selection to an extreme, treating the characteristics of an organism almost as independent actors under selection. Indeed, the power of selection to produce extravegantly exaggerated individual characters is obvious. Nevertheless, developmental biologists such as D'Arcy Thompson bristled when confronted with notions of unreservedly fragmenting the organismal unit ([Thompson, 1917]). While acknowledging that the study of organismal complexity is a challenge that requires some degree of abstraction, Thompson's hollistic view of the organism suggested that when considering the evolution of forms, some traits must evolve synchronously because the functioning of the whole organism imposes its own set of constraints.

At about the same time Thompson published his seminal work on that subject, another key concept of organismal integration was being developed in the burgeoning realm of genetics. This phenomenon was called 'pleiotropy,' and refers to the condition in which a change at one genetic locus has an effect on multiple traits ([Paaby and Rockman, 2013]). Although the idea of pleiotropy had been known for some time even before the term was

coined, its mechanisms and implications have been expanded upon immensely over the past century ([Stearns, 2010]). Pleiotropy is often discussed in terms of the constraints on evolution that might be imposed by traits that share a genetic basis, as mutations in a shared locus might be favorable for one trait but detrimental to another ([Lande and Arnold, 1983, Arnold, 1992, Roff and Fairbairn, 2007]). However, a great deal of work over the years has also developed the idea that pleiotropy, by maintaining some level of genetic integration of developmental programs, is foundational to the maintenance of organismal modularity—the ability for clusters of related traits to evolve together while remaining mostly independent from other such clusters ([Welch and Waxman, 2003, Li et al., 2006, Wagner et al., 2008]). Thus, while pleiotropy can impose constraints in many cases, the genetic integration of some traits may be a key part of the explanation for their ability to evolve more freely in the context of a whole organism ([Wagner, 2014]).

When we observe correlations between two or more organismal traits across species or populations, we must therefore consider the possibility of both functional and genetic integration when devising possible explanations for their evolutionary trajectories. Only in recent years have the tools been developed to investigate the latter of these two axes of integration in any great detail, and the question of how correlated traits evolve at the molecular level remains a dynamic one to this day.

An exceptionally interesting system to consider in this light is the evolution of male and female genitalia in animals. In both males and females, the genitalia of internally fertilizing taxa are wildly diverse, often being the only reliable distinghishing features amoung closely related species ([Eberhard, 1985, Hosken and Stockley, 2004, Simmons, 2014, Sloan and Simmons, 2019]). There is a general consensus that the diversity of animal genitalia is most often the result of sexual selection on these characters [Simmons, 2014, Brennan and Prum, 2015, Sloan and Simmons, 2019], although the stunning prevalence of the trend of rapid evolution of genitalia still leaves open many questions [Yassin, 2016].

The critical need to maintain the functional relationship of male and female genitalia seems almost to require that the genitalia of males and females coevolve ([Brennan and Prum, 2015]). Indeed, correlations between the sizes or shapes of male and female genitalia across species or populations have been observed in numerous cases (e.g. [Evans et al., 2011,

Kotrba et al., 2014, Genevcius et al., 2017, Yao et al., 2019]). Mismatches in genital fit have been shown to cause harm in mating, and be detrimental to fitness ([Masly and Kamimura, 2014, Barnard et al., 2017, Tanaka et al., 2018]). We therefore expect that the evolution of the genitalia in one sex would be under some degree of constraint due to the need for functional compatibility with the genitalia of the other sex.

With regard to genetic integration between the genitalia of males and females, we currently know much less compared to our knowledge of functional interactions. Pleiotropy could theorethically contribute to correlations between male and female genitalia if these traits were part of the same pleiotropic "clusters." It has in fact been noted that one possible way that male and female genitalia could maintain a good fit would be to have linked genetics associated with beneficial covariation ([Brennan and Prum, 2015, Simmons and Garcia-Gonzalez, 2011]). This state is predicted by theory and is sometimes seen for other sexually selected traits (e.g. [Marcillac et al., 2005, Kronforst et al., 2006]). Yet because we know very little about the developmental genetics of genitalia generally ([Shapiro and Porter, 1989]), and of female genitalia especially (as the study of female genitalia has historically been underemphasized (see [Ah-King et al., 2014, Méndez and Córdoba-Aguilar, 2004])), it is difficult to assess the degree to which correlations observed across species or within populations might be influenced by pleiotropy.

A small but growing body of work has provided some insight in this area. Several studies have found evidence for genetic covariance between male and female genital traits (e.g. [Simmons and Garcia-Gonzalez, 2011, André et al., 2020, Genevcius et al., 2020]). Two studies, both conducted in the same system, found evidence of one pair of weakly-linked QTL between a male and female genital measure ([Sasabe et al., 2010, Fujisawa et al., 2019]. In *Drosophila*, it has been demonstrated that there is a higher degree of shared genetic programming between males and females early on in the development of the genital segments ([Gorfinkiel et al., 1999, Keisman and Baker, 2001, Sánchez and Guerrero, 2001]), but that the male and female genital discs diverge more in their genetics later ([Chatterjee et al., 2011]). Many more studies such as these are needed in order to gain a more general picture of how genetic linkage might influence the coevolutionary trajectories of genitalia.

Finally, it is interesting to consider whether there might be differences between how

pleiotropy influences correlated *quantitative* changes in ancient structures, such as correlated shifts in the length of the adeagus and spermatheca, versus how shared genetics might manifest in correlations between males and females following *qualitative* changes in developmental genetics that might occur when novel genital features evolve. Genital evolution often involves increases in "complexity" ([Brennan et al., 2007, Kuntner et al., 2009, Rowe and Arnqvist, 2012]) involving alterations such as the emergence of novel shapes or addition of novel structures on the genitalia. If such qualitative changes to genitalia are pleiotropic, such that the genitalia of both sexes are affected, the shorter evolutionary period of differentiation could mean that the evolutionary trajectories of the resultant novel male and female strucures would be more strongly influenced by that linkage than ancient structures, which might have long since drifted or been driven by selection into separate developmental "modules."

For this reason, we chose to investigate the possibility that shared genetic architecture contributed to correlated divergence of a pair of recently evolved male and female genital structures in the *D. melanogaster* subgroup. Males in this subgroup possess a cuticular outgrowth on their genitalia called the posterior lobe. This structure is newly-evolved, present in only the four most closely-related species of this clade (**Figure 5, A**). Females in this group also possess a newly-evolved genital structure, the oviscapt pouch ([Yassin and Orgogozo, 2013]) (**Figure 5, B**). Like the posterior lobe, the oviscapt pouch is absent from species outside the D. melanogaster clade. The interaction of the posterior lobe with the female ovipositor during copulation has been reported previously ([Jagadeeshan and Singh, 2006, Mattei et al., 2015]) but it is not known whether the oviscapt pouch contacts the posterior lobe during genital coupling. Yassin et al (2013) noted that a positive size correlation exists between the posterior lobe and the oviscapt pouch in the three species where both structures were measured (although the trend was not significant given the limited number of species). A number of studies have investigated the possible functions of the posterior lobe ([LeVasseur-Viens et al., 2015, Frazee and Masly, 2015, Frazee et al., 2020]). Nothing is currently known about the oviscapt pouch beyond its size relationship to the posterior lobe ([Yassin and Orgogozo, 2013]).

To look for evidence of genetic linkage contributing to the correlated structure sizes

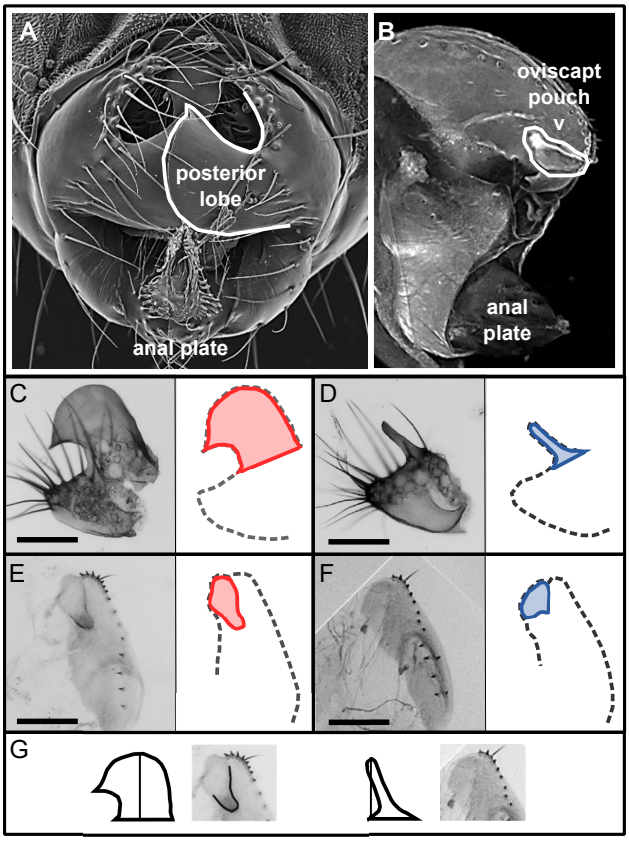

Figure 5: **A.** Electron micrograph of *D. simulans* male genitalia, showing the location of the posterior lobe and anal plate. **B.** Confocal image of female ovipostor from *D. sechellia*, another species in this group, showing the location of the oviscapt pouch and anal plate. **C.** *D. simulans* posterior lobe cuticle preparation (Left), and tracing (Right). **D.** *D. mauritiana* posterior lobe cuticle preparation (Left), and tracing (Right). **E.** *D. simulans* oviscapt pouch cuticle preparation (Left), and tracing (Right). **F.** *D. mauritiana* oviscapt pouch cuticle preparation (Left), and tracing (Right). **G.** Outlines of posterior lobes with vertical line to indicate how height was measured, and oviscapt pouches with a black line to indicate the oviscapt ridge. *D. simulans* posterior lobe and oviscapt on left, *D. mauritiana* posterior lobe and oviscapt on right. It can be seen that the oviscapt ridge is absent in the *D. mauritiana* sample

across species, we performed a QTL analysis of the size and shape of the posterior lobe and oviscapt pouch of *D. simulans* and *D. mauritiana*, two sister species that have the largest (*D. simulans*, **Figure 5, C,E**) and smallest (*D. mauritiana*, **Figure 5, D,F**) posterior lobes and oviscapt pouches in this group, respectively. We looked for colocalizing QTL associated with the phenotypic difference in size and shape for both structures, which might indicate that linked or shared genetics contributed to the correlated divergence observed across these two species. In addition, for a set of genes that occured within shared QTL and which were known from previous work to be deployed during the development of the posterior lobe [Glassford et al., 2015], we investigated whether these genes were also active in the developing female genitalia, and whether the posterior lobe regulatory elements for these genes also drive expression in or around the developming oviscapt pouch. Shared regulatory information for these developmental genes suggests that pleiotropy could be implicated in QTL colocalization, as opposed to, or in addition to, selection for closely physically linked but distinct genetic programming. Our results highlight the likely role of pleiotropy in generating correlations in these genital structures.

2.2 METHODS

2.2.1 Experimental animals and phenotyping

Crossing scheme: For the QTL analysis, the parental *D. simulans* strain used was Dsim [MD15], an isofemale D. simulans line first collected in 2002 [Rogers et al., 2014]. The parental *D. mauritiana* strain was the inbred line Dmau w [12] (National Drosophila Species Stock Center 14021-0241.60). All flies for this experiment were reared in narrow (8-dram) vials containing standard fly media. Animals were housed in a 25 deg C incubator. Parental cross virgin females from Dsim [MD15] were collected and aged at least 3 days to sexual maturity. These females were crossed to males of Dmau w[12] in vials containing eight virgin *D. simulans* females and five *D. mauritiana* males. F1 females from this cross were backcrossed to males of both parental lines as follows: eight F1 females and five *D. simulans*

males, eight F1 females and five *D. mauritiana* males. Twelve males and twelve females were taken from each backcrosss vial, unless the vial did not produce 12 of each, in which case fewer individuals were taken.

Dissection and Phenotyping: Animals were anesthetized under CO2. The genitalia and 1 femur were removed with surgical dissection scissors and placed on a glass slide in halocarbon oil 27 (Sigma). The remainder of the carcass was transferred to a 96-well cell culture plate (Corning), and the position of the individual fly was noted directly on the glass slide containing that individual's genitalia. Rows of each plate alternated between males and females. Plates were sealed with a rubber seal and stored at -20C until needed for extraction and genotyping. Genitalia were dissected immediately using tungsten wires. For males, posterior lobes were left attached to the lateral plate and laid flat. For females, the two halves of the ovipositors were separated almost entirely but left connected slightly at the distal tip of the ovipositor, so that the samples could be flattened (butterflied) but kept together on the slide. Femurs from each individual were mounted on the same slide. Slides were finished with glass cover slips and stored in slide boxes at room temperature for imaging. Imaging was performed using a Leica DM2000, with a 10x objective for all structures except the femurs, which were captured using a 5X objective. All images were taken at a resolution of 2560x1920. Images were processed using ImageJ ([Rueden et al., 2017]) or FIJI ([Schindelin et al., 2012]). From each individual, only one posterior lobe or oviscapt pouch was measured. For the posterior lobe, the structure area and height were measured as indicated in **Figure 5, G**. For the oviscapt pouch, the structure area was measured. A binary oviscapt pouch phenotype was also noted indicating the presence or absence of a visible ridge of cuticle along the proximal edge of the oviscapt pouch (**Figure 5, G, left center, black line**), as this is one of the most obvious distinguishing features between the oviscapt pouches of these two species. We chose to use the binary phenotype because the linear measure of this feature was quite noisy and likely depended heavily on small variation in mounting from sample to sample. Area and height measures were checked for normality, and transformed where necessary using the Box Cox transformation (**Table 6, appendix**). As there is some debate about whether transformations improve the ability to detect QTL to a greater extent than they reduce detection sensitivity, the transformation was only performed in cases that failed

normality at a level of $p < 0.0001$, which was the case only for the area measure for the *D. mauritiana* backcross direction. Individuals were not included in the analysis if poor sample quality (e.g. the sample was chipped or broken) prevented proper phenotyping.

Shape analysis As the posterior lobe and oviscapt pouch have few true landmarks, we chose to assess shape differences using elliptical Fourier analysis. Elliptical Fourier analysis is a curve-fitting technique that approximates the outline of a structure using the weighted sum of multiple sine and cosine waves [Kuhl and Giardina, 1982]. Each posterior lobe and oviscapt pouch was first outlined from images in ImageJ. For each structure, the outlines from each individual were imported into R, and the package Momocs [Bonhomme et al., 2014] was used to generate outline objects with 100 evenly distributed pseudolandmarks. The outlines were aligned and their size normalized using procrustes analysis with the package "shapes"[Dryden, 2019]. The alligned and normalized outlines were used to generate Fourier coefficients by elliptical Fourier analysis, with the number of harmonics selected corresponding to the number needed to acheive 99% harmonic power. This was nine harmonics for males and seven harmonics for females. The normalization step was not performed for the elliptical Fourier analysis, as the Procrustes analysis had already been employed. A principal components analysis was then performed using the Fourier coefficients as input, to generate a small number of parameters describing the majority of shape variation. The total number of principal components produced by the analysis was 36 for males and 28 for females.

Investigation of allometry Measurements were assessed for a correlation with body size using the femur measurement (**Table 6, appendix**). Where a correlation was found, the residuals from the body size regression were used in place of the raw measures to mitigate the detection of QTL associated only with general body size differences.

Corrections for normality, adjustments to account for differences in body size, and statistical analyses investigating differences between parental and backcross directions were performed in R using base functions ([R Core Team, 2017]).

2.2.2 Genotyping

Multiplexed Shotgun Sequencing (MSG) ([Andolfatto et al., 2011]) was used to carry out genome-wide ancestry assignment of backcross progeny. Genomic DNA was extracted from the remainder of the carcasses after the removal of genitalia and femurs using the Quick-DNA (Zymo Research, Irvine, California, USA, catalog No. D3010) extraction kit. A Tn5-based method ([Picelli et al., 2014]) was used to generate Illumina libraries. Sequencing of libraries was performed with 150bp single-end reads on an Illumina HiSeq 2500 at the Princeton University Genomics Core Facility. Median coverage was 221,839 reads per individual. Individuals with $<$ 20,000 reads were excluded. In the *D. simulans* run, there were 1,168,177 markers, and in the *D. mauritiana* run, there were 1,158,828 markers

Reads were processed using MSG with the following parameters:

D. mauritiana backcross: Priors: 0.5, 0.5, 0 (homozygous for *D. mauritiana*, heterozygous, homozygous for *D. simulans*) deltapar1= 0.006 deltapar2 = 0.008 recRate = 3 rfac = 0.1

D simulans backcross: Priors: 0, 0.5, 0.5 (homozygous for *D. mauritiana*, heterozygous, homozygous for *D. simulans*) deltapar1= 0.025 deltapar2 = 0.007 recRate = 3 rfac = 0.1 Prior probability that the allelic state of the parental reference is incorrect was tuned using 12 individuals each of the parental *D. simulans* and *D. muritiana* lines.

2.2.3 QTL mapping

QTL analyses were performed using the package rqtl ([W. et al., 2003]), with the rqtl implimentation of MQM ([Arends et al., 2010]). Because previous work had indicated that the divergence in posterior lobe morphology between these two species was likely highly polygenic ([Liu et al., 1996, Zeng et al., 2000]), we chose to employ the multiple QTL mapping (MQM) tecnique for our analysis. MQM offers higher statistical power for QTL detection, especially for complex traits that may have many QTL close to each other on the chromosome ([Jansen, 1994, Arends et al., 2010]). For MQM, the computational limit is 600 markers, and so the marker set was thinned to 120 markers per chromosome arm. Our sample sizes for this analysis were moderate such that this number of markers still meets or

exceeds the expected possible resolution of QTL, and therefore should not greatly impact our ability to detect QTL. We did not include the fourth chromosome in our analysis because of the small size and small number of known functional genes on that chromosome in these species. We also did not include the Y chromosome, as the direction of our backcross only produces males with unrecombined Y chromosomes from a single parent. We first used 492 *D. simulans* backcross male individuals and 508 *D. simulans* backcross female individuals to generate a genetic map for our analysis of the *D. simulans* backcross direction. Similarly for the *D. mauritiana* backcross direction, 504 *D. mauritiana* backcross male individuals and 506 *D. mauritiana* backcross female individuals were used to generate a genetic map. Sexes were used jointly to generate the genetic maps for each backcross so that the identified QTL locations could be directly compared between males and females within a backcross direction. MQM is a three step process. First, data is "augmented", to fill in any missing genotype information. The genotyping of individuals in our data was exceptionally good, with very little missing genotype information (<1%). Therefore, we chose the most minimal augmentation in which only the most likely genotype for an individual was used (more complex approaches can be employed in which multiple pseudo-individuals with a range of likely genotypes are included). After augmentation, a set of markers is selected via multiple regression and backward elimination to be used as "cofactors". Cofactors essentially inform the interval mapping step (the third step of MQM) of regions that may be associated with QTL, so that as a QTL is moved along the chromosome during interval mapping the model can take into account these regions of interest and avoid spuriously eliminating QTL that may be close together, while also enabling the elimination of cofactors that do not improve the model. The list of QTL and positions produced by the mqm interval mapping scan was used to fit a final multiple QTL model using rqtl multiple QTL model functions. We ran 1000 permutations to determine a threshold of significance. Any QTL that were below this significance threshold were removed, until a model was acheived that consisted of only QTL above our calculated threshold. When two QTL were less than 1cM apart, we also removed whichever QTL had the lower lod score to retain only one QTL in that region, as a distance of less than 1cM is beyond a resolution we could acheive with our methods. The cofactor selection is done in an automated step as part of the MQM for rqtl pipeline, however, the

user chooses how many cofactors to start with for backward elimination. We chose to start backward elimination with the number of cofactors that would correspond to approximately 1 cofactor/cM, as it is unlikly that we could have confidence resolving QTL any more closely related than this. For the *D. simulans* backcross direction, this number was 209. For the *D. mauritiana* backcross this number was 236. Automated cofactor selection is not entirely deterministic, therefore, there was some varaition in the cofactor selection that caused some variation in our final result. To mitigate this problem, we ran 1000 iterations of the cofactor selection and interval mapping steps, recording the markers selected for the locations of QTL in each run. For every genetic position identified as a QTL in any run, we determined the marker that was most commonly associated with that QTL. We then took all of these markers and used them to generate a single set of cofactors, which we used for the final mqm scan. MQM for rqtl is not currently capable of treating sex chromosomes differently than somatic chromsomes. To lend ourselves additional confidence that this did not greatly impact our ability to detect QTL on the X chromsome, we compared our MQM results to two other mapping techniques. First, we performed genetic mapping with a single QTL model that maps only one QTL per chromosome using the Haley-Knott regression method ([Haley and Knott, 1992]). Because our phenotypes have more than one QTL per chromosome, this model suits our data poorly. Still, we used these scans to roughly corroborate our findings, especially on the X chromosome where we found only a few QTL even by the MQM technique. We performed 1,000 permutations to determine the statistically-signficant lod threshold for the analysis ([Churchill and Doerge, 1994]). We also performed composite interval mapping (CIM; [Zeng, 1994]), again using the Haley-Knott regression method ([Haley and Knott, 1992]) to perform the forward and backward selection process to arrive at a multiple QTL model. We again used 1,000 permutations to determine statistcal significance ([Churchill and Doerge, 1994]). Although the single QTL model and CIM can accomodate far more markers than MQM, we thinned our marker sets using a custom Python ([Van Rossum and Drake Jr, 1995]) script in order to save computational time. These marker sets consisted of 14,724 markers for the *D. simulans* backcross and 14,411 markers for the *D. mauritiana* backcross.

2.2.4 Simulation study

We designed a custom perl script 'QTL-overlap-sim' to assess the likelihood of finding the observed overlaps of male and female QTL intervals by random chance. We used the data from the *D. simulans* backcross for our analysis. The script operates in two steps, first calculating the degree of overlap between two lists of intervals in an exhaustive interval-by-interval comparison. For this analysis, a non-redundant/non-overlapping set of male and female intervals were collated. In the second step, randomized, but non-overlapping intervals are produced, and the degree of overlap among this randomized set is calculated for each simulation. The number of simulations which equal or exceed the measured data are calculated and reported.

2.2.5 Investigation of previously known posterior lobe genes

As the posterior lobe and oviscapt pouch arose concomitantly in this lineage, we wanted to know whether some of these same genes implicated in the origination of the posterior lobe were also active in female genitalia during the development of the oviscapt pouch. We viusalized the activity of these genes in the female genitalia of *D. melanogaster*. For each of six genes, we either performed immunostaining to viusalize gene products (*eyes absent* (*eya*), *crumbs* (*crb*)) or *in situ* hybridization to visualize mRNA transcripts (*unpaired* (*upd*), *Pox-neuro* (*Poxn*), *Rho guanine nucleotide exchange factor at 64C* (*Gef64C*), *Cadherin 86C* (*Cad86C*)). We then performed GFP reporter assays using previously identified posterior lobe enhancer regions ([Glassford et al., 2015]) to determine whether these known posterior lobe regulatory elements also drive gene expression in the developing female genitalia.

For all experiments, I collected white prepupa from a *D. melanogaster* line that is mutant for the genes *yellow* and *white* (Bloomington Stock Center No. 1495). Prepuae were sexed and incubated at 25 degrees C for between 48 and 60 hours. The developmental timing of the female genital structures is slightly longer than for the male structures, and the window that we used for developmental aging reflects that shift. Pupae were dissected in cold PBS by cutting the pupa in half with surgical scissors and carefully pulling the pupal membrane out of the pupal case. The internal organs and fat bodies were then flushed out

of the pupal membrane with gentle pipetting. Samples were fixed in a solution of PBS with 0.3% Triton-X and 4% paraformadehyde at room temperature for about 30 min. For male genital samples in the previous work, the fixation solution contained only 0.1% Triton-X ([Glassford et al., 2015]. However, we found that it was necessary to use a slightly higher detergent concentration for our samples due to the beginning of cuticle formation at these slightly later time points. If a sample was being prepared for *in situ* hybridization, the samples were washed twice with methanol and then stored at -20 degrees C until needed. For immunostaining or reporter analysis, samples were washed three times in a PBS solution containing 0.1% Triton-X for 5-10 minutes and then stored at 4 deg until used for staining or imaging within 24 hours. For immunostaining, female pupal genitalia were were incubated overnight at 4 deg C with primary antibodies. All antibody diutions and washes used a PBS solution with 0.1% Trition. After staining with the primary antibody, the samples were washed five times and the diluted secondary antibody was added. Samples were again incubated overnight at 4 deg C. After staining with the secondary antibody, the samples were washed five times. The samples were then washed for 5 min in a solution of 50% PBS with 0.1% Triton, and 50% glycerol, before being transferred to a solution of 80% glycerol and 0.1 M Tris-HCL and then mounted on glass slides. Primary antibodies used were mouse anti-*Crb* 1:25 dilution (Tepass and Knust, 1993) and mouse anti-*Eya* 1:100 diluation (Bonini et al., 1997). The secondary antibody was donkey anti-mouse Alexa 488 (Molecular Probes), 1:200 dilution. *In situ* hybridizations were performed using an InsituPro VSi robot (Intavis Bioanalytical Instruments) as described in Vincent et al 2019. The alkaline phosphatase stain was developed in a solution of nitro-blue tetrazolium chloride and 5-bromo-4-chloro-30-indolyphosphate p-toluidine salt (NBT/BCIP) (Promega). Samples were then washed in PBS solution containing 0.1% Triton-X, transferred to 80% glycerol with 0.1 M Tris-HCL, and mounted on glass slides. Sequences of probes used for *upd*, *poxn*, *cad86c* and *gef64c* are listed in the appendix (**Table 7**).

GFP reporter assays were conducted using the *D. melanogaster* posterior lobe enhancers of *cad86c*, *crb*, *eya*, *gef64c*, and *poxn*, which are described in Glassford et al 2015 ("*Cad86C* enhancer", "*crb* enhancer", "*eya* enhancer"', "*Gef64C* Intron + Exon 1", and "*Poxn* posterior lobe enhancer", respectively, injected into the attP2 landing site). For the gene *upd*,

the enhancer region was identified separately by Sarah J. Smith (Unpublished data. *D. melanogaster* forward primer: TTCCGggcgcgccCGTATCAGTTTGCAATGGGTGGTG, reverse primer:TTGCCcctgcaggCACCACCCATTGCAAACTGATACG). Timed female pupae were dissected as described above, then immediately washed for 5 min in a solution of 50% PBS with 0.1% Triton, and 50% glycerol, before being transferred to a solution of 80% glycerol and 0.1 M Tris-HCL and then mounted on glass slides for imaging.

Immunostained and fluorescent reporter samples were imaged on an Olympus Fluoview 1000 confocal microscope with a 20x objective. *In situ* samples were imaged on a Leica DM2000 at 20x.

2.3 RESULTS

2.3.1 Phenotype and shape analysis

As expected, the size of the posterior lobe and oviscapt pouch are both significantly different by backcross direction, and also differed from the parental strains (**Table 1, Figure 6**), although there is more overlap between the two directions and the parental strains in the females. This may be a result of the manner in which oviscapt pouches were phenotyped for this experiment. Because in *D. mauritiana* the oviscapt pouch is extremely reduced, it was difficult to define the distal edge of the structure (in **Figure 5, E** and **F**, the left hand side of the structure) consistently. In order to be comparable across all samples, the distal boundary of the female structure was therefore defined as the dorsal ovipositor edge, as this was a relaible boundary. However, it is almost certain that this underestimated the difference in structure size between the backcross directions, as in some cases of the *D. mauritiana* backcross the oviscapt pouch was so shallow that it is somewhat inaccurate to claim that it extended all the way to the edge of the ovipostor. Nevertheless, the oviscapt pouches of the two backcross directions were still measurably distinct in size, indicating that variance in structure size was captured by our methods. Posterior lobe height was also significantly different between backcross directions. The binary metric indicating presence

or absence of a ridge on the proximal edge of the oviscapt pouch was found to be a useful metric for the purposes of the QTL analysis only in the *D. mauritiana* backcross direction. This is because only four of the 508 *D. simulans* oviscapt pouch samples did not possess this feature, so it lacked the power to be informative, whereas the ridge was absent in roughly one quarter (123/506) of the *D. mauritiana* samples, and thus provided some recombination breakpoint information that could be used for the QTL analysis.

Table 1: Posterior lobe and Oviscapt pouch morphology across parental and backcross strains.

	Posterior lobe					Oviscapt pouch		
Strain	N	Area (X 10^{-3} mm^2)	sd (X 10^{-3} mm^2)	Height (mm)	sd (mm)	N	Area (X 10^{-3} mm^2)	sd (X 10^{-3} mm^2)
D. mauritiana (mau12)	12	1.71961	0.16638	0.07240	0.00235	12	1.67466	0.13545
D. mauritiana backcross	504	4.02528	1.14619	0.08948	0.00657	506	2.07018	0.31961
D. simulans (simMD15)	17	11.57190	0.54127	0.11516	0.00538	17	3.00166	0.18704
D. simulans backcross	492	8.86960	1.30634	0.10744	0.00581	508	2.71823	0.32798

For the posterior lobe, the first six principal components each accounted for greater than 1% (each) of the variation in structure shape. The cumulative percentage of variation explained for the first six PCs was 97.6% (**Table 2**). The first principal component accounted for the vast majority of this, explaining 79.5% of the variation. For the oviscapt pouch also, the first six principal components each accounted for greater than 1% of the shape variation, with a cumulative percentage of 95.6% for these first six components. The first principal component accounted for 75.8% of the varaition (**Table 2**).

When comparing the principal component scores across the backcross groups for the posterior lobe, only PC1 and PC3 are significantly different by backcross direction (**Figure 7, A**). **Figure 8, A** shows the results of the separation of the two backcross directions in shape space according to these two principal components. For the purposes of our QTL analysis, we chose to investigate only PC1 to search for loci of divergence, as this compontent was highly significantly different between groups, and explained the vast majority of the variance in shape. **Figure 9** shows outlines of individuals from the analysis to demonstrate changes to the shape of the posterior lobe associated with changes to PC1.

The results were similar for the oviscapt pouch, except that PC5 was also significantly

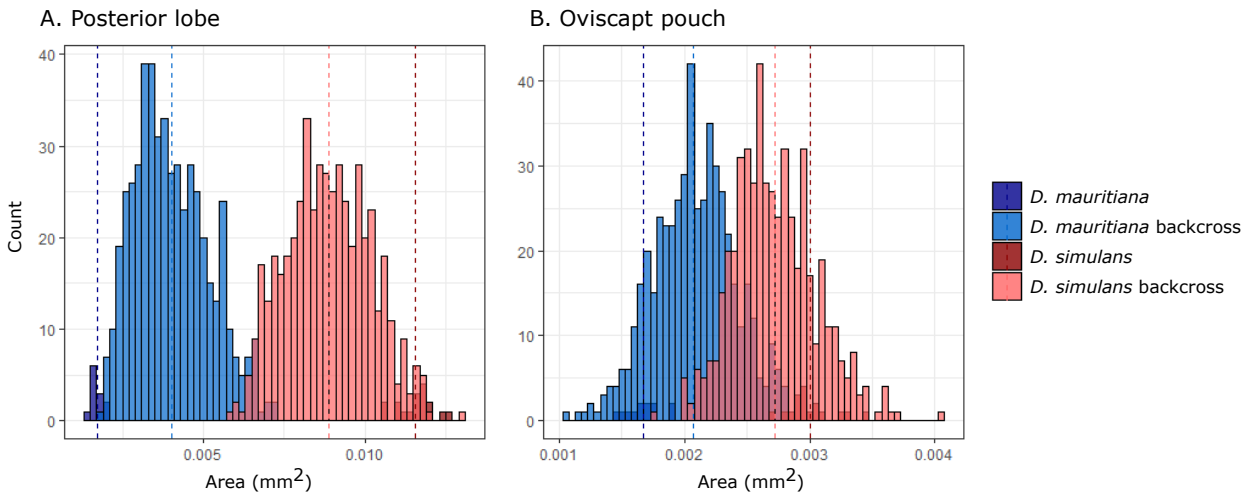

Figure 6: Histograms showing the range and frequency of sizes for the male posterior lobe (**A**), and female oviscapt pouch (**B**), across parental lines and the backcross directions. Dashed vertical lines are the group means. The spread of size has much more overlap for the female structure

Table 2: Shape analysis of male and female genital structures. Percent of variance in shape explained by the first 6 principal components of shape for the posterior lobe and oviscapt pouch.

	Posterior lobe			Oviscapt pouch		
Shape variable	Variance explained (%)	Standard deviation	Cumulative var. explained (%)	Variance explained (%)	Standard deviation	Cumulative var. explained (%)
PC1	79.5	1.83	79.5	75.8	1.60	75.8
PC2	8.6	0.60	88.1	9.9	0.58	85.7
PC3	3.9	0.41	92.0	3.9	0.36	89.6
PC4	2.9	0.35	94.9	2.7	0.30	92.3
PC5	1.6	0.26	96.5	1.8	0.25	94.1
PC6	1.1	0.21	97.6	1.5	0.22	95.6

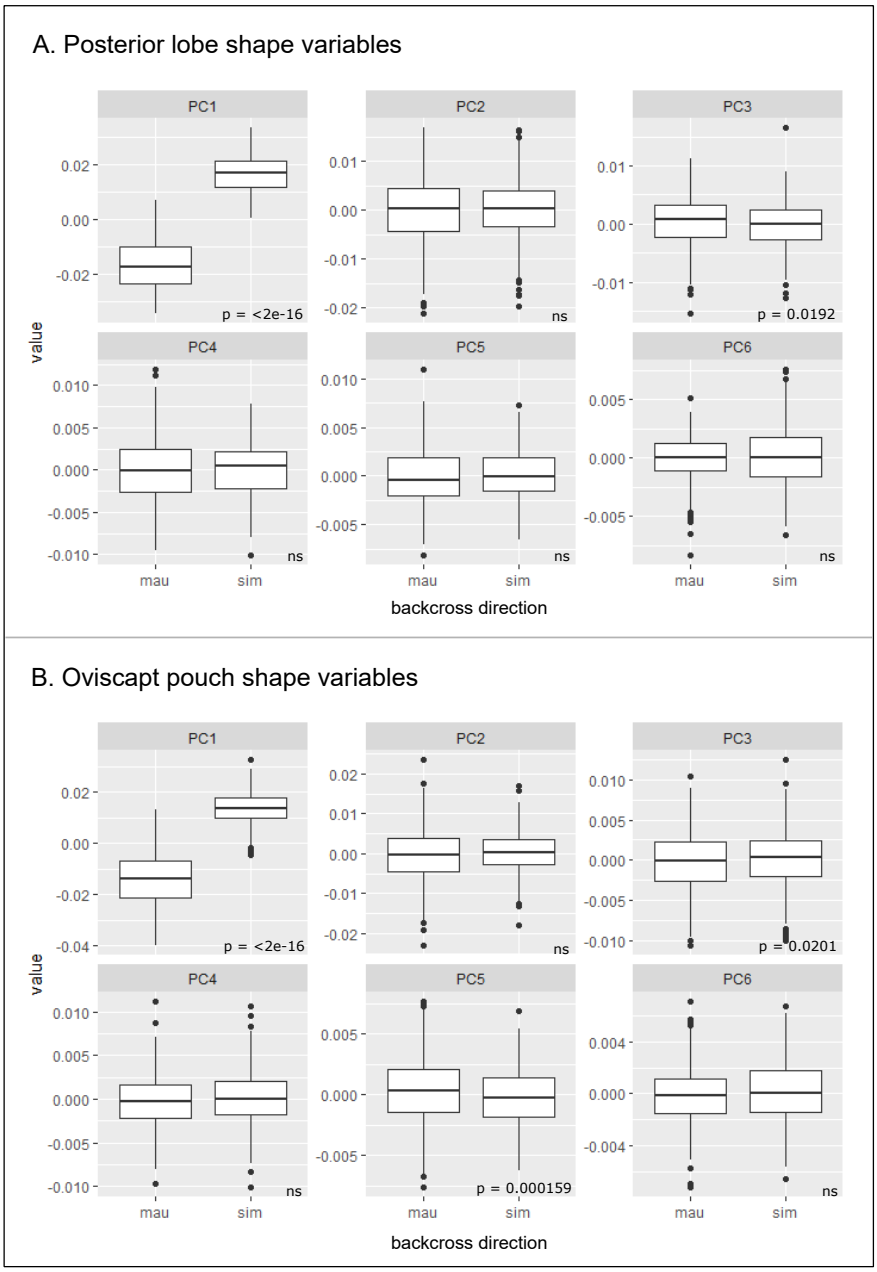

Figure 7: Shape variables by backcross direction for male and female genital structures. Boxplots comparing the first six principal components of shape variation by backcross direction for the male posterior lobe (**A**), and female oviscapt pouch (**B**). The p-value is shown where significant.

different between backcross groups (**Figure 7, B**). **Figure 8, B** shows the results of the separation of the two backcross directions in shape space according to PC1 and PC3. As with the posterior lobe, due to the small amount of variance explained by PC3 and PC5, we elected to limit our investigation to PC1 for the QTL analysis. **Figure 10** shows outlines of individuals from the analysis to demonstrate changes to the shape of the oviscapt pouch associated with changes to PC1.

2.3.2 The posterior lobe and oviscapt pouch are both highly polygenic traits, with male phenotypes generally having more QTL than female traits

We detected a large number of QTL associated with the variation in posterior lobe size and shape in our backcross populations with our Multiple QTL modeling (MQM, **Table 3**; **Figure 11**; **Figure 12**). This indicates that this trait is highly polygenic in its differentiation between these two species. Our result is consistent with a previous analysis, which identified a total of 19 putatitve QTL for the posterior lobe shape and size differences between these two species ([Zeng et al., 2000]). In the *D. simulans* backcross direction we detected twelve QTL associated with lobe area, ten associated with posterior lobe height, and nine associated with PC1. All but two of the QTL associated with height overlap with area QTL, which is expected given that posterior lobe height should be captured in the area metric (FIgure S1). The good correspondence between area and height QTL therefore gives us additional confidence that our models are accurately identifiying QTL. We obtained a similar result for the *D. mauritiana* backcross direction, finding fourteen QTL for area, ten for height, and thirteen QTL for PC1. All but one height QTL overlap with area QTL in the *D. mauritiana* direction (Supplementary figure). Many of the QTL for shape parameters also colocalize with the size measures, which may indicate shared underlying developmental mechanisms. While each QTL explains only a small percentage of the variation (**Table 4**), the overall QTL models for the male phenotypes generally explain a large percentage of the trait variation (average of 74.6%; **Table 3**). Because the two backcross directions do not have the same genetic maps, we cannot directly compare the genetic positons of QTL across the two analyses. However, we can qualitatively assess that the results of the two backcross

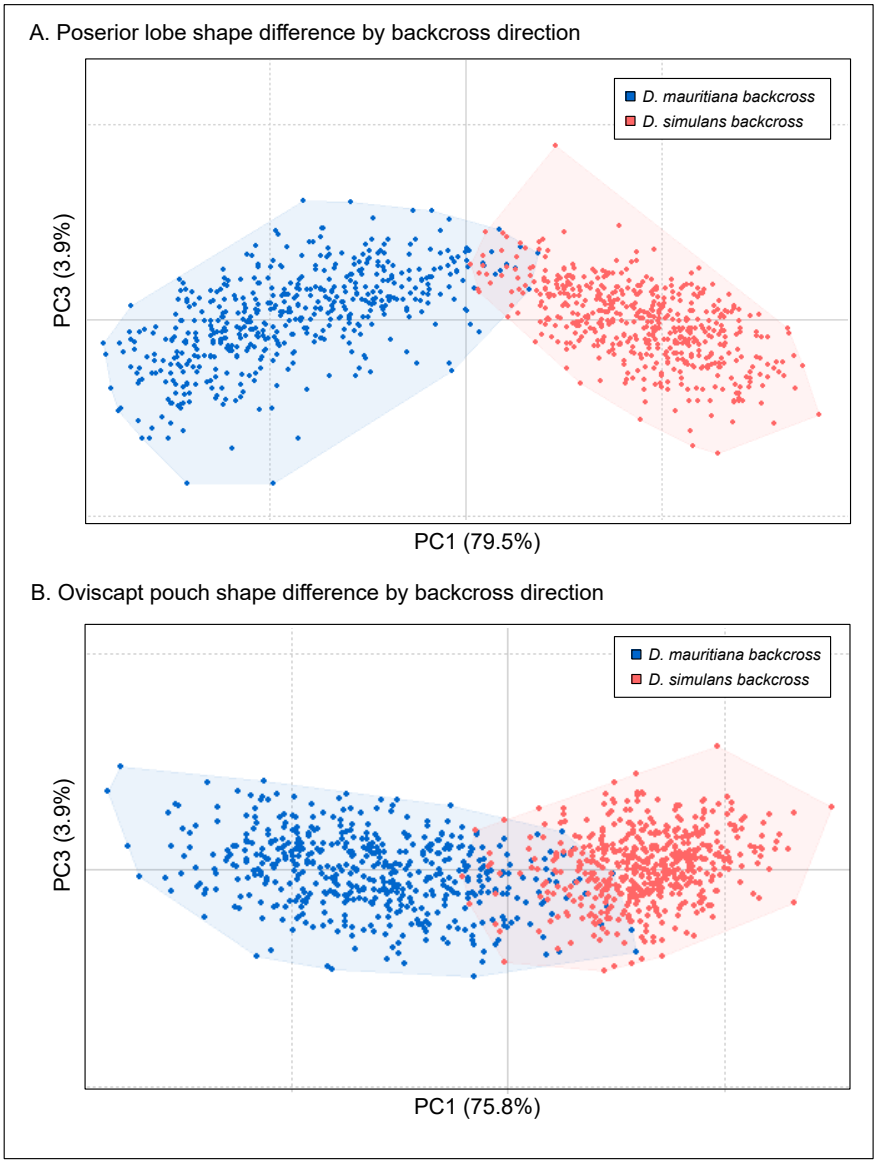

Figure 8: Separation of backcross progeny in shape space, color coded by backcross direction. For both male posterior lobes (**A**) and female oviscapt pouches (**B**). PC1 and PC3 significantly differ between groups. Both the male and female structures are separated best by PC1, which explains the majority of the shape variance in both cases.

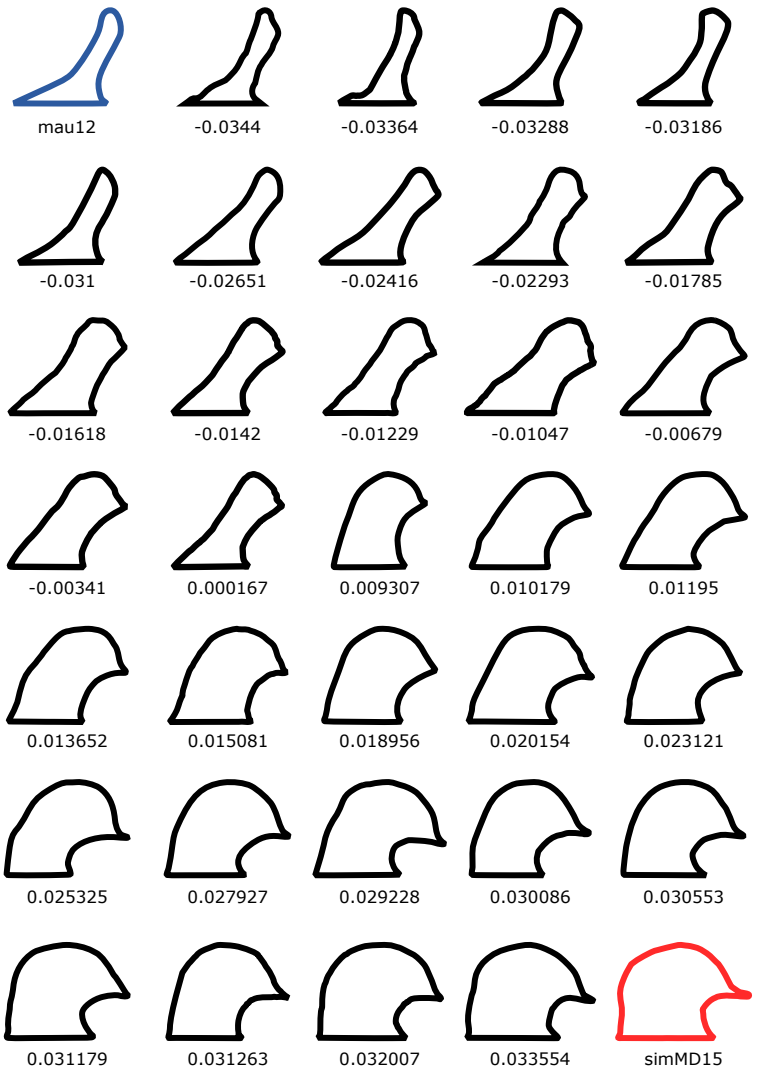

Figure 9: Oulines of posterior lobes from backcross individuals showing the association of changes in shape with variation in PC1, score indicated below each tracing. The first (blue) and last (red) outlines are example individuals from the parental *D. mauritiana* and *D. simulans* lines, respectively.

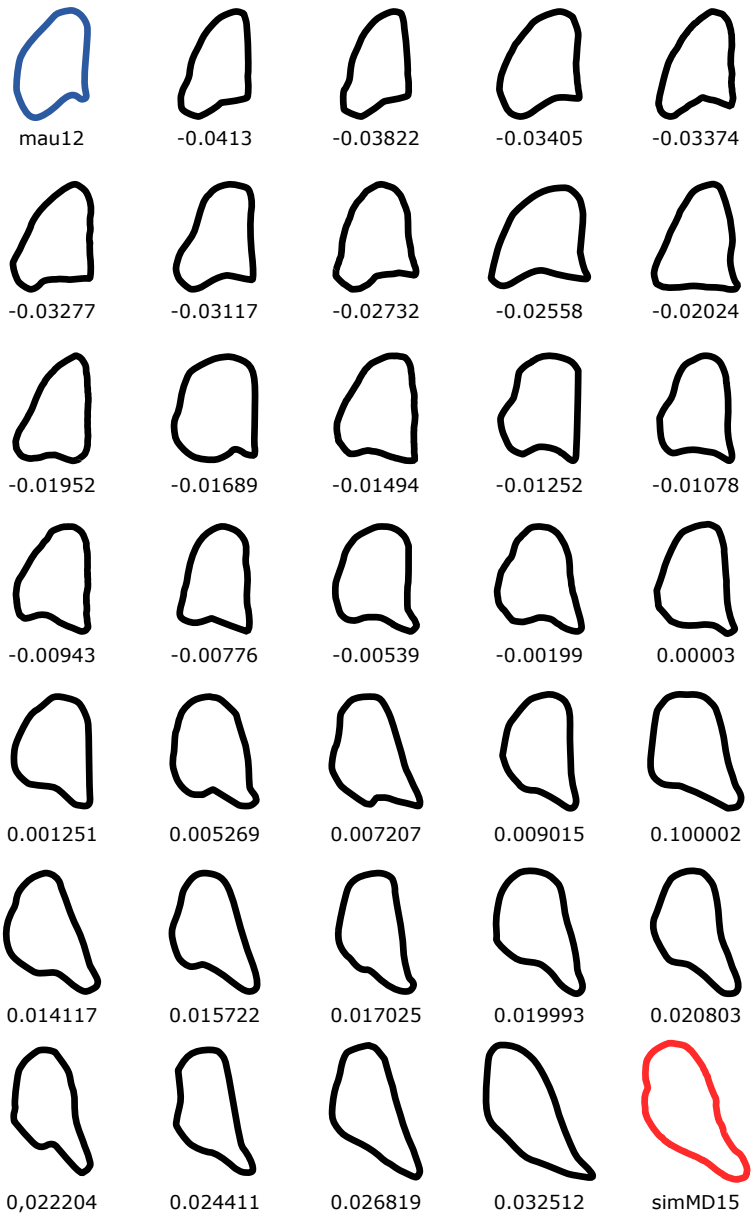

Figure 10: Oulines of oviscapt pouches from backcross individuals showing the association of changes in in shape with variation in PC1, score indicated below each tracing. The first (blue) and last (red) outlines are example individuals from the parental *D. mauritiana* and *D. simulans* lines, respectively.

directions corroborate quite well.

We found that our single QTL model scan, which searches for only one QTL per chromosome, and our composite interval mapping (CIM, a distinct multi-QTL interval mapping technique) results also roughly corroborated our MQM data, given the limitations of those two methods (**Figure 15, Figure 16, appendix**). One clear conclusion is that the single QTL models are quite inaccurate in determining QTL location when multiple QTL occur close together. For instance, the single QTL scan for posterior lobe area shows a region of high LOD in the same area as those we found with our MQM scan. However, the peak maximum in the single QTL scan falls directly in between the two QTL we identified with MQM. This indicates the caution that must be exercised when using QTL scans to identify regions if the goal is to look for candidate genes within those intervals. In this case, use of a poorly representative single QTL model could have resulted in an entirely fruitless search for a causitive locus.

Fewer QTL were identified for the female traits. In the *D. simulans* backcross, seven QTL were detected for the oviscapt pouch area, and eight were detected for PC1. In the *D. maurtiana* direction, six QTL were detected for the oviscapt pouch area, six were detected for PC1, and six were found for the oviscapt ridge binary phenotype. The QTL models for the female phenotypes also explain less of the variation in phenotype (average of 26.0%; **Table 3**) than was captured for the QTL models associated with the male phenotypes. This may reflect a genuine disparity in the number of loci involved in the differentiation of these two structures, or it may in part be due to greater difficulty in phenotyping the oviscapt pouch. Again, however, the QTL identified for oviscapt pouch phenotypes for the two backcross directions seem to corroborate each other fairly well, although less so for shape. For PC1, several QTL on the X chromosome were identified in only one direction and in no other female parameter in either direction. This does not necessarily mean that these are not true QTL, but suggests that their effects may be more strongly influenced by dominance or epistasis. The ridge phenotype scan was predominated by the first large QTL on chromosome 3 (**Figure 14, appendix**), which coincided with a large QTL found for oviscapt pouch area.

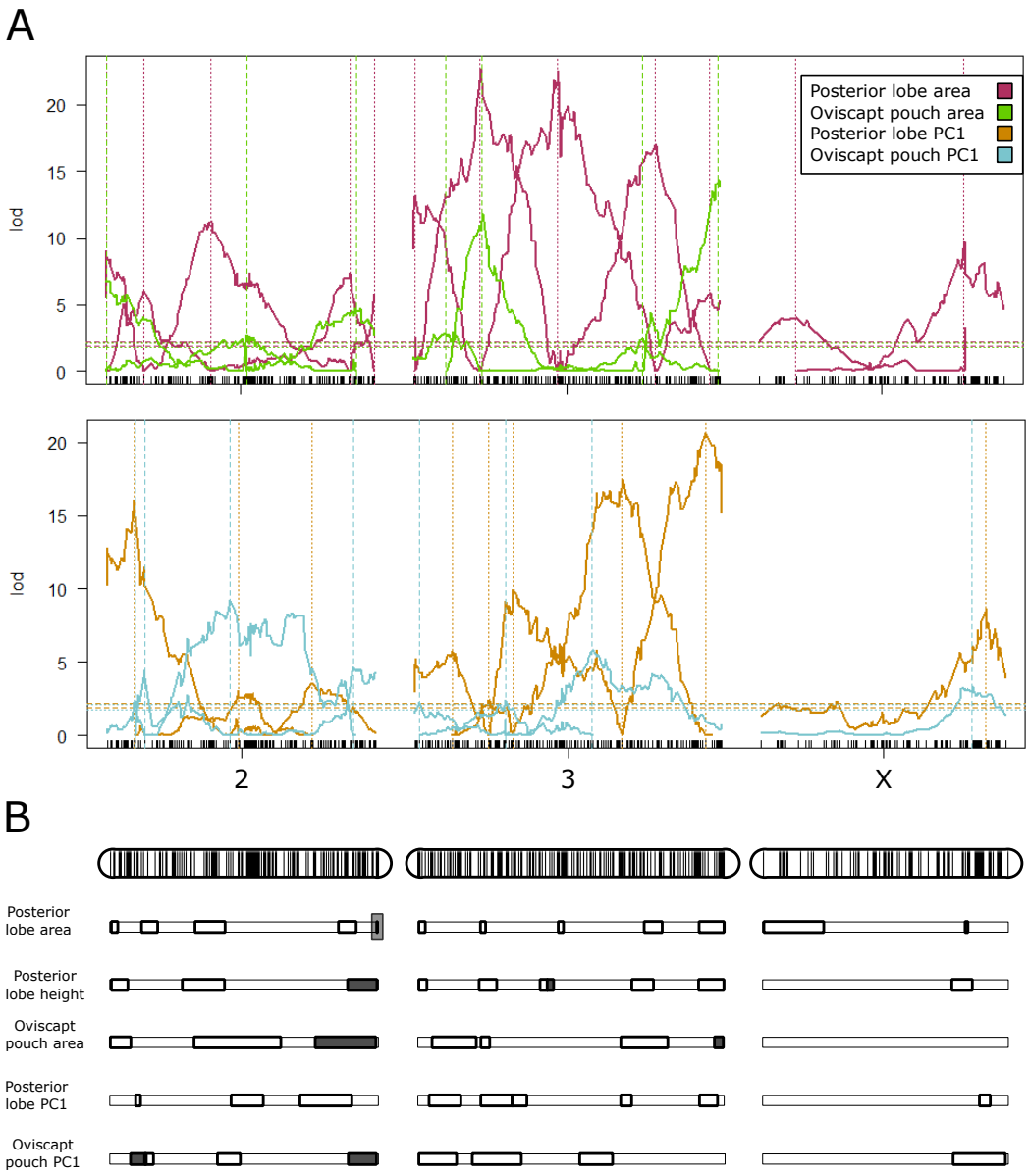

Figure 11: A. Traces from MQM scans of area and shape parameters for the posterior lobe and oviscapt pouch in the *D. simulans* backcross direction.Vertical dashed lines represent locations of QTL. Horizontal dashed lines are the LOD threshold values determined via 1,000 permutations of the data. MQM scan for posterior lobe height not shown (**Figure 14, A. appendix**) **B**. The QTL intervals for each trait visualized with boxes over their locations on the chromosome for each trait. Shaded boxes indicate a postive sign for effect size, with white boxes indicating a negative sign. In one case a shaded box is shown overlaying the QTL as the QTL was too small to show the shading itself.

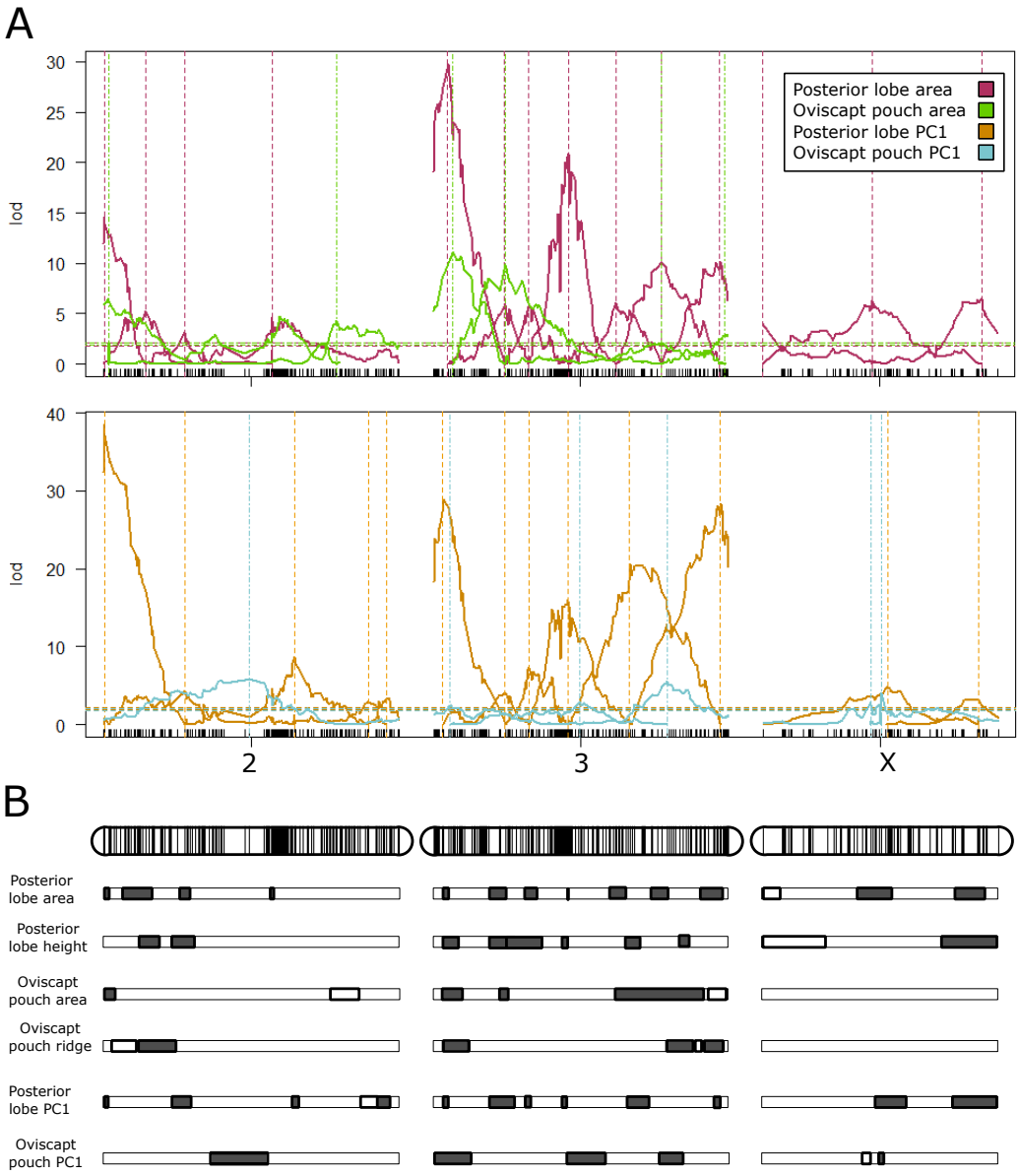

Figure 12: **A**. Traces from MQM scans of area and shape parameters for the posterior lobe and oviscapt pouch in the *D. mauritiana* backcross direction.Vertical dashed lines represent locations of QTL. Horizontal dashed lines are the LOD threshold values determined via 1,000 permutations of the data. MQM scan for posterior lobe height and oviscapt ridge not shown (**Figure 14, B-C. appendix**) **B**. The QTL intervals for each trait visualized with boxes over their locations on the chromosome for each trait. Shaded boxes indicate a postive sign for effect size, with white boxes indicating a negative sign.

Table 3: Summary information about MQM models for male and female genital phenotypes.

	Phenotype	No. QTL detected	Variance explained (%)	LOD threshold (5%)
D. simulans direction	Posterior lobe area	12	76.14348	2.26
	Posterior lobe PC1	9	71.40896	2.16
	Posterior lobe height	10	45.46946	2.10
	Oviscapt pouch area	7	38.53451	2.17
	Oviscapt pouch PC1	8	24.66418	2.09
D. mauritiana direction	Posterior lobe area	14	82.25678	2.08
	Posterior lobe PC1	13	79.83387	2.16
	Posterior lobe height	10	63.29186	2.22
	Oviscapt pouch area	6	30.55953	2.12
	Oviscapt pouch PC1	6	19.92995	2.09
	Oviscapt pouch fold	6	16.51348	2.09

Table 4: Positions and effect sizes of QTL found using Multiple QTL Models for the posterior lobe and oviscapt pouch.

Phenotype	QTL No.	Position (chr:bp)	Position (cM)	Max LOD	1.5 LOD support interval (cM)		1.5 LOD support interval (bp)		Additive allelic effects		
					From	To	From	To	Effect size	SE	Variance (%)
D. simulans backcross direction											
Posterior lobe area	**1**	2:465560	0.00	9.03	0.00	2.22	2:55258	2:2565601	-2.48E-04	3.89E-05	2.02
	2	2:6133919	9.74	5.99	8.00	12.71	2:5365409	2:7895822	-2.12E-04	4.27E-05	1.23
	3	2:14556642	26.75	11.22	21.59	29.85	2:11250851	2:16242545	-2.45E-04	3.39E-05	2.60
	4	2:43628179	62.45	7.40	58.37	63.56	2:40568707	2:44146095	-2.48E-04	4.29E-05	1.66
	5	2:47323270	68.53	5.78	68.03	68.53	2:46519904	2:47323270	2.42E-04	4.61E-05	1.38
	6	3:738782	0.48	13.15	0.00	2.10	3:28515	3:1741605	-2.52E-04	3.22E-05	3.04
	7	3:10738791	17.18	22.73	16.03	17.73	3:10315951	3:10910759	-3.73E-04	3.49E-05	5.68
	8	3:22092020	36.99	22.52	35.68	37.39	3:20217859	3:22256464	-3.52E-04	3.30E-05	5.68
	9	3:44961894	61.92	16.93	57.61	62.87	3:42888846	3:45245664	-3.46E-04	3.77E-05	4.20
	10	3:51803606	75.86	5.89	71.38	78.48	3:49616389	3:54624755	-1.86E-04	3.68E-05	1.27
	11	X:3217075	9.47	4.04	0.00	16.06	X:193188	X:4941381	-1.31E-04	3.07E-05	0.91
	12	X:16591130	52.45	9.75	51.55	52.76	X:15907441	X:16378390	-2.47E-04	3.54E-05	2.43
Posterior lobe PC1	**1**	2:4822901	6.80	16.08	6.42	8.00	2:4576653	2:5365409	-1.39E-02	1.56E-03	4.73
	2	2:18921446	33.44	3.10	30.86	39.39	2:17055908	2:30465638	-6.24E-03	1.72E-03	0.78
	3	2:37916736	52.17	3.54	48.14	62.20	2:35790863	2:43364897	-6.87E-03	1.69E-03	0.97
	4	3:6711420	9.76	5.72	2.76	11.32	3:2095722	3:7855970	-1.02E-02	2.01E-03	1.52
	5	3:11177767	19.10	2.44	16.03	24.67	3:10315951	3:13619051	-1.06E-02	2.91E-03	0.79
	6	3:14039721	25.24	9.96	24.06	28.58	3:13235278	3:15795810	-1.77E-02	2.58E-03	2.80
	7	3:40422824	53.08	17.53	51.98	55.09	3:39745093	3:41265996	-1.63E-02	1.76E-03	5.11
	8	3:51011821	74.39	20.66	72.14	77.07	3:49983430	3:52470555	-1.68E-02	1.66E-03	6.10
	9	X:19432526	57.52	8.66	55.44	58.62	X:18321529	X:19948377	-1.23E-02	1.88E-03	2.53
Posterior lobe height	**1**	2:465560	0.00	3.50	0.00	5.11	2:55258	2:3868839	-9.73E-04	2.01E-04	2.66
	2	2:11250851	21.59	6.37	18.39	29.85	2:10077897	2:16242545	-1.10E-03	2.03E-04	3.29
	3	2:47070599	68.43	2.38	60.66	68.53	2:42393961	2:47323270	7.27E-04	2.00E-04	1.50
	4	3:816904	0.59	4.79	0.00	2.76	3:28515	3:2095722	-1.00E-03	2.16E-04	2.45
	5	3:10768486	17.18	7.53	15.53	20.59	3:9971901	3:11746673	-1.50E-03	2.51E-04	4.01
	6	3:18682967	32.75	3.48	31.05	33.80	3:17174960	3:19274594	-2.71E-03	7.25E-04	1.59
	7	3:19274594	33.80	2.16	32.75	34.93	3:18682967	3:19741735	2.09E-03	7.16E-04	0.97
	8	3:42769763	57.30	7.53	54.52	60.60	3:41041263	3:44084836	-1.31E-03	2.31E-04	3.66
	9	3:54624755	78.48	2.14	71.38	78.48	3:49616389	3:54624755	-7.01E-04	2.20E-04	1.15
	10	X:16480653	52.13	5.16	48.30	54.17	X:15907441	X:17555489	-1.10E-03	2.30E-04	2.60
Oviscapt pouch area	**1**	2:55258	0.00	6.81	0.00	6.17	2:55258	2:4348545	-1.37E-04	2.42E-05	3.92
	2	2:20609385	36.06	2.72	21.59	44.50	2:11250851	2:33739279	-8.62E-05	2.44E-05	1.54
	3	2:44266871	63.86	4.73	52.68	68.53	2:38160681	2:47323270	1.13E-04	2.42E-05	2.69
	4	3:5777239	8.36	2.90	3.54	15.53	3:2610087	3:9971901	-1.12E-04	3.06E-05	1.64
	5	3:10910759	17.73	11.81	16.03	18.75	3:10315951	3:11084826	-2.31E-04	3.08E-05	6.94
	6	3:43322117	58.74	2.42	51.88	64.52	3:39717201	3:46184560	-8.06E-05	2.42E-05	1.36
	7	3:53317887	78.10	14.40	75.86	78.48	3:51803606	3:54624755	2.00E-04	2.40E-05	8.57
Oviscapt pouch PC1	**1**	2:5105189	7.15	2.42	5.11	9.49	2:3868839	2:6069563	2.57E-02	7.74E-03	1.67
	2	2:6069563	9.49	4.45	8.58	11.43	2:5591123	2:7060051	-3.53E-02	7.78E-03	3.10
	3	2:17570712	31.36	9.22	27.31	33.44	2:14869478	2:18921446	-2.21E-02	3.34E-03	6.57
	4	2:43944291	63.00	4.74	60.66	68.53	2:42393961	2:47323270	1.51E-02	3.22E-03	3.31

51

	5	3:1452844	1.29	2.28	0.00	10.41	3:28515	3:7246609	-1.06E-02	3.27E-03	1.57
	6	3:12646940	23.26	2.38	13.75	26.83	3:9167246	3:15119960	-1.13E-02	3.43E-03	1.65
	7	3:34637918	45.46	5.84	41.22	50.21	3:32157416	3:38539840	-1.73E-02	3.32E-03	4.09
	8	X:17497699	53.97	3.35	48.30	62.42	X:15907441	X:22850051	-1.22E-02	3.10E-03	2.32
D. muaritiana backcross direction											
Posterior lobe area	**1**	2:476527	0.27	14.57	0.00	1.48	2:55518	2:1291852	5.56E-02	6.66E-03	2.53
	2	2:7409985	12.27	5.21	4.88	14.05	2:3752019	2:8423756	3.79E-02	7.76E-03	0.86
	3	2:11832113	23.37	3.18	21.18	24.85	2:10865136	2:12627889	2.84E-02	7.48E-03	0.52
	4	2:23461453	48.64	4.60	47.65	48.96	2:19635608	2:24148834	2.95E-02	6.44E-03	0.76
	5	3:3614588	4.08	29.82	2.67	4.73	3:2868531	3:3893128	7.18E-02	5.80E-03	5.56
	6	3:11739022	20.51	5.97	15.84	21.11	3:10388398	3:11866931	4.48E-02	8.57E-03	0.99
	7	3:15967829	27.39	5.64	26.31	29.91	3:15316501	3:17453693	4.82E-02	9.48E-03	0.94
	8	3:31426943	38.82	20.93	38.61	39.03	3:30946660	3:31609902	7.86E-02	7.74E-03	3.74
	9	3:38832531	52.41	6.11	50.55	55.69	3:38008980	3:40542459	3.76E-02	7.09E-03	1.02
	10	3:45109259	65.42	10.09	62.30	67.65	3:43680732	3:46065875	4.63E-02	6.74E-03	1.71
	11	3:51201490	82.33	10.16	76.59	83.45	3:50176245	3:53234674	3.88E-02	5.63E-03	1.73
	12	X:194201*	0.00	4.01	0.00	5.44	X:194201	X:390479	-2.30E-02	5.39E-03	0.66
	13	X:9284003	31.58	6.32	26.97	37.53	X:8081165	X:11096918	3.05E-02	5.66E-03	1.05
	14	X:20732896	62.96	6.53	55.08	64.20	X:16551114	X:22762202	2.98E-02	5.43E-03	1.09
Posterior lobe PC1	**1**	2:476527	0.27	38.60	0.00	1.48	2:55518	2:1291852	1.44E-02	2.47E-03	1.39
	2	2:11832113	23.37	4.36	19.55	25.53	2:10251242	2:13040585	1.13E-02	2.49E-03	0.84
	3	2:33036056	54.93	8.67	53.88	56.49	2:32645216	2:33694211	1.34E-02	1.55E-03	3.09
	4	2:42636906	76.10	2.99	73.79	79.20	2:41031501	2:44941788	-9.57E-03	2.30E-03	0.71
	5	2:45788504	81.19	3.34	78.60	82.85	2:44244908	2:46297438	9.63E-03	2.22E-03	0.77
	6	3:2868531	2.67	28.93	2.44	4.73	3:2702378	3:3893128	1.92E-02	1.67E-03	5.41
	7	3:11739022	20.51	4.08	15.84	23.71	3:10388398	3:12931898	9.21E-03	2.30E-03	0.66
	8	3:15967829	27.39	7.31	26.31	28.25	3:15316501	3:16366036	1.64E-02	2.52E-03	1.74
	9	3:30752499	38.50	16.02	36.69	38.61	3:27473090	3:30946660	1.92E-02	2.07E-03	3.54
	10	3:40912622	56.27	20.64	55.69	62.30	3:40542459	3:43680732	1.28E-02	1.63E-03	2.51
	11	3:52481865	82.33	28.37	80.52	82.80	3:51833229	3:52782293	1.78E-02	1.51E-03	5.70
	12	X:10290922	35.97	4.89	32.10	41.62	X:8778446	X:12324288	7.65E-03	1.55E-03	1.00
	13	X:19890071	61.96	3.31	54.54	67.53	X:16362491	X:22851029	4.76E-03	1.56E-03	0.39
Posterior lobe height	**1**	2:7409985	12.27	7.48	10.06	16.46	2:6306103	2:9390057	1.21E-03	2.07E-04	2.55
	2	2:11832113	23.37	5.88	19.55	26.32	2:10251242	2:13718662	9.87E-04	2.08E-04	1.67
	3	3:5406279	7.41	7.73	2.67	7.86	3:2868531	3:5751576	1.26E-03	2.03E-04	2.88
	4	3:11221288	18.79	2.70	15.84	23.71	3:10388398	3:12931898	1.20E-03	2.69E-04	1.48
	5	3:14770755	25.38	2.14	20.61	31.56	3:11746673	3:18591618	8.62E-04	2.88E-04	0.67
	6	3:30946660	38.61	8.35	36.93	39.03	3:28073394	3:31609902	1.28E-03	2.41E-04	2.10
	7	3:40912622	56.27	8.98	54.97	59.77	3:40263056	3:42815426	1.30E-03	2.01E-04	3.11
	8	3:47766769	72.32	6.51	70.46	73.68	3:47268800	3:48404965	1.02E-03	1.88E-04	2.19
	9	X:3453718	11.69	6.60	0.00	18.47	X:194201	X:5311637	-9.77E-04	1.70E-04	2.47
	10	X:20160809	62.31	2.82	51.35	67.53	X:15827870	X:22851029	5.46E-04	1.69E-04	0.77
Oviscapt pouch area	**1**	2:1291852	1.48	6.53	0.00	3.64	2:55518	2:3118794	9.89E-05	2.28E-05	2.61
	2	2:38089523	67.10	4.33	64.88	73.79	2:37168230	2:41031501	-7.56E-05	2.30E-05	1.50
	3	3:4156412	5.63	11.02	2.67	8.69	3:2868531	3:6383404	1.64E-04	2.63E-05	5.40
	4	3:11746673	20.61	9.86	18.99	21.90	3:11336157	3:12140025	1.77E-04	2.69E-05	6.03
	5	3:45109259	65.42	2.11	52.09	77.86	3:38725359	3:50594385	7.37E-05	2.46E-05	1.25
	6	3:53821941	83.80	2.86	78.84	84.35	3:51201490	3:54625977	-8.33E-05	2.47E-05	1.59
Oviscapt pouch PC1	**1**	2:19160607	41.88	5.80	30.23	47.45	2:16349618	2:19499996	4.47E-02	8.59E-03	4.34
	2	3:3898088	4.83	2.47	0.00	10.76	3:28899	3:7647265	2.71E-02	8.04E-03	1.82
	3	3:33858086	41.96	2.93	37.30	49.20	3:28485401	3:37293050	3.04E-02	8.26E-03	2.17
	4	3:45964104	67.03	5.57	71.54	73.52	3:47530554	3:48322853	4.17E-02	8.18E-03	4.17
	5	X:8530282	30.94	2.83	28.48	31.63	X:8339996	X:8652094	-1.20E-01	3.31E-02	2.09
	6	X:9284003	33.92	4.13	33.37	35.17	X:8984569	X:10048921	1.44E-01	3.30E-02	3.06
Oviscapt	**1**	2:5745663	9.12	2.66	2.34	9.91	2:2404305	2:6271170	-2.32E+00	7.59E-01	1.97

52

pouch ridge	**2**	2:6564425	10.62	3.87	9.64	21.13		2:6102290	2:10864340		2.74E+00	7.64E-01	2.82
	3	3:4854894	6.52	14.97	2.67	10.76		3:2868531	3:7647265		1.89E+00	2.49E-01	12.04
	4	3:47766769	72.32	2.51	66.88	75.39		3:45938719	3:49586796		2.18E+00	8.13E-01	1.21
	5	3:50106847	76.29	2.95	75.29	77.55		3:49585811	3:50524535		-2.87E+00	9.11E-01	1.69
	6	3:52087132	81.44	2.13	77.55	83.80		3:50525059	3:53821306		1.19E+00	4.35E-01	1.31

53

2.3.3 The additive effect of the majority of QTL for posterior lobe and oviscapt pouch phenotypes are in the same direction (within a backcross group), but the trend is much stronger for the posterior lobe

For the posterior lobe, almost all QTL for all parameters were consistent in their direction within a backcross (**Table 4**). This is in agreement with previous work [Zeng et al., 2000] and suggests a history of directional selection on the posterior lobe. For the *D. simulans* backcross, of the 31 QTL identified for area, height, and PC1, all effect sizes were negative (indicating higher phenotypic values for *D. simulans*) except 3, or 90.3% of QTL. For the *D. mauritiana*, backcross direction, the opposite sign dominated (indicating lower phenotypic values for *D. mauritiana*), with 34 of 37 (or 91.9%) of QTL identified for area, height, and PC1 having positive effect sizes.

The same trend was observed for females, but was not as strong. The additive genetic effect of oviscapt pouch QTL in the *D. simulans* backcross direction was generally negative (11/15 or 73.3% of QTL). The result was similar for the *D. mauritiana* direction with 13/18 (or 72.2%) oviscapt pouch QTL with positive additive genetic effect.

2.3.4 The majority of QTL for the oviscapt pouch colocalize with posterior lobe QTL

It was immediately apparent to us that the majority of QTL identified for the oviscapt pouch size in both backcross directions coincided with posterior lobe QTL (**Figure 11**, **Figure 12**). In fact, in the *D. simulans* direction, all oviscapt pouch QTL intervals overlap with a QTL interval for either posterior lobe area or PC1. In two cases, the coincidence of the peak maximum itself was also extremely good, with the third and fourth oviscapt pouch area QTL having maxima within 2cM of the fourth posterior lobe area QTL and fourth posterior lobe PC1 QTL, respectively. In other cases, the coincidence was nearly perfect, with the maxima of the first and fifth oviscapt pouch area QTL being within 1cM of the first and seventh posterior lobe area QTL, respectively. Notably, the same pattern was observed in the other backcross direction, with five of the six oviscapt pouch area QTL intervals overlapping with posterior lobe QTL intervals. Peak maxima were also very closely

associated in these cases. The first, third, and sixth oviscapt pouch area QTL were within 2cM of of the first, fifth, and eleventh posterior lobe QTL (respectively), and the fourth and fifth oviscapt area QTL were within 1cM of the sixth and tenth posterior lobe QTL (respectively).

While this observation was striking, given the large number of QTL we detected we wanted to test how likely it is that we could have arrived at this degree of coincidence, at least with respect to amount of overlap, by chance alone. Our simulation analysis of the *D. simulans* backcross revealed that the male QTL in total occupy 61,705,367 bp of genomic space, which is 48% of the genome. The female QTL occupy 76,840,503 bp of genomic space, or 60% of the genome. We found an overlap between male and female QTLs of 48,849,913 bp, or 38% of the genome. This is slightly higher than the expectation of 29%. Over 10,000 simulations, the average overlap of simulated QTL was 38,450,357 bp, which is close to the expected value of 29% of the genome. Overlap of 48,849,913 bp or greater occurred in just 48 of the simulations ($p = 0.0048$). This result indicates that it is very unlikely that such a degree of overlap between male and female trait QTL that we observed would occur by chance alone.

2.3.5 Genes in shared candidate regions that have known activity in the development of the posterior lobe also show activity in the developing oviscapt pouch in Drosophila melanogaster

Previous work had uncovered a gene regulatory network that is involved in the development of the posterior lobe ([Glassford et al., 2015]). The results of that paper provided evidence that this network was an instance of "network co-option", wherein an existing regulatory network is redeployed in a novel context during development ([True and Carroll, 2002, Olson, 2006, Shubin et al., 2009, Monteiro, 2012, Peter and Davidson, 2015]). In this case, it was suggested that the co-option event described was involved in the origination of this novel structure, that is, the redeployment of the ancestral regulatory network in the genitalia (a novel location) of the ancestor of flies in this group contributed to the morphological changes which resulted in the initial formation of the posterior lobe. Given this knowledge

of the history and developmental genetics of the posterior lobe, we wanted to investigate two further questions: First, we wanted to know whether any of those developmental genes, which were identified as possibly involved in the origination of the posterior lobe, could also be implicated in the subsequent structural divergence of the genitalia. To do this, we determined whether any of these co-opted posterior lobe network genes fell within QTL intervals for any of our phenotypes.

We found that many of the genes that were already known to be involved in posterior lobe development and possibly origination ([Glassford et al., 2015]) were within QTL identified for either the posterior lobe, oviscapt pouch, or both (**Table 5**). While not all instances in which the gene occurs within the interval are particularly compelling (e.g. when the gene occurs far from the peak maximum or is in a part of the interval that is below the lod threshold for significance), several are promising candidates. For instance, the gene *Gef64C* occurs near the peak maximum in one of the major QTL for females in both backcross directions (QTL3 in the *D. mauritiana* direction, and QTL4 in the *D. simulans* direction).

Our second question was to investigate with greater rigor whether any of these known posterior lobe development genes were also associated with the development of the oviscapt pouch, providing stronger evidence for pleiotropic linkage between these two structures. For the six of the posterior lobe development genes that occurred in QTL for both males and females, we used expression analysis to investigate whether these genes are deployed in the developing female genitalia. In the instances of *eya* and *crb*, we performed antibody stains in pupal genitalia to visualize the presence of the gene product. We performed *in situ* hybridizations of *upd*, *Poxn*, *Gef64C* and *Cad86C* to visualize the presence of transcripts. In all six cases, it appears that these genes are deployed in the female genitalia during development (**Figure 13**, **A-F**, arrows). In addition, for *eya*, *crb*, *upd*, *Gef94C*, and *Cad86C*, the GFP reporter assays of posterior lobe enhancer sequences drive expression in a pattern that at least partially overlaps with the observed pattern of gene expression **Figure 13**, **G-K**, arrows). These results suggest that the posterior lobe enhancer also contains regulatory information directing gene expression in the developing female genitalia in or around the developing oviscapt pouch, indicating that these enhancers may be pleiotropic. Only in the case of *Poxn* did the reporter fail to drive expression in the female genitalia (**Figure 13**, **L**).

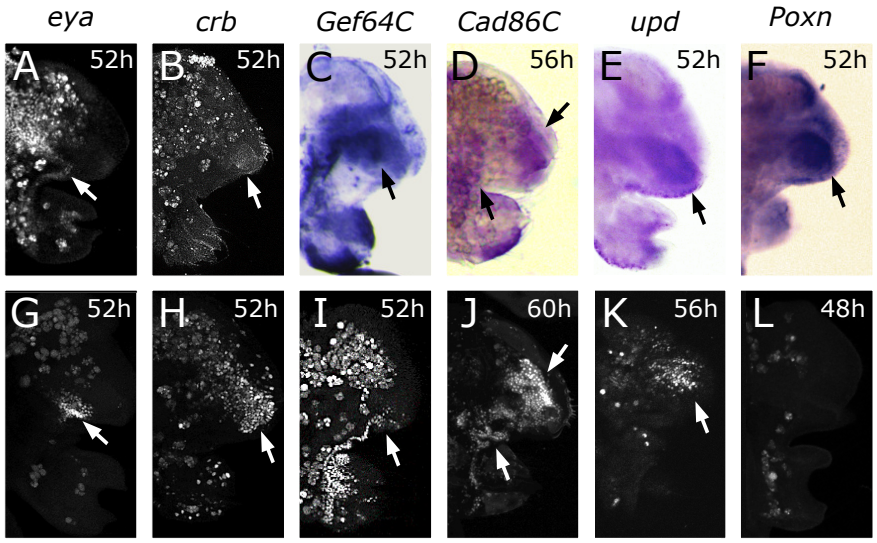

Figure 13: Pupal ovispositors visualizing the presence of posterior lobe network genes in the developing genitalia. **A.** Immunostaining of *eya*, **B.** Immunostaining of *crb*, **C.** *in situ* hybridization for *Gef64C*, **D.** *in situ* hybridization for *Cad86C*, **E.** *in situ* hybridization for *upd*, **F.** *in situ* hybridization for *Poxn*. GFP reporter expression driven in the female pupal genitalia of transgenic *D. melanogaster* flies by posterior lobe network enhancers: **G.** *eya*, **H.** *crb*, **I.** *Gef64C*, **J.** *Cad86C*, **K.** *upd*, **L.** *Poxn*. Arrows point to patterns of interest.

Table 5: Posterior lobe network genes and their locations on the backcross maps.

Posterior lobe network gene	Flybase Number	Chromosome	D. simulans backcross map postion (nearest marker)	D. simulans backcross map postion (cM)	D. mauritiana backcross map postion (nearest marker)	D. mauritiana backcross map postion (cM)	Falls within QTL interval posterior lobe	Falls within QTL interval oviscapt pouch
upd1	FBgn0004956	X	1X:18115762	55.09022	1X:17759908	57.66462	yes	yes
eya	FBgn0000320	2L	2:6133919	9.74016	2:6329451	10.1548	yes	yes
salm	FBgn0261648	2L	2:11203892	21.44297	2:11831221	23.32341	yes	no
En	FBgn0000577	2R	2:33739279	44.50303	2:33694211	56.48907	yes	no
Poxn	FBgn0000577	2R	2:38097131	52.42457	2:38089523	67.1026	yes	yes
Gef64C	FBgn0035574	3L	3:4850228	6.914839	3:4854894	6.523867	yes	yes
STAT	FBgn0016917	3R	3:32157416	41.21803	3:31972617	39.51191	no	no
AbdB	FBgn0000015	3R	3:35440729	46.97362	3:35716676	46.98888	no	yes
ems	FBgn0259685	3R	3:39195258	51.39621	3:38832531	52.41458	yes	no
Cad86C	FBgn0261053	3R	3:41774450	55.50125	3:41867465	57.51818	yes	yes
crumbs	FBgn0000015	3R	3:47054691	66.7049	3:47268800	70.4565	yes	yes

2.4 DISCUSSION AND CONCLUSION

In this work, we assessed the possibility that the size correlation between a pair of recently evolved male and female gential traits in the *D. melanogaster* subgroup might be caused by genetic covariation, which could be a consequence of pleiotropy. That pleiotropy ([Mayr, 1963, Arnold, 1973, Langerhans et al., 2016]), or genetic correlations ([Simmons and Garcia-Gonzalez, 2011, House et al., 2020]) could contribute to patterns of evolution in male and female genitalia has long been recognized but has rarely been tested ([Langerhans et al., 2016, House et al., 2020]). To our knowledge, only two other studies, both in *Ohomopterus* ground beetles, have directly compared the QTL of male and female genital stuctures ([Sasabe et al., 2010, Fujisawa et al., 2019]). These studies focused on relatively ancient genital structures, and found only minimal genetic linkage. Furthermore, in our study we looked within colocalizing male and female QTL and examined a number of genes known to be involved in posterior lobe development, as well as their posteror lobe-specific regulatory elements, to directly test the molecular mechanims of genetic correlation or pleiotropy. Our findings revealed strong support for the plausibility of pleiotropy acting to shape patterns of covariation in this system, and provide novel insight into longstanding questions about co-evolving parts and correlated patterns of morphological variation.

Our results for the posterior lobe confirm previous work ([Zeng et al., 2000]) that found a large number of QTL for the posterior lobe divergence between these two species. For the oviscapt pouch of females, we also identified a number of QTL associated with divergence. When comparing our results across the sexes, we found that all but one QTL identified for oviscapt pouch area colocalized with QTL for posterior lobe metrics. This was a striking result. The number of QTL is large and therefore a substantial percentage of the genome falls within our QTL regions, however our simulation study provides evidence that the degree of overlap between the oviscapt pouch QTL and posterior lobe QTL is unlikely to have occurred simply by chance. The fact that these QTL not only overlap but that a significant percentage (8/13) of the predicted oviscapt pouch QTL maxima (that is, the location of highest liklihood for the true location of the QTL) fall within just a few cM of posterior lobe QTL is also noteworthy. While further functional work will be required to identify the genes

that underlie our QTL regions and to determine whether they are in fact pleiotropic, our results are consistent with a scenario in which either pleiotropy or genetic linkage contribute to the across species phenotypic correllations in the posterior lobe and oviscapt pouch.

It is interesting to consider which mechanisms might be contributing to the evolution of these structures in light of our data. Sexual selection has predominated the discourse on the evolution of genitalia in recent years, and many studies have found support for sexual selection driving evolution of genital varaition in both males ([Eberhard, 1985, Simmons, 2014]) and females ([Sloan and Simmons, 2019]). Our results are consistent with the previous study which found that the direction of QTL effects were nearly always uniform within a backcross direction for posterior lobe metrics ([Zeng et al., 2000]). This pattern suggests a history of directional selection on the posterior lobe ([Orr, 1998]. Given prevoius work in this system, sexual selection indeed seems a likely explanation ([LeVasseur-Viens et al., 2015, Frazee and Masly, 2015, Frazee et al., 2020]), although one experimental evolution study found that natural selection was also capable of driving changes in posterior lobe shape in *D. simulans* ([House et al., 2013]). The majority of oviscapt pouch QTL effects were also consistent with backcross direction, but the trend was less strong. This perhaps suggests that directional selection on trait value is weaker for the oviscapt pouch.

Theory predicts that if functionally related traits of males and females require some degree of phenotypic matching, the establishment of genetic linkage should be favored ([Lande and Arnold, 1983, Simmons and Garcia-Gonzalez, 2011]). Such a scenario could explain our QTL results if the posterior lobe and oviscapt pouch have a functional relationship. We do not currently know whether the oviscapt pouch and the posterior lobe interact physically. We know at least that they are not in contact after genital coupling is acheived, as after that point the posterior lobe is inserted between the eigth and ninth female tergites ([Robertson, 1988, Jagadeeshan and Singh, 2006]), which is some distance from the oviscapt pouch. Contact between these two structures would therefore only be possible during copulation attempts or in the first moments of mounting.

Various mechanisms of genital co-evolution that have been proposed could conceivably apply to this system. One study showed that modifications to the *D. simulans* posterior lobe had an effect on precopulatory selection [LeVasseur-Viens et al., 2015]. Males in this

group often make several copulation attempts while courting a female, during which the genitalia make contact but coupling is not acheived. If the posterior lobe and oviscapt pouch come into contact during such copulation attempts, a functional interaction between them could lead to genital coevolution due to sexual selection by female choice [Brennan and Prum, 2015], or as a mechanism of hybridization avoidance, commonly referred to as "lock and key" ([Dufour, 1844, Masly, 2012]). Sexual conflict leading to sexually antagonistic coevolution of male and female genitalia is also well-documented ([Brennan and Prum, 2015]). Genital wounding is often observed in cases of sexually antagonistic evolution [Rönn et al., 2007, Tatarnic and Cassis, 2010], and indeed one study using introgression lines between *D. simulans* and *D. mauritiana* showed that posterior lobes with phenotypes outside the species norm can cause wounding to females ([Masly and Kamimura, 2014]). However, the observed wounding did not appear to be in the area of the oviscapt pouch, so it seems unlikely that this would be the explanation for coevolution of these particular structures. If the posterior lobe and oviscapt pouch do not interact at all, functional coevolution of any kind would be difficult to invoke as an explanation as to why these two structures both track each other phenotypically across species and seem to share loci of divergence in that process [Thompson, 2014]. Futher behavioral and functional experiments incorporating both strucures will be required to determine whether or how functional coevolution is occurring in this system.

Our developmental genetics data provides strong support that the oviscapt pouch and posterior lobe have pleiotropic genetic architechture at least in some respects. Our unidirectional QTL effects are consistent with a history of directional selection, however selection on only one structure would still generate patterns of phenotypic covariance between these two genital traits if pleiotropy between them is extensive. Thus, even in the absence of functional coevolution, pleiotropy alone could be sufficient to explain a pattern of correlation across species such as the one we observe here. Of course, we do not know whether the pleiotropic posteior lobe network genes that occur in the QTL intervals are in fact the causitive loci for those QTL, and there are many more QTL for which we have no candidate genes at present. Therefore, we cannot yet say how widespread pleiotropy is in this system. However, our discovery of pleiotropy at the level of regulation for genes thought to be involved in the evolutionary origins of the posterior lobe, combined with the fact that the oviscapt pouch

arose at the same point in this lineage, strongly suggests that these two structures have a long history of linked developmental programming.

A final intriguing note with respect to this data is that in only one case, for the gene *Poxn*, the female QTL that contains the posterior lobe network gene (QTL 2 for oviscapt pouch area in the *D. mauritiana* backcross direction) has the opposite sign of effect for the backcross direction, whereas the male QTL in that region (QTL 3 for posterior lobe PC1 in the *D. simulans* backcross direction) has the expected sign of effect for that backcross. This is also the only gene we tested for which the posterior lobe enhancer had no activity in the females (i.e. is not pleiotropic). A possible explanation for this observation could be that there was once intralocus sexual conflict over the *Poxn* enhancer, such that the optimal trait values for the male and female were not in the same direction. Theory predicts that such a scenario should resolve over time ([Bonduriansky and Chenoweth, 2009]), which could be acheived via differentiation of regulatory control. More detailed functional analyses would be required to support such a claim in this case.

Whether what we have observed here is an instance of "true" pleiotropy, or closely linked developmental gentics, or a combination of both, is still to be determined. However, either scenario is quite exciting with regard to the implications for how these two sturctures might evolve or coevolve. To our knowledge our work is the first to give evidence for such a high degree of genetic linkage between a male and female genital structure, and represents an important contribution to the literature on genital coevolution.

Appendix

Table 6: Corrections for normality and allometry for posterior lobe and oviscapt pouch phenotypes.

	Phenotype	Test normality (Shapiro-Wilk)	lambda (Boxcox)	transformation	Normality after transform (Shapiro-Wilk)	Significance of correlation with femur (pearson)	pearson coeff	Normality of regression residuals (Shapiro-Wilk)
D. simulans direction	**Posterior lobe area**	0.0111	na	none	na	0.7279	-0.0157	na
	Posterior lobe PC1	0.0087	na	none	na	0.0001	-0.1770	0.0104
	Posterior lobe height	0.7516	na	none	na	0.1454	0.0657	na
	Oviscapt pouch area	0.0012	na	none	na	0.4008	0.0374	na
	Oviscapt pouch PC1	2.16E-09	na	none	na	0.2407	-0.0521	na
D. mauritiana direction	**Posterior lobe area**	1.08E-07	0.1685	LOG(Y)	0.0025	¡ 2.2e-16	0.3625	0.2015
	mau m PC1	4.15E-06	na	none	na	2.307E-11	0.2920	0.0036
	mau m H	0.5212	na	none	na	¡ 2.2e-16	0.3950	0.1898
	mau f Area	0.3528	na	none	na	1.333E-12	0.3083	0.4926
	mau f PC1	0.1748	na	none	na	2.999E-15	0.3411	0.7267

Table 7: Primer sequences for *in situ* hybridization probes.

Gene	Species	Forward primer	Reverse primer
Cad86C	D. melanogaster	ACAACAACGGCACGTTCGAGATCAG	taatacgactcactataggCATCACTTCGCGATCGAAGCCATGC
Gef64C	D. melanogaster	GAGACGGAGCTCTTGAAGATTCTTC	taatacgactcactataggGAAATCGAAGAGCTCGTAGTTGTGG
unpaired	D. melanogaster	TTCTAGTCACATAAGAGCAACCGC	taatacgactcactatagggagaTCAAGCACTATATCACAGAT
Poxn	D. melanogaster	ACCGTGGTGAAGAAGGATCATCC	taatacgactcactataggCAGATCAAAACTGGGTCAGTGG

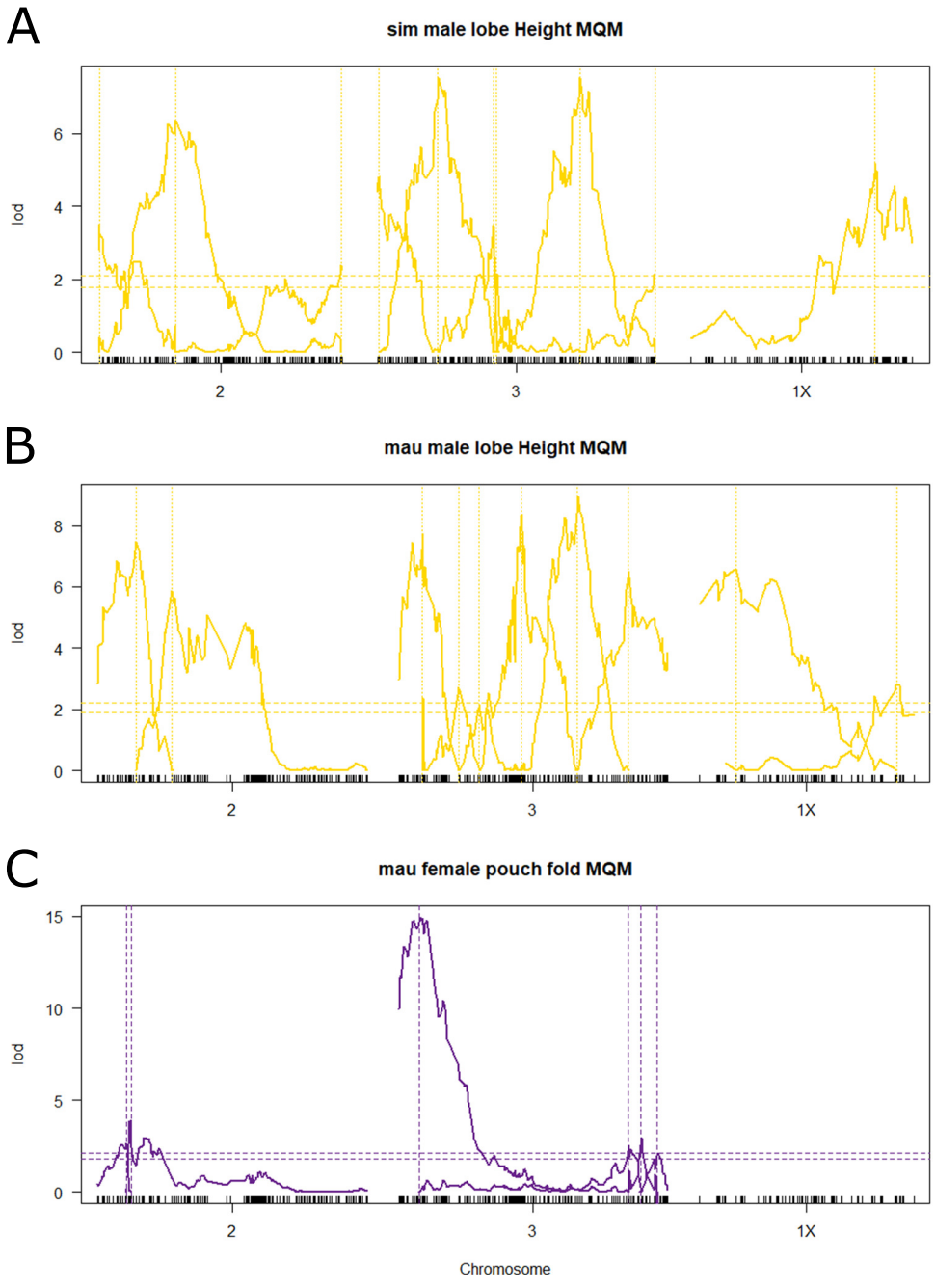

Figure 14: **A**. Trace from MQM scan of height for the posterior lobe in the *D. simulans* backcross direction. **B**. Trace from MQM scan of height for the posterior lobe in the *D. mauritiana* backcross direction. **C**. Trace from MQM scan of oviscapt ridge for the oviscapt pouch in the *D. mauritiana* backcross direction.Vertical dashed lines represent locations of QTL. Horizontal dashed lines are the LOD threshold values determined via 1,000 permutations of the data.

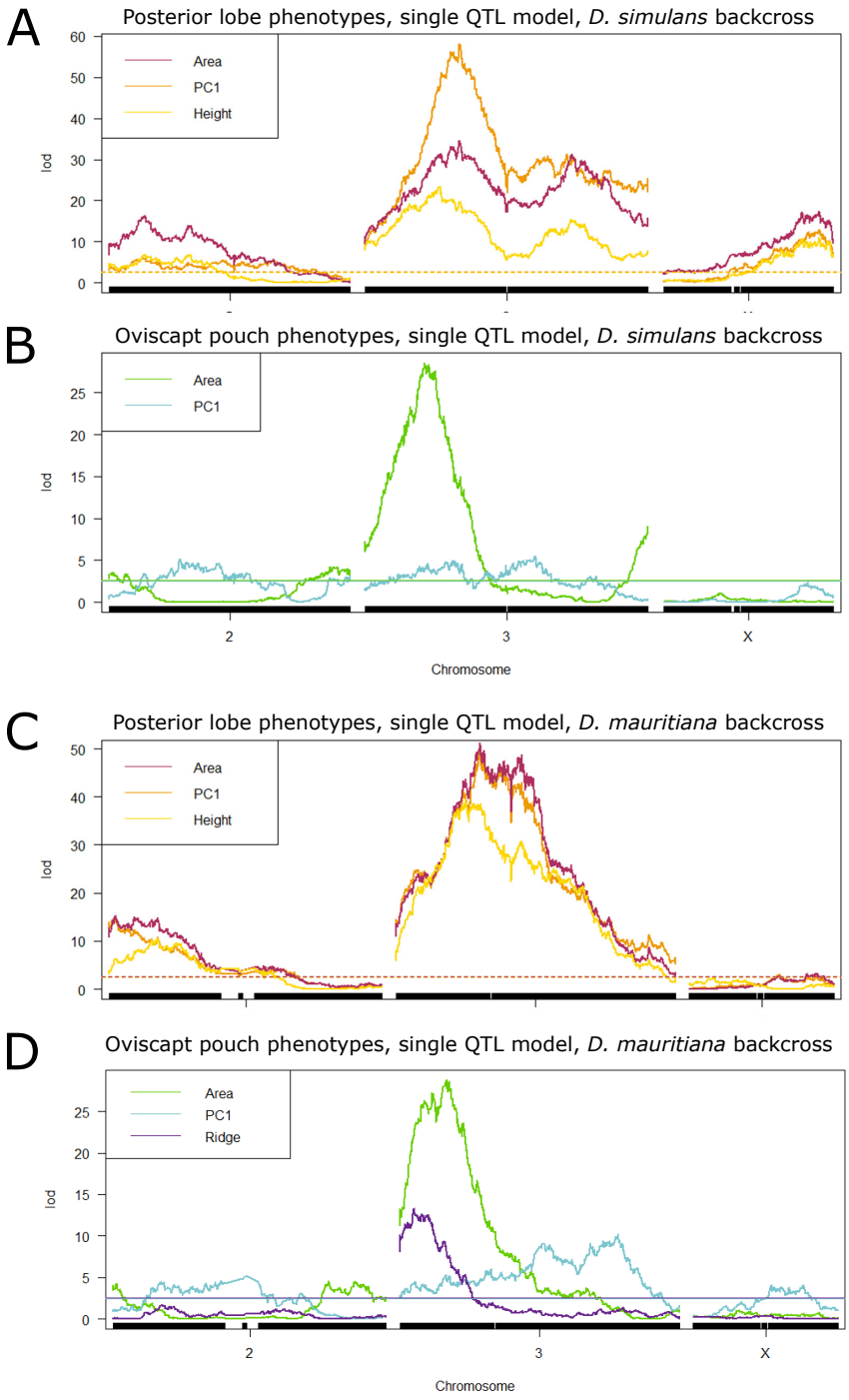

Figure 15: Scans for QTL using single QTL models for *D. simulans* posterior lobe phenotypes (**A**) and oviscapt pouch phenotypes (**B**), and for *D. mauritiana* posterior lobe phenotypes (**C**) and oviscapt pouch phenotypes (**D**)

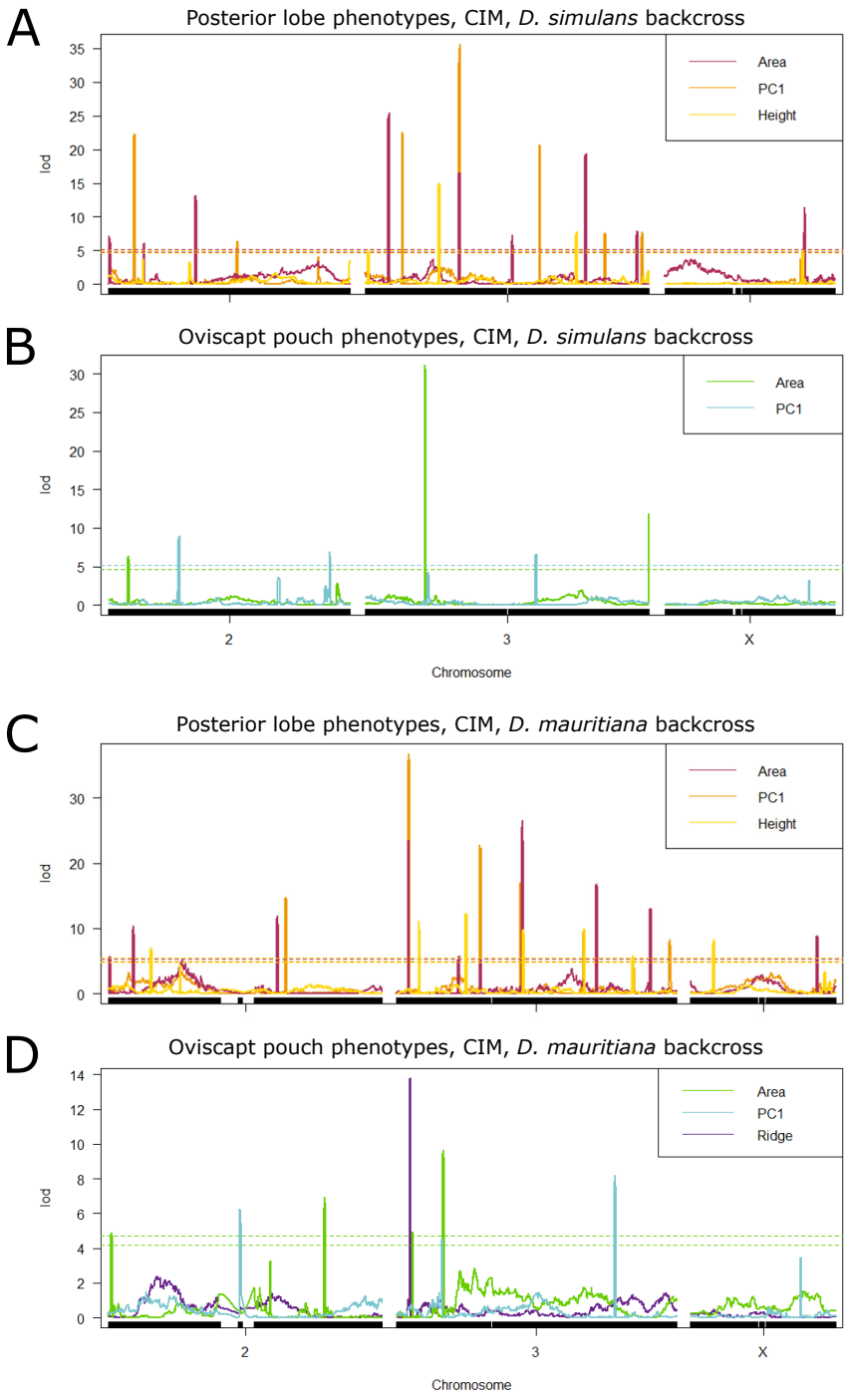

Figure 16: Composite interval mapping for *D. simulans* posterior lobe phenotypes (**A**) and oviscapt pouch phenotypes (**B**), and for *D. mauritiana* posterior lobe phenotypes (**C**) and oviscapt pouch phenotypes (**D**)

Bibliography

Malin Ah-King, Andrew B. Barron, and Marie E. Herberstein. Genital evolution: Why are females still understudied? *PLOS Biology*, 12(5):e1001851, May 2014. ISSN 1545-7885. doi: 10.1371/journal.pbio.1001851.

Roger P. Alexander, Philip M. Kim, Thierry Emonet, and Mark B. Gerstein. Understanding modularity in molecular networks requires dynamics. *Science Signaling*, 2(81):44– 44, 2009. doi: 10.1126/scisignal.281pe44. URL `https://doi.org/10.1126/scisignal.281pe44.`

Ana Maria Rocha de Almeida, Roxana Yockteng, James Schnable, Elena R. Alvarez-Buylla, Michael Freeling, and Chelsea D. Specht. Co-option of the polarity gene network shapes filament morphology in angiosperms. *Scientific Reports*, 4(1):1–9, 2014. doi: 10.1038/srep06194. URL `https://doi.org/10.1038/srep06194.`

P. Andolfatto, D. Davison, D. Erezyilmaz, T. T. Hu, J. Mast, T. Sunayama-Morita, and D. L. Stern. Multiplexed shotgun genotyping for rapid and efficient genetic mapping. *Genome Research*, 21(4):2990–2992, 2011. doi: 10.1101/gr.115402.110.

Gonçalo I. André, Renée C. Firman, and Leigh W. Simmons. The coevolution of male and female genitalia in a mammal: A quantitative genetic insight. *Evolution*, 74(7):1558–1567, 2020. ISSN 1558-5646. doi: 10.1111/evo.14031.

D. Arends, P. Prins, R. C. Jansen, and K. W. Broman. R/qtl: high-throughput multiple qtl mapping. *Bioinformatics*, 26(23):2990–2992, 2010. doi: 10.1093/bioinformatics/btq565.

Edwin Nicholas Arnold. Relationships of the palaeartic lizards assigned to the genera lacerta, algyroides and psammodromus (reptilia: Lacertidae). 1973.

Stevan J Arnold. Constraints on phenotypic evolution. *The American Naturalist*, 140:S85–S107, 1992.

Alexandra A Barnard, Ola M Fincke, Mark A McPeek, and John P Masly. Mechanical and tactile incompatibilities cause reproductive isolation between two young damselfly species. *Evolution*, 71(10):2410–2427, 2017.

Scott Barolo. Shadow enhancers: Frequently asked questions about distributed cis-regulatory information and enhancer redundancy. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 34(2):135–41, 2012. doi: 10.1002/bies.201100121. URL `https://doi.org/10.1002/bies.201100121.`

Scott Barolo and James W. Posakony. Three habits of highly effective signaling pathways: Principles of transcriptional control by developmental cell signaling. *Genes and Development*, 16(10):1167–81, 2002. doi: 10.1101/gad.976502. URL `https://doi.org/10.1101/gad.976502.`

R. Bonduriansky and S. F. Chenoweth. Intralocus sexual conflict. *Trends Ecol Evol*, 24(5): 280–8, 2009. ISSN 0169-5347 (Print) 0169-5347 (Linking). doi: 10.1016/j.tree.2008.12.005. URL http://www.ncbi.nlm.nih.gov/pubmed/19307043.

Vincent Bonhomme, Sandrine Picq, Cedric Gaucherel, and Julien Claude. Momocs: Outline analysis using r. *Journal of Statistical Software*, 56(13):1–24, 2014. URL http://www.jstatsoft.org/v56/i13/.

Patricia L.R. Brennan and Richard O. Prum. Mechanisms and evidence of genital coevolution: The roles of natural selection, mate choice, and sexual conflict. *Cold Spring Harbor Perspectives in Biology*, 7(7), Jul 2015. ISSN 1943-0264. doi: 10.1101/cshperspect.a017749. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4484975/.

Patricia LR Brennan, Richard O Prum, Kevin G McCracken, Michael D Sorenson, Robert E Wilson, and Tim R Birkhead. Coevolution of male and female genital morphology in waterfowl. *PLoS one*, 2(5):e418, 2007.

C.R. Brunetti, J.E. Selegue, A. Monteiro, V. French, P.M. Brakefield, and S.B. Carroll. The generation and diversification of butterfly eyespot color patterns. *Current Biology: CB*, 11(20):1578–85, 2001. doi: 10.1016/s0960-9822(01)00502-4. URL https://doi.org/10.1016/s0960-9822(01)00502-4.

S.B. Carroll, J. Gates, D.N. Keys, S.W. Paddock, G.E. Panganiban, J.E. Selegue, and J.A. Williams. Pattern formation and eyespot determination in butterfly wings. *Science*, 5168: 265, 1994. doi: 10.1126/science.7912449. URL https://doi.org/10.1126/science.7912449.

Sean B. Carroll. Evo-devo and an expanding evolutionary synthesis: A genetic theory of morphological evolution. *Cell*, 134(1):25–36, 2008. doi: 10.1016/j.cell.2008.06.030. URL https://doi.org/10.1016/j.cell.2008.06.030.

Ana Catalán, Adriana D. Briscoe, and Sebastian Höhna. Drift and directional selection are the evolutionary forces driving gene expression divergence in eye and brain tissue of heliconius butterflies. *Genetics*, 213(2):581–94, 2019. doi: 10.1534/genetics.119.302493. URL https://doi.org/10.1534/genetics.119.302493.

Judith Cebra-Thomas, Fraser Tan, Seeta Sistla, Eileen Estes, Gunes Bender, Christine Kim, Paul Riccio, and Scott F. Gilbert. How the turtle forms its shell: A paracrine hypothesis of carapace formation. *Journal of Experimental Zoology Part B: Molecular and Developmental Evolution*, 304B (6:558–69, 2005. doi: 10.1002/jez.b.21059. URL https://doi.org/10.1002/jez.b.21059.

R. Chaix, M. Somel, D.P. Kreil, P. Khaitovich, and G.A. Lunter. Evolution of primate gene expression: Drift and corrective sweeps? *Genetics*, 180(3):1379–89, 2008. doi: 10.1534/genetics.108.089623. URL https://doi.org/10.1534/genetics.108.089623.

Sujash S Chatterjee, Locke D Uppendahl, Moinuddin A Chowdhury, Pui-Leng Ip, and Mark L Siegal. The female-specific doublesex isoform regulates pleiotropic transcription factors to pattern genital development in drosophila. *Development*, 138(6):1099–1109, 2011.

G. A. Churchill and R. W. Doerge. Empirical threshold values for quantitative trait mapping. *Genetics*, 138:963–971, 1994.

Enrico S. Coen and Elliot M. Meyerowitz. The war of the whorls: Genetic interactions controlling flower development. *Nature*, 353(6339):31–37, 1991. doi: 10.1038/353031a0. URL `https://doi.org/10.1038/353031a0`.

Eric H. Davidson. *The Regulatory Genome: Gene Regulatory Networks In Development And Evolution*. Elsevier, 2010.

Eric H. Davidson and Douglas H. Erwin. Gene regulatory networks and the evolution of animal body plans. *Science*, 311(5762):796–800, 2006.

Lance A. Davidson. Epithelial machines that shape the embryo. *Trends in Cell Biology*, 22 (2):82–87, 2012. doi: 10.1016/j.tcb.2011.10.005. URL `https://doi.org/10.1016/j.tcb.2011.10.005`.

James DiFrisco and Johannes Jaeger. Beyond networks: Mechanism and process in evo-devo. *Biology and Philosophy*, 34(6):54, 2019. doi: 10.1007/s10539-019-9716-9. URL `https://doi.org/10.1007/s10539-019-9716-9`.

A. Dinwiddie and S. Rachootin. Patterning of a compound eye on an extinct dipteran wing. *Biology Letters*, 7(2):281–284, 2011. doi: 10.1098/rsbl.2010.0809. URL `https://doi.org/10.1098/rsbl.2010.0809`.

Marcelo C. Dornelas, Camila Maistro Patreze, Gerco C. Angenent, and Richard G.H. Immink. Mads: The missing link between identity and growth? *Trends in Plant Science*, 16 (2):89–97, 2011. doi: 10.1016/j.tplants.2010.11.003. URL `https://doi.org/10.1016/j.tplants.2010.11.003`.

I. L. Dryden. `shapes` *package*. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL `http://www.R-project.org`. Contributed package, Version 1.2.5.

Denis Duboule and Adam S. Wilkins. The evolution of 'bricolage'. *Trends in Genetics*, 14 (2):54–59, 1998. doi: 10.1016/S0168-9525(97)01358-9. URL `https://doi.org/10.1016/S0168-9525(97)01358-9`.

L. Dufour. Anatomie générale des diptères. *Annales des Sciences Naturelles*, 1:244–264, 1844.

W.G. Eberhard. *Sexual Selection and Animal Genitalia*. Harvard University Press, 1985. ISBN 978-0-674-80283-4.

Douglas J. Emlen, Laura Corley Lavine, and Ben Ewen-Campen. On the origin and evolutionary diversification of beetle horns. *Proceedings of the National Academy of Sciences of the United States of America*, 1(May):8661–68, 2007. doi: 10.1073/pnas.0701209104. URL https://doi.org/10.1073/pnas.0701209104.

Douglas H. Erwin. Evolutionary dynamics of gene regulation. *Current Topics in Developmental Biology*, 139(1):407–431, 2020. doi: 10.1016/bs.ctdb.2020.02.006. URL https://doi.org/10.1016/bs.ctdb.2020.02.006.

Carlos Espinosa-Soto and Andreas Wagner. Specialization can drive the evolution of modularity. *PLOS Computational Biology*, 6(3):1000719, 2010. doi: 10.1371/journal.pcbi.1000719. URL https://doi.org/10.1371/journal.pcbi.1000719.

Jonathan P Evans, Clelia Gasparini, Gregory I Holwell, Indar W Ramnarine, Trevor E Pitcher, and Andrea Pilastro. Intraspecific evidence from guppies for correlated patterns of male and female genital trait diversification. *Proceedings of the Royal Society B: Biological Sciences*, 278(1718):2611–2620, 2011.

Emma K. Farley, Katrina M. Olson, and Michael S. Levine. Regulatory principles governing tissue specificity of developmental enhancers. *Cold Spring Harbor Symposia on Quantitative Biology*, 80:27–32, 2015. doi: 10.1101/sqb.2015.80.027227. URL https://doi.org/10.1101/sqb.2015.80.027227.

Tyler B. Faust, Jennifer M. Binning, John D. Gross, and Alan D. Frankel. Making sense of multifunctional proteins: Human immunodeficiency virus type 1 accessory and regulatory proteins and connections to transcription. *Annual Review of Virology*, 4(1):241–60, 2017. doi: 10.1146/annurev-virology-101416-041654. URL https://doi.org/10.1146/annurev-virology-101416-041654.

Cera R. Fisher, Jill L. Wegrzyn, and Elizabeth L. Jockusch. Co-option of wing-patterning genes underlies the evolution of the treehopper helmet. *Nature Ecology and Evolution*, page 1–11, 2019. doi: 10.1038/s41559-019-1054-4. URL https://doi.org/10.1038/s41559-019-1054-4.

Ronald A. Fisher. *The Genetical Theory of Natural Selection*. Clarendon Press, Oxford, 1930.

Nicolás Frankel, Gregory K. Davis, Diego Vargas, Shu Wang, François Payre, and David L. Stern. Phenotypic robustness conferred by apparently redundant transcriptional enhancers. *Nature*, 466(7305):490–93, 2010. doi: 10.1038/nature09158. URL https://doi.org/10.1038/nature09158.

Stephen R. Frazee and John P. Masly. Multiple sexual selection pressures drive the rapid evolution of complex morphology in a male secondary genital structure. *Ecology and Evolution*, 5(19):4437–4450, Oct 2015. ISSN 2045-7758. doi: 10.1002/ece3.1721.

Stephen R. Frazee, Angelica R. Harper, Mehrnaz Afkhami, Michelle L. Wood, John C. McCrory, and John P. Masly. Interspecific introgression reveals a role of male genital morphology during the evolution of reproductive isolation in drosophila. *bioRxiv*, page 2020.06.03.132100, Jun 2020. doi: 10.1101/2020.06.03.132100.

Tomochika Fujisawa, Masataka Sasabe, Nobuaki Nagata, Yasuoki Takami, and Teiji Sota. Genetic basis of species-specific genitalia reveals role in species diversification. *Science advances*, 5(6):eaav9939, 2019.

Feng Gao and Eric H. Davidson. Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution. *Proceedings of the National Academy of Sciences*, 105(16):6091–96, 2008. doi: 10.1073/pnas.0801201105. URL `https://doi.org/10.1073/pnas.0801201105.`

A. García-Bellido. Genetic control of wing disc development in drosophila. *Ciba Foundation Symposium*, 0(29):161–82, 1975. doi: 10.1002/9780470720110.ch8. URL `https://doi.org/10.1002/9780470720110.ch8.`

BC Genevcius, DS Caetano, and CF Schwertner. Rapid differentiation and asynchronous coevolution of male and female genitalia in stink bugs. *Journal of evolutionary biology*, 30 (3):461–473, 2017.

Bruno C. Genevcius, Monique N. Simon, Tamara Moraes, and Cristiano F. Schwertner. Copulatory function and development shape modular architecture of genitalia differently in males and females. *Evolution*, 74(6):1048–1062, 2020. ISSN 1558-5646. doi: 10.1111/evo.13977.

William J. Glassford, Winslow C. Johnson, Natalie R. Dall, Sarah Jacquelyn Smith, Yang Liu, Werner Boll, Markus Noll, and Mark Rebeiz. Co-option of an ancestral hox-regulated network underlies a recently evolved morphological novelty. *Developmental Cell*, 34(5): 520–31, 2015. doi: 10.1016/j.devcel.2015.08.005. URL `https://doi.org/10.1016/j.devcel.2015.08.005.`

Nicolas Gompel, Benjamin Prud'homme, Patricia J. Wittkopp, Victoria A. Kassner, and Sean B. Carroll. Chance caught on the wing: Cis -regulatory evolution and the origin of pigment patterns in drosophila. *Nature*, 433(7025):481–87, 2005. doi: 10.1038/nature03235. URL `https://doi.org/10.1038/nature03235.`

Nicole Gorfinkiel, Lucas Sánchez, and Isabel Guerrero. Drosophila terminalia as an appendage-like structure. *Mechanisms of development*, 86(1-2):113–123, 1999.

Sara Green and Robert Batterman. Biology meets physics: Reductionism and multi-scale modeling of morphogenesis. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, 61(February):20–34, 2017. doi: 10.1016/j.shpsc.2016.12.003. URL `https://doi.org/10.1016/j.shpsc.2016.12.003.`

G. Halder, P. Callaerts, and W.J. Gehring. Induction of ectopic eyes by targeted expression of the eyeless gene in drosophila. *Science*, 267(5205):1788–92, 1995. doi: 10.1126/science.7892602. URL https://doi.org/10.1126/science.7892602.

C. S. Haley and S. A. Knott. A simple regression method for mapping quantitative trait loci in line crosses using flanking markers. *Heredity*, 69:315–324, 1992.

Benedikt Hallgrímsson, Heather Jamniczky, Nathan M. Young, Campbell Rolian, Trish E. Parsons, Julia C. Boughner, and Ralph S. Marcucio. Deciphering the palimpsest: Studying the relationship between morphological integration and phenotypic covariation. *Evolutionary Biology*, 36(4):355–76, 2009. doi: 10.1007/s11692-009-9076-5. URL https://doi.org/10.1007/s11692-009-9076-5.

Thomas F. Hansen. Is modularity necessary for evolvability?: Remarks on the relationship between pleiotropy and evolvability. *Biosystems*, 69(2):83–94, 2003. doi: 10.1016/S0303-2647(02)00132-6. URL https://doi.org/10.1016/S0303-2647(02)00132-6.

Matthew P. Harris, John F. Fallon, and Richard O. Prum. Shh-bmp2 signaling module and the evolutionary origin and diversification of feathers. *Journal of Experimental Zoology*, 294 (2):160–76, 2002. doi: 10.1002/jez.10157. URL https://doi.org/10.1002/jez.10157.

Matthew P. Harris, Scott Williamson, John F. Fallon, Hans Meinhardt, and Richard O. Prum. Molecular evidence for an activator–inhibitor mechanism in development of embryonic feather branching. *Proceedings of the National Academy of Sciences*, 102(33): 11734–39, 2005. doi: 10.1073/pnas.0500781102. URL https://doi.org/10.1073/pnas.0500781102.

Veronica F. Hinman, Albert Nguyen, and Eric H. Davidson. Caught in the evolutionary act: Precise cis-regulatory basis of difference in the organization of gene networks of sea stars and sea urchins. *Developmental Biology*, 312(2):584–95, 2007. doi: 10.1016/j.ydbio.2007.09.006. URL https://doi.org/10.1016/j.ydbio.2007.09.006.

Joung-Woo Hong, David A. Hendrix, and Michael S. Levine. Shadow enhancers as a source of evolutionary novelty. *Science*, page 321, 2008. doi: 10.1126/science.1160631. URL https://doi.org/10.1126/science.1160631. 5894): 1314.

D. J. Hosken and P. Stockley. Sexual selection and genital evolution. *Trends in Ecology and Evolution*, 19(2):87–93, Feb 2004. ISSN 0169-5347. doi: 10.1016/j.tree.2003.11.012.

C. M. House, Z. Lewis, D. J. Hodgson, N. Wedell, M. D. Sharma, J. Hunt, and D. J. Hosken. Sexual and natural selection both influence male genital evolution. *PLoS One*, 8 (5):e63807, 2013. ISSN 1932-6203 (Electronic) 1932-6203 (Linking). doi: 10.1371/journal.pone.0063807. URL http://www.ncbi.nlm.nih.gov/pubmed/23717488.

Clarissa House, Philip Tunstall, James Rapkin, Mathilda J. Bale, Matthew Gage, Enrique del Castillo, and John Hunt. Multivariate stabilizing sexual selection and the evolution

of male and female genital morphology in the red flour beetle. *Evolution*, 74(5):883–896, May 2020. ISSN 0014-3820. doi: 10.1111/evo.13912.

David J. Irons and Nicholas A.M. Monk. Identifying dynamical modules from genetic regulatory systems: Applications to the segment polarity network. *BMC Bioinformatics*, 8(1):413, 2007. doi: 10.1186/1471-2105-8-413. URL `https://doi.org/10.1186/1471-2105-8-413.`

François Jacob. Evolution and tinkering. *Science*, 196(4295):1161–66, 1977.

S. Jagadeeshan and R. S. Singh. A time-sequence functional analysis of mating behaviour and genital coupling in drosophila: role of cryptic female choice and male sex-drive in the evolution of male genitalia. *Journal of Evolutionary Biology*, 19(4):1058–1070, 2006. ISSN 1420-9101. doi: 10.1111/j.1420-9101.2006.01099.x.

R. C. Jansen. Controlling the type i and type ii errors in mapping quantitative trait loci. *Genetics*, 138(3):871–881, Nov 1994. ISSN 0016-6731.

Hilde Janssens, Shuling Hou, Johannes Jaeger, Ah-Ram Kim, Ekaterina Myasnikova, David Sharp, and John Reinitz. Quantitative and predictive model of transcriptional control of the drosophila melanogaster even skipped gene. *Nature Genetics*, 38(10):1159–65, 2006. doi: 10.1038/ng1886. URL `https://doi.org/10.1038/ng1886.`

Eric L Keisman and Bruce S Baker. The drosophila sex determination hierarchy modulates wingless and decapentaplegic signaling to deploy dachshund sex-specifically in the genital imaginal disc. *Development*, 128(9):1643–1656, 2001.

David N. Keys, David L. Lewis, Jane E. Selegue, Bret J. Pearson, Lisa V. Goodrich, Ronald L. Johnson, Julie Gates, Matthew P. Scott, and Sean B. Carroll. Recruitment of a hedgehog regulatory circuit in butterfly eyespot evolution. *Science*, 283(5401):532–34, 1999.

Philipp Khaitovich, Svante Pääbo, and Gunter Weiss. Toward a neutral evolutionary model of gene expression. *Genetics*, 170(2):929–39, 2005. doi: 10.1534/genetics.104.037135. URL `https://doi.org/10.1534/genetics.104.037135.`

Marion Kotrba, Julia Huber, and Hans R Feijen. Coevolution of male and female genitalia in stalk-eyed flies (diptera: Diopsidae). *Organisms Diversity & Evolution*, 14(2):187–201, 2014.

Marcus R Kronforst, Laura G Young, Durrell D Kapan, Camille McNeely, Rachel J O'Neill, and Lawrence E Gilbert. Linkage of butterfly mate preference and wing color preference cue at the genomic location of wingless. *Proceedings of the National Academy of Sciences*, 103(17):6575–6580, 2006.

Frank P Kuhl and Charles R Giardina. Elliptic fourier features of a closed contour. *Computer Graphics and Image Processing*, 18(3):236–258, Mar 1982. ISSN 0146-664X. doi: 10.1016/0146-664X(82)90034-X.

Matjaž Kuntner, Jonathan A Coddington, and Jutta M Schneider. Intersexual arms race? genital coevolution in nephilid spiders (araneae, nephilidae). *Evolution: International Journal of Organic Evolution*, 63(6):1451–1463, 2009.

Russell Lande and Stevan J. Arnold. The measurement of selection on correlated characters. *Evolution*, 37(6):1210–26, 1983. doi: 10.2307/2408842. URL `https://doi.org/10.2307/2408842`.

R. Brian Langerhans, Christopher M. Anderson, and Justa L. Heinen-Kay. Causes and consequences of genital evolution. *Integrative and Comparative Biology*, 56(4):741–751, Oct 2016. ISSN 1540-7063. doi: 10.1093/icb/icw101.

Derek Lemons, Jens H. Fritzenwanker, John Gerhart, Christopher J. Lowe, and William McGinnis. Co-option of an anteroposterior head axis patterning system for proximodistal patterning of appendages in early bilaterian evolution. *Developmental Biology*, 344 (1):358–62, 2010. doi: 10.1016/j.ydbio.2010.04.022. URL `https://doi.org/10.1016/j.ydbio.2010.04.022`.

Hélène LeVasseur-Viens, Michal Polak, and Amanda J. Moehring. No evidence for external genital morphology affecting cryptic female choice and reproductive isolation in drosophila. *Evolution; International Journal of Organic Evolution*, 69(7):1797–1807, Jul 2015. ISSN 1558-5646. doi: 10.1111/evo.12685.

Mike Levine. Transcriptional enhancers in animal development and evolution. *Current Biology : CB*, 20(17):754–63, 2010. doi: 10.1016/j.cub.2010.06.070. URL `https://doi.org/10.1016/j.cub.2010.06.070`.

Renhua Li, Shirng-Wern Tsaih, Keith Shockley, Ioannis M Stylianou, Jon Wergedal, Beverly Paigen, and Gary A Churchill. Structural model analysis of multiple quantitative traits. *PLoS Genet*, 2(7):e114, 2006.

Jianjun Liu, John M Mercer, Lynn F Stam, Gregory C Gibson, Zhao-Bang Zeng, and Cathy C Laurie. Genetic analysis of a morphological shape difference in the male genitalia of drosophila simulans and d. mauritiana. *Genetics*, 142(4):1129–1145, 1996.

Grace A. Loredo, Alla Brukman, Matthew P. Harris, David Kagle, Elizabeth E. Leclair, Rachel Gutman, and Erin Denney. Development of an evolutionarily novel structure: Fibroblast growth factor expression in the carapacial ridge of turtle embryos. *Journal of Experimental Zoology*, 291(3):274–81, 2001. doi: 10.1002/jez.1103. URL `https://doi.org/10.1002/jez.1103`.

C.Owen Lovejoy, Richard S. Meindl, James C. Ohman, Kingsbury G. Heiple, and Tim D. White. The maka femur and its bearing on the antiquity of human walking: Applying

contemporary concepts of morphogenesis to the human fossil record. *American Journal of Physical Anthropology*, 119(2):97–133, 2002. doi: 10.1002/ajpa.10111. URL https://doi.org/10.1002/ajpa.10111.

Nicholas M. Luscombe, M.Madan Babu, Haiyuan Yu, Michael Snyder, Sarah A. Teichmann, and Mark Gerstein. Genomic analysis of regulatory network dynamics reveals large topological changes. *Nature*, 431(7006):308–12, 2004. doi: 10.1038/nature02782. URL https://doi.org/10.1038/nature02782.

Elena R. Álvarez Buylla, Barbara A. Ambrose, Eduardo Flores-Sandoval, Marie Englund, Adriana Garay-Arroyo, Berenice García-Ponce, and Eduardo Torre-Bárcena. B-function expression in the flower center underlies the homeotic phenotype of lacandonia schismatica (triuridaceae. *The Plant Cell*, 22(11):3543–59, 2010. doi: 10.1105/tpc.109.069153. URL https://doi.org/10.1105/tpc.109.069153.

Trudy F.C. Mackay. Epistasis and quantitative traits: Using model organisms to study gene–gene interactions. *Nature Reviews Genetics*, 15(1):22–33, 2014. doi: 10.1038/nrg3627. URL https://doi.org/10.1038/nrg3627.

Richard S. Mann and Sean B. Carroll. Molecular mechanisms of selector gene function and evolution. *Current Opinion in Genetics and Development*, 12(5):592–600, 2002. doi: 10.1016/S0959-437X(02)00344-1. URL https://doi.org/10.1016/S0959-437X(02)00344-1.

Sylvain Marcellini and Pat Simpson. Two or four bristles: Functional evolution of an enhancer of scute in drosophilidae. *PLoS Biology*, 4(12):386, 2006. doi: 10.1371/journal.pbio.0040386. URL https://doi.org/10.1371/journal.pbio.0040386.

Fabrice Marcillac, Yaël Grosjean, and Jean-François Ferveur. A single mutation alters production and discrimination of drosophila sex pheromones. *Proceedings of the Royal Society B: Biological Sciences*, 272(1560):303–309, 2005.

John P Masly. 170 years of "lock-and-key": genital morphology and reproductive isolation. *International Journal of Evolutionary Biology*, 2012, 2012.

John P Masly and Yoshitaka Kamimura. Asymmetric mismatch in strain-specific genital morphology causes increased harm to drosophila females. *Evolution*, 68(8):2401–2411, 2014.

Alexandra L. Mattei, Mark L. Riccio, Frank W. Avila, and Mariana F. Wolfner. Integrated 3d view of postmating responses by the drosophila melanogaster female reproductive tract, obtained by micro-computed tomography scanning. *Proceedings of the National Academy of Sciences*, 112(27):8475–8480, Jul 2015. ISSN 0027-8424, 1091-6490. doi: 10.1073/pnas.1505797112.

E. Mayr. *Animal Species and Evolution*. Belknap Press. Belknap Press of Harvard University Press, 1963.

Eden McQueen and Mark Rebeiz. *Chapter Twelve - On the specificity of gene regulatory networks: How does network co-option affect subsequent evolution?*, volume 139 of *Current Topics in Developmental Biology*, page 375–405. Academic Press, 2020. doi: 10.1016/bs.ctdb.2020.03.002. URL http://www.sciencedirect.com/science/article/pii/S0070215320300491.

Luigi Minafra, Valentina Bravatà, Giusi Irma Forte, Francesco Paolo Cammarata, Maria Carla Gilardi, and Cristina Messa. Gene expression profiling of epithelial–mesenchymal transition in primary breast cancer cell culture. *Anticancer Research*, 34 (5):2173–83, 2014.

Vivian Méndez and Alex Córdoba-Aguilar. Sexual selection and animal genitalia. *Trends in Ecology and Evolution*, 19(5):224–225, May 2004. ISSN 0169-5347. doi: 10.1016/j.tree.2004.03.012.

A.P. Moczek. On the origins of novelty and diversity in development and evolution: A case study on beetle horns. *Cold Spring Harbor Symposia on Quantitative Biology*, 74:289–96, 2009. doi: 10.1101/sqb.2009.74.010. URL https://doi.org/10.1101/sqb.2009.74.010.

Armin P. Moczek and Lisa M. Nagy. Diverse developmental mechanisms contribute to different levels of diversity in horned beetles. *Evolution and Development*, 7(3):175–85, 2005. doi: 10.1111/j.1525-142X.2005.05020.x. URL https://doi.org/10.1111/j.1525-142X.2005.05020.x.

Armin P. Moczek and Debra J. Rose. Differential recruitment of limb patterning genes during development and diversification of beetle horns. *Proceedings of the National Academy of Sciences of the United States of America*, 106(22):8992–97, 2009. doi: 10.1073/pnas.0809668106. URL https://doi.org/10.1073/pnas.0809668106.

A. Monteiro and M.D. Gupta. Identifying coopted networks and causative mutations in the origin of novel complex traits. *Current Topics in Developmental Biology*, 119:205–26, 2016. doi: 10.1016/bs.ctdb.2016.03.003. URL https://doi.org/10.1016/bs.ctdb.2016.03.003.

Antónia Monteiro. Gene regulatory networks reused to build novel traits. *BioEssays*, 34 (3):181–86, 2012. doi: 10.1002/bies.201100160. URL https://doi.org/10.1002/bies.201100160.

Olga Nagy, Isabelle Nuez, Rosina Savisaar, Alexandre E. Peluffo, Amir Yassin, Michael Lang, David L. Stern, Daniel R. Matute, Jean R. David, and Virginie Courtier-Orgogozo. Correlated evolution of two copulatory organs via a single cis-regulatory nucleotide change. *Current Biology*, 28(21):3450–3457 13, 2018. doi: 10.1016/j.cub.2018.08.047. URL https://doi.org/10.1016/j.cub.2018.08.047.

Tetsuya Nakamura, Jeff Klomp, Joyce Pieretti, Igor Schneider, Andrew R. Gehrke, and Neil H. Shubin. Molecular mechanisms underlying the exceptional adaptations of batoid

fins. *Proceedings of the National Academy of Sciences*, 112(52):15940–45, 2015. doi: 10.1073/pnas.1521818112. URL https://doi.org/10.1073/pnas.1521818112.

H.F. Nijhout. Metaphors and the role of genes in development. *BioEssays : News and Reviews in Molecular, Cellular and Developmental Biology*, 1990.

Jeffrey C. Oliver, Xiao-Ling Tong, Lawrence F. Gall, William H. Piel, and Antónia Monteiro. A single origin for nymphalid butterfly eyespots followed by widespread loss of associated gene expression. *PLoS Genetics*, 8(8):1002893, 2012. doi: 10.1371/journal.pgen.1002893. URL https://doi.org/10.1371/journal.pgen.1002893.

Eric N. Olson. Gene regulatory networks in the evolution and development of the heart. *Science*, 5795:313, 2006. doi: 10.1126/science.1132292. URL https://doi.org/10.1126/science.1132292.

H Allen Orr. Testing natural selection vs. genetic drift in phenotypic evolution using quantitative trait locus data. *Genetics*, 149(4):2099–2104, 1998.

H.A. Orr. Adaptation and the cost of complexity. *Evolution; International Journal of Organic Evolution*, 54(1):13–20, 2000. doi: 10.1111/j.0014-3820.2000.tb00002.x. URL https://doi.org/10.1111/j.0014-3820.2000.tb00002.x.

Sarah P. Otto. Two steps forward, one step back: The pleiotropic effects of favoured alleles. In *Proceedings of the Royal Society B: Biological Sciences*, page 271, 2004.

Annalise B Paaby and Matthew V Rockman. The many faces of pleiotropy. *Trends in Genetics*, 29(2):66–73, 2013.

Mihaela Pavlicev, Jane P. Kenney-Hunt, Elizabeth A. Norgard, Charles C. Roseman, Jason B. Wolf, and James M. Cheverud. Genetic variation in pleiotropy: Differential epistasis as a source of variation in the allometric relationship between long bone lengths and body weight. *Evolution*, 62(1):199–213, 2008. doi: 10.1111/j.1558-5646.2007.00255.x. URL https://doi.org/10.1111/j.1558-5646.2007.00255.x.

Michael W. Perry, Alistair N. Boettiger, Jacques P. Bothma, and Michael Levine. Shadow enhancers foster robustness of drosophila gastrulation. *Current Biology: CB*, 20(17):1562–67, 2010. doi: 10.1016/j.cub.2010.07.043. URL https://doi.org/10.1016/j.cub.2010.07.043.

Isabelle S. Peter and Eric H. Davidson. *Genomic Control Process: Development and Evolution*. Academic Press, London, UK ; San Diego, CA, USA, 1 edition, 2015.

Patrick C. Phillips. Epistasis–the essential role of gene interactions in the structure and evolution of genetic systems. *Nature Reviews. Genetics*, 9(11):855–67, 2008. doi: 10.1038/nrg2452. URL https://doi.org/10.1038/nrg2452.

S. Picelli, A. K. Björklund, B. Reinius, S. Sagasser, G. Winberg, and R. Sandberg. Tn5 transposase and tagmentation procedures for massively scaled sequencing projects. *Genome Research*, 24(12):2033–2040, 2014. doi: 10.1101/gr.177881.1140.

Ella Preger Ben-Noon, Gonzalo Sabarís, Daniela M. Ortiz, Jonathan Sager, Anna Liebowitz, David L. Stern, and Nicolás Frankel. Comprehensive analysis of a cis-regulatory region reveals pleiotropy in enhancer function. *Cell Reports*, 22(11):3021–31, 2018. doi: 10.1016/j.celrep.2018.02.073. URL `https://doi.org/10.1016/j.celrep.2018.02.073.`

B. Prud'homme, C. Minervino, M. Hocine, J. D. Cande, A. Aouane, H. D. Dufour, V. A. Kassner, and N. Gompel. Body plan innovation in treehoppers through the evolution of an extra wing-like appendage. *Nature*, 473(7345):83–86, 2011. doi: 10.1038/nature09977. URL `https://doi.org/10.1038/nature09977.`

Richard O. Prum. Evolution of the morphological innovations of feathers. *Journal of Experimental Zoology. Part B, Molecular and Developmental Evolution*, 304(6):570–79, 2005. doi: 10.1002/jez.b.21073. URL `https://doi.org/10.1002/jez.b.21073.`

Richard O. Prum and Alan H. Brush. The evolutionary origin and diversification of feathers. *The Quarterly Review of Biology*, 77(3):261–95, 2002. doi: 10.1086/341993. URL `https://doi.org/10.1086/341993.`

R.O. Prum. Development and evolutionary origin of feathers. *The Journal of Experimental Zoology*, 285(4):291–306, 1999.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2017. URL `https://www.R-project.org/`.

Mark Rebeiz and Miltos Tsiantis. Enhancer evolution and the origins of morphological novelty. *Current Opinion in Genetics and Development*, 45(August):115–23, 2017. doi: 10.1016/j.gde.2017.04.006. URL `https://doi.org/10.1016/j.gde.2017.04.006.`

Mark Rebeiz, Nick Jikomes, Victoria A. Kassner, and Sean B. Carroll. Evolutionary origin of a novel gene expression pattern through co-option of the latent activities of existing regulatory sequences. *Proceedings of the National Academy of Sciences of the United States of America*, 108(25):10036–43, 2011. doi: 10.1073/pnas.1105937108. URL `https://doi.org/10.1073/pnas.1105937108.`

Mark Rebeiz, Nipam H. Patel, and Veronica F. Hinman. Unraveling the tangled skein: The evolution of transcriptional regulatory networks in development. *Annual Review of Genomics and Human Genetics*, 16(1):103–31, 2015. doi: 10.1146/annurev-genom-091212-153423. URL `https://doi.org/10.1146/annurev-genom-091212-153423.`

Gavin Rice and Mark Rebeiz. Evolution: How many phenotypes do regulatory mutations affect? *Current Biology*, 29(1):21–23, 2019. doi: 10.1016/j.cub.2018.11.027. URL `https://doi.org/10.1016/j.cub.2018.11.027.`

Hugh M. Robertson. Mating asymmetries and phylogeny in the drosophila melanogaster species complex. *Pacific Science*, 42(1-2):72–80, 1988.

Derek A Roff and DJ Fairbairn. The evolution of trade-offs: where are we? *Journal of evolutionary biology*, 20(2):433–447, 2007.

Rebekah L. Rogers, Julie M. Cridland, Ling Shao, Tina T. Hu, Peter Andolfatto, and Kevin R. Thornton. Landscape of standing variation for tandem duplications in drosophila yakuba and drosophila simulans. *Molecular Biology and Evolution*, 31(7):1750–1766, Jul 2014. ISSN 0737-4038. doi: 10.1093/molbev/msu124.

Johanna Rönn, Mari Katvala, and Göran Arnqvist. Coevolution between harmful male genitalia and female resistance in seed beetles. *Proceedings of the National Academy of Sciences*, 104(26):10921–10925, 2007.

Robert S. Rosenthal. The file drawer problem and tolerance for null results. In *In*. 1979. doi: 10.1037/0033-2909.86.3.638. URL `https://doi.org/10.1037/0033-2909.86.3.638`.

Locke Rowe and Göran Arnqvist. Sexual selection and the evolution of genital shape and complexity in water striders. *Evolution: International Journal of Organic Evolution*, 66 (1):40–54, 2012.

Curtis T Rueden, Johannes Schindelin, Mark C Hiner, Barry E DeZonia, Alison E Walter, Ellen T Arena, and Kevin W Eliceiri. Imagej2: Imagej for the next generation of scientific image data. *BMC bioinformatics*, 18(1):529, 2017.

T. Ruths and L. Nakhleh. Neutral forces acting on intragenomic variability shape the escherichia coli regulatory network topology. *Proceedings of the National Academy of Sciences of the United States of America*, 110(19):7754–7759, 2013. doi: 10.1073/pnas.1217630110. URL `https://doi.org/10.1073/pnas.1217630110`.

Gonzalo Sabarís, Ian Laiker, Ella Preger-Ben Noon, and Nicolás Frankel. Actors with multiple roles: Pleiotropic enhancers and the paradigm of enhancer modularity. *Trends in Genetics*, 35(6):423–33, 2019. doi: 10.1016/j.tig.2019.03.006. URL `https://doi.org/10.1016/j.tig.2019.03.006`.

J.P.J. Saeij, S. Coller, J.P. Boyle, M.E. Jerome, M.W. White, and J.C. Boothroyd. Toxoplasma co-opts host gene expression by injection of a polymorphic kinase homologue. *Nature*, 445(7125):324–27, 2007. doi: 10.1038/nature05395. URL `https://doi.org/10.1038/nature05395`.

Lucas Sánchez and Isabel Guerrero. The development of the drosophila genital disc. *Bioessays*, 23(8):698–707, 2001.

Masataka Sasabe, Yasuoki Takami, and Teiji Sota. Qtl for the species-specific male and female genital morphologies in ohomopterus ground beetles. *Molecular Ecology*, 19(23): 5231–5239, 2010.

Johannes Schindelin, Ignacio Arganda-Carreras, Erwin Frise, Verena Kaynig, Mark Longair, Tobias Pietzsch, Stephan Preibisch, Curtis Rueden, Stephan Saalfeld, Benjamin Schmid, et al. Fiji: an open-source platform for biological-image analysis. *Nature methods*, 9(7): 676–682, 2012.

Stephan Schneuwly, Roman Klemenz, and Walter J. Gehring. Redesigning the body plan of drosophila by ectopic expression of the homoeotic gene antennapedia. *Nature*, 325(6107): 816–18, 1987. doi: 10.1038/325816a0. URL `https://doi.org/10.1038/325816a0`.

Karen E. Sears, Terence D. Capellini, and Rui Diogo. On the serial homology of the pectoral and pelvic girdles of tetrapods. *Evolution*, 69(10):2543–55, 2015.

Sandeep N Shah, Leslie Cope, Weijie Poh, Amy Belton, Sujayita Roy, C Conover Talbot Jr, Saraswati Sukumar, David L Huso, and Linda MS Resar. Hmga1: A master regulator of tumor progression in triple-negative breast cancer cells. *PLOS ONE*, 8(5):63419, 2013. doi: 10.1371/journal.pone.0063419. URL `https://doi.org/10.1371/journal.pone.0063419`.

Arthur M Shapiro and Adam H Porter. The lock-and-key hypothesis: evolutionary and biosystematic interpretation of insect genitalia. *Annual review of entomology*, 34(1):231–245, 1989.

Neil Shubin, Cliff Tabin, and Sean Carroll. Deep homology and the origins of evolutionary novelty. *Nature*, 457(7231):818–23, 2009. doi: 10.1038/nature07891. URL `https://doi.org/10.1038/nature07891`.

Leigh W. Simmons. Sexual selection and genital evolution. *Austral Entomology*, 53(1):1–17, 2014. ISSN 1558-5646. doi: 10.1111/aen.12053.

Leigh W. Simmons and Francisco Garcia-Gonzalez. Experimental coevolution of male and female genital morphology. *Nature Communications*, 2(11):374, Jul 2011. ISSN 2041-1723. doi: 10.1038/ncomms1379.

Nadia S. Sloan and Leigh W. Simmons. The evolution of female genitalia. *Journal of Evolutionary Biology*, 32(9):882–899, 2019. ISSN 1420-9101. doi: 10.1111/jeb.13503.

Alexander V. Spirov, Marat A. Sabirov, and David M. Holloway. In silico evolution of gene cooption in pattern-forming gene networks. *The Scientific World Journal*, 2012. doi: 10.1100/2012/560101. URL `https://doi.org/10.1100/2012/560101`.

Max V. Staller, Dong Yan, Sakara Randklev, Meghan D. Bragdon, Zeba B. Wunderlich, Rong Tao, Lizabeth A. Perkins, Angela H. DePace, and Norbert Perrimon. Depleting gene activities in early drosophila embryos with the 'maternal-gal4–shrna' system. *Genetics*, 193(1):51–61, 2013. doi: 10.1534/genetics.112.144915. URL `https://doi.org/10.1534/genetics.112.144915`.

Frank W. Stearns. One hundred years of pleiotropy: A retrospective. *Genetics*, 186(3): 767–773, Nov 2010. ISSN 0016-6731, 1943-2631. doi: 10.1534/genetics.110.122549.

David L. Stern and Virginie Orgogozo. The loci of evolution: How predictable is genetic evolution? *Evolution; International Journal of Organic Evolution*, 62(9):2155–77, 2008. doi: 10.1111/j.1558-5646.2008.00450.x. URL https://doi.org/10.1111/j.1558-5646.2008.00450.x.

David L. Stern and Virginie Orgogozo. Is genetic evolution predictable? *Science*, 323(5915): 746–51, 2009.

Kushal Suryamohan, Casey Hanson, Emily Andrews, Saurabh Sinha, Molly Duman Scheel, and Marc S. Halfon. Redeployment of a conserved gene regulatory network during aedes aegypti development. *Developmental Biology*, 416(2):402–13, 2016. doi: 10.1016/j.ydbio.2016.06.031. URL https://doi.org/10.1016/j.ydbio.2016.06.031.

Kentaro M Tanaka, Yoshitaka Kamimura, and Aya Takahashi. Mechanical incompatibility caused by modifications of multiple male genital structures using genomic introgression in drosophila. *Evolution*, 72(11):2406–2418, 2018.

Basile Tarchini, Denis Duboule, and Marie Kmita. Regulatory constraints in the evolution of the tetrapod limb anterior–posterior polarity. *Nature*, 443(7114):985–88, 2006. doi: 10.1038/nature05247. URL https://doi.org/10.1038/nature05247.

Nikolai J Tatarnic and G Cassis. Sexual coevolution in the traumatically inseminating plant bug genus coridromius. *Journal of evolutionary Biology*, 23(6):1321–1326, 2010.

D.A.W. Thompson. *On Growth and Form*. Dover Books on Biology Series. Dover Publications, 1917. ISBN 978-0-486-67135-2.

J.N. Thompson. *Interaction and Coevolution*. Interaction and Coevolution. University of Chicago Press, 2014. ISBN 978-0-226-12732-3.

Yoshinori Tomoyasu, Yasuyuki Arakane, Karl J. Kramer, and Robin E. Denell. Repeated co-options of exoskeleton formation during wing-to-elytron evolution in beetles. *Current Biology: CB*, 19(24):2057–65, 2009. doi: 10.1016/j.cub.2009.11.014. URL https://doi.org/10.1016/j.cub.2009.11.014.

John R. True and Sean B. Carroll. Gene co-option in physiological and morphological evolution. *Annual Review of Cell and Developmental Biology*, 18(1):53–80, 2002. doi: 10.1146/annurev.cellbio.18.020402.140619. URL https://doi.org/10.1146/annurev.cellbio.18.020402.140619.

Guido Van Rossum and Fred L Drake Jr. *Python reference manual*. Centrum voor Wiskunde en Informatica Amsterdam, 1995.

Berta Verd, Nicholas A.M. Monk, and Johannes Jaeger. Modularity, criticality, and evolvability of a developmental gene regulatory network, 2019. URL https://doi.org/10.7554/eLife.42832. ELife 8 (June): e42832.

Ben J. Vincent, Max V. Staller, Francheska Lopez-Rivera, Meghan D.J. Bragdon, Edward C.G. Pym, Kelly M. Biette, Zeba Wunderlich, Timothy T. Harden, Javier Estrada, and Angela H. DePace. Hunchback is counter-repressed to regulate even-skipped stripe 2 expression in drosophila embryos. *PLoS Genetics*, 14(9):1007644, 2018. doi: 10.1371/journal.pgen.1007644. URL https://doi.org/10.1371/journal.pgen.1007644.

Ben J. Vincent, Gavin R. Rice, Gabriella M. Wong, William J. Glassford, Kayla I. Downs, Jessica L. Shastay, Kenechukwu Charles-Obi, Malini Natarajan, Madelaine Gogol, Julia Zeitlinger, and et al. An atlas of transcription factors expressed in male pupal terminalia of drosophila melanogaster. *G3: Genes, Genomes, Genetics*, 9(12):3961–3972, Dec 2019. ISSN 2160-1836. doi: 10.1534/g3.119.400788.

Broman K. W., H. Wu, S. Sen, and G.A. Churchill. R/qtl: Qtl mapping in experimental crosses. *Bioinformatics*, 19(7):889–890, 2003. doi: 10.1093/bioinformatics/btg112.

Günter P. Wagner and Jianzhi Zhang. The pleiotropic structure of the genotype–phenotype map: The evolvability of complex organisms. *Nature Reviews Genetics*, 12(3):204–13, 2011. doi: 10.1038/nrg2949. URL https://doi.org/10.1038/nrg2949.

Günter P. Wagner, Jane P. Kenney-Hunt, Mihaela Pavlicev, Joel R. Peck, David Waxman, and James M. Cheverud. Pleiotropic scaling of gene effects and the 'cost of complexity.'. *Nature*, 452(7186):470–72, 2008. doi: 10.1038/nature06756. URL https://doi.org/10.1038/nature06756.

G.P. Wagner. *Homology, Genes, and Evolutionary Innovation*. Princeton University Press, 2014. ISBN 978-1-4008-5146-1.

Bethany R. Wasik and Armin P. Moczek. Decapentaplegic (dpp) regulates the growth of a morphological novelty, beetle horns. *Development Genes and Evolution*, 221 (1):17–27, 2011. doi: 10.1007/s00427-011-0355-7. URL https://doi.org/10.1007/s00427-011-0355-7.

John J. Welch and David Waxman. Modularity and the cost of complexity. *Evolution*, 57 (8):1723–34, 2003. doi: 10.1111/j.0014-3820.2003.tb00581.x. URL https://doi.org/10.1111/j.0014-3820.2003.tb00581.x.

Mary Jane West-Eberhard. Developmental plasticity and the origin of species differences. *Proceedings of the National Academy of Sciences of the United States of America*, 1(May): 6543–49, 2005. doi: 10.1073/pnas.0501844102. URL https://doi.org/10.1073/pnas.0501844102.

Andrew Whitehead and Douglas L. Crawford. Neutral and adaptive variation in gene expression. *Proceedings of the National Academy of Sciences*, 103(14):5425–30, 2006. doi: 10.1073/pnas.0507648103. URL https://doi.org/10.1073/pnas.0507648103.

J. G. Wideman, A. Novick, S. A. Muñoz-Gómez, and W. F. Doolittle. Neutral evolution of cellular phenotypes. *Current Opinion in Genetics and Development*, 103:58–59, 2019. doi: 10.1016/j.gde.2019.09.004. URL https://doi.org/10.1016/j.gde.2019.09.004.

Molly C. Womack, Marissa J. Metz, and Kim L. Hoke. Larger genomes linked to slower development and loss of late-developing traits. *The American Naturalist*, page 000–000, 2019. doi: 10.1086/705897. URL https://doi.org/10.1086/705897.

Gregory A. Wray. The evolutionary significance of cis-regulatory mutations, 2007.

Zeba Wunderlich, Meghan D.J. Bragdon, Ben J. Vincent, Jonathan A. White, Javier Estrada, and Angela H. DePace. Krüppel expression levels are maintained through compensatory evolution of shadow enhancers. *Cell Reports*, 12(11):1740–47, 2015. doi: 10.1016/j.celrep.2015.08.021. URL https://doi.org/10.1016/j.celrep.2015.08.021.

F Yao, B Shi, X Wang, J Cheng, M Bai, J Yan, N Cumberlidge, and H Sun. Rapid divergent co-evolution of freshwater crab genitalia facilitates a burst of species diversification. *Integrative Zoology*, 2019.

Amir Yassin. Unresolved questions in genitalia coevolution: bridging taxonomy, speciation, and developmental genetics. *Organisms Diversity & Evolution*, 16(4):681–688, 2016.

Amir Yassin and Virginie Orgogozo. Coevolution between male and female genitalia in the drosophila melanogaster species subgroup. *PLOS ONE*, 8(2):e57158, Feb 2013. ISSN 1932-6203. doi: 10.1371/journal.pone.0057158.

Nathan M. Young and Benedikt Hallgrímsson. Serial homology and the evolution of mammalian limb covariation structure. *Evolution*, 59(12):2691–2704, 2005. doi: 10.1111/j.0014-3820.2005.tb00980.x. URL https://doi.org/10.1111/j.0014-3820.2005.tb00980.x.

Jozsef Zakany and Denis Duboule. The role of hox genes during vertebrate limb development. *Current Opinion in Genetics and Development*, 17(4):359–66, 2007. doi: 10.1016/j.gde.2007.05.011. URL https://doi.org/10.1016/j.gde.2007.05.011.

Eduardo E. Zattara, Hannah A. Busey, David M. Linz, Yoshinori Tomoyasu, and Armin P. Moczek. Neofunctionalization of embryonic head patterning genes facilitates the positioning of novel traits on the dorsal head of adult beetles. *Proceedings of the Royal Society B: Biological Sciences*, 283, 2016. doi: 10.1098/rspb.2016.0824. URL https://doi.org/10.1098/rspb.2016.0824.

Z. B. Zeng. Precision mapping of quantitative trait loci. *Genetics*, 136:1457–1468, 1994.

Zhao-Bang Zeng, Jianjun Liu, Lynn F Stam, Chen-Hung Kao, John M Mercer, and Cathy C Laurie. Genetic architecture of a morphological shape difference between two drosophila species. *Genetics*, 154:299–310, Jan 2000.

J. Zhang. Neutral theory and phenotypic evolution. *Molecular Biology and Evolution*, 35 (6):1327–1331, 2018. doi: 10.1093/molbev/msy065. URL https://doi.org/10.1093/molbev/msy065.

Jia Zheng, Joshua L. Payne, and Andreas Wagner. Cryptic genetic variation accelerates evolution by opening access to diverse adaptive peaks. *Science*, 365(6451):347–53, 2019. doi: 10.1126/science.aax1837. URL https://doi.org/10.1126/science.aax1837.

Nesibe Özsu and Antónia Monteiro. Wound healing, calcium signaling, and other novel pathways are associated with the formation of butterfly eyespots. *BMC Genomics*, 18, 2017. doi: 10.1186/s12864-017-4175-7. URL https://doi.org/10.1186/s12864-017-4175-7.