

**Spring 2017 – Epigenetics and Systems Biology**  
**Discussion Session (Epigenetics and Evolutionary Biology)**  
**Michael K. Skinner – Biol 476/576**  
**Week 16 (April 27)**

**Epigenetics and Evolutionary Biology**

**Primary Papers**

1. Soyer and O'Malley (2013) Bioessays 35:696-705
2. Skinner, et al. (2014) BMC Genomics 15:377
3. Skinner (2015) GBE 7:1296-1302

**Discussion**

Student 32 – Ref #1 above

- What is the evolutionary systems biology referred to?
- How can systems biology help understand evolution?

Student 36 – Ref #2 above

- What is sexual selection?
- How did epigenetic transgenerational inheritance change mate preference?
- Does this provide evidence for environmental induction of epigenetic transgenerational inheritance and a role in evolutionary biology?

Student \_\_ – Ref #3 above

- What are the neo-Lamarckian concepts and role of environmental epigenetics?
- What are the neo-Darwinian concepts and role phenotypic variation?
- What is the integration involved in the unified theory of evolution and role of epigenetics?

# Prospects & Overviews

## Evolutionary systems biology: What it is and why it matters

Orkun S. Soyer<sup>1)</sup> and Maureen A. O'Malley<sup>2)</sup>

Evolutionary systems biology (ESB) is a rapidly growing integrative approach that has the core aim of generating mechanistic and evolutionary understanding of genotype-phenotype relationships at multiple levels. ESB's more specific objectives include extending knowledge gained from model organisms to non-model organisms, predicting the effects of mutations, and defining the core network structures and dynamics that have evolved to cause particular intracellular and intercellular responses. By combining mathematical, molecular, and cellular approaches to evolution, ESB adds new insights and methods to the modern evolutionary synthesis, and offers ways in which to enhance its explanatory and predictive capacities. This combination of prediction and explanation marks ESB out as a research manifesto that goes further than its two contributing fields. Here, we summarize ESB via an analysis of characteristic research examples and exploratory questions, while also making a case for why these integrative efforts are worth pursuing.

**Keywords:**

■ evolution; genotype-phenotype mapping; integration; molecular networks; system dynamics

DOI 10.1002/bies.201300029

<sup>1)</sup> Warwick Centre for Synthetic Biology, School of Life Sciences, University of Warwick, Coventry, UK

<sup>2)</sup> Department of Philosophy, University of Sydney, Sydney, NSW, Australia

**Corresponding authors:**

Orkun S. Soyer

E-mail: o.s.soyer@exeter.ac.uk

Maureen A. O'Malley

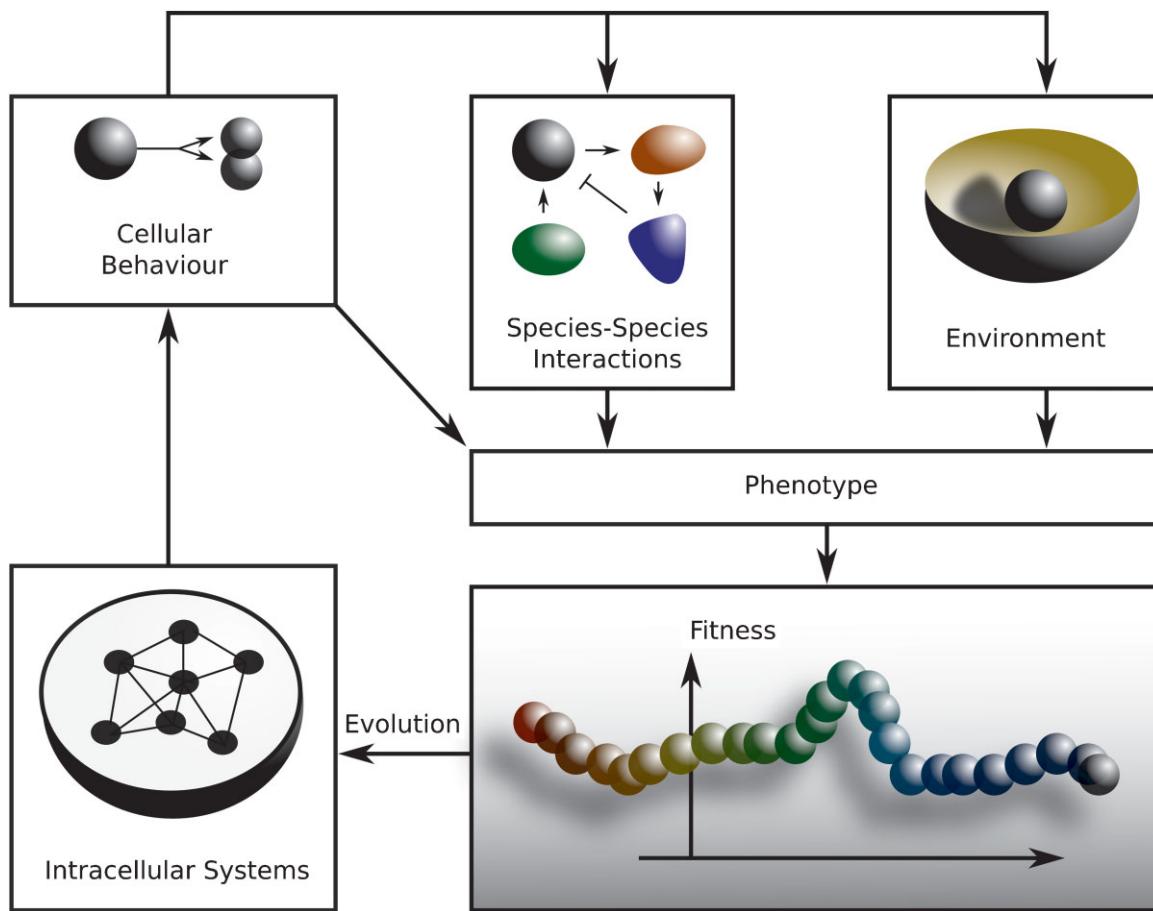
E-mail: maureen.omalley@sydney.edu.au

### The emergence of ESB

At its most basic, evolutionary systems biology (ESB) is the synthesis of system-level approaches to biological function with evolutionary explanations of multilevel properties. "System" in this context refers to dynamically interacting components that produce behavior not revealed by analyses of isolated components. Cellular interactions of signaling, regulatory, and metabolic components are all considered as systems or networks that can display structural intricacies and nonlinear dynamics. ESB recognizes that the system-level properties of cellular networks are subject to evolutionary change, and that evolved network properties will variously influence the future evolutionary course of the organism. It is this interdependency between evolutionary processes and system properties that ESB aims to understand.

One of the earliest articles to describe ESB was published in 2005 [1]. While this overview and subsequent elaborations (e.g. [2]) were milestones in the development of ESB as a research programme, their scope was restricted to comparative analyses. Since 2005 ESB has flourished, to the extent that comparative methods currently constitute only one facet of its diverse research strategies (in fact, many discussions of ESB prefer to exclude this comparative work in favor of more dynamic accounts; we make a case for including it below). ESB also goes beyond existing efforts to merge molecular and evolutionary biology (e.g. [3]). Those projects aim to understand evolutionary processes in the narrower context of individual genes and structure of their protein products, whereas ESB is committed to studying phenotypes as the results of evolving intracellular interaction networks (Fig. 1). At present, only a few ESB researchers refer self-consciously to their work with the label "evolutionary systems biology" (see [4]). However, this situation is changing as increasing numbers of evolutionarily minded, systems-focused researchers forge connections between a range of research questions (e.g. [5, 6]).

In contrast to some views of ESB, which suggest it is defined by a specific research agenda [1, 2, 7], we see the distinguishing capacity of ESB as its ability to integrate theoretical tools, experimental methods, and extensive datasets within an evolutionary framework. This integration is



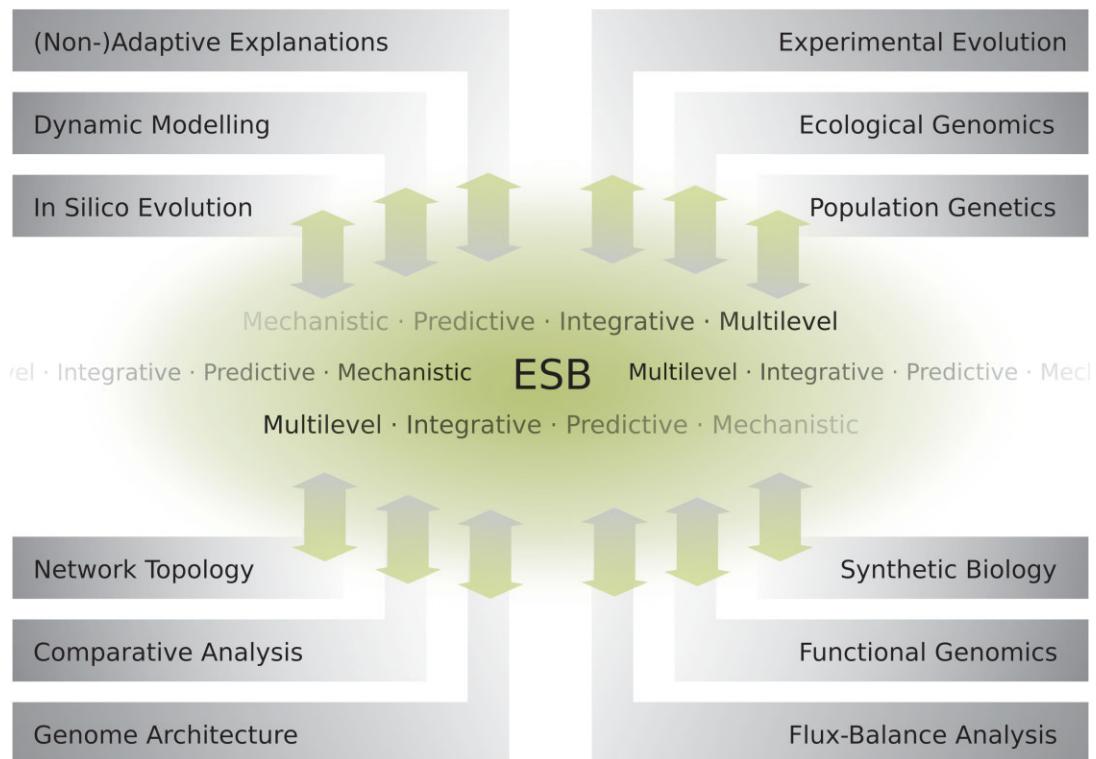
**Figure 1.** Systems under evolutionary forces. At the core of ESB lies the aim of achieving a deep mechanistic understanding of genotype-phenotype mappings in biological systems. While these mappings can be drawn at different levels in different combinations of ESB, a major area of interest is currently intracellular systems. These systems give rise to cellular physiology, which – in the case of unicellular organisms – directly determines species' interactions with their environment and other organisms. These higher-level interactions are responsible for the fitness of organisms. Evolutionary processes (i.e. neutral drift and adaptation) move populations of these organisms on this dynamic fitness landscape by altering the properties of their intracellular systems. Credit: Arno Steinacher created this figure.

occurring in a highly pragmatic manner, and is driven by the core ambition of ESB practitioners to develop closer insight into evolving genotype-phenotype mappings across different biological scales (Fig. 1). Researchers with this goal seize upon tools and datasets as they become available (e.g. dynamical models, gene-knockout studies, flux balance analyses, *in silico* evolution, reverse engineering, comparative omics data; Fig. 2) to address questions as old as biology or to reformulate new ones in light of system-level insight.

Some readers may find this description of ESB too methodologically inclusive. It could be argued that to be a viable research approach, ESB should be more restricted in what it does and does not do. But methodological flexibility

and a wide investigative agenda are what characterize many new movements in biology (e.g. systems biology itself); historically, molecular biology was not the product of a precisely organized research agenda [8]. It seems unlikely to us that biological research programmes come about as the planned products of rational intellectual design and social engineering. Instead, they emerge from communities of researchers as piecemeal responses to the needs and constraints experienced in existing fields of research. ESB is such a product. It may indeed resist any tight definitions, but as we will show, it has a synthesizing mission that distinguishes it from its contributing fields.

While systems biology may have initially overlooked evolution because modeling network dynamics is already demanding enough, there are also explicit arguments *against* taking an evolutionary approach to systems biology. One of these is that to understand the organization of systems, an evolutionary perspective is irrelevant. While specific organizational features may have evolved to cope with environmental perturbations, it is getting a grip on “how” these are implemented at the molecular level rather than “why” that is considered the important task (e.g. [9]). We will address this objection below by showing that organizational properties have dimensions appropriately and necessarily addressed by ESB. Another argument for leaving evolutionary inquiries out of systems biology might be that population genetic approaches are sufficient to understand mutations and their effects. Again, we will show how the combination of



**Figure 2.** A representation of methodological inputs into ESB. There is great variety in the methods and perspectives that can and do contribute to ESB, and this figure is merely indicative rather than an exhaustive list of contributions. The examples outlined in the text depict various combinations of these methodological inputs, although no single ESB study is likely to combine all of them. It is important to note that the interaction between ESB and these methodological contributions is two-way, with new insights and innovations achieved by ESB having feedback effects on the contributing tools and methods. ESB itself is conceptualized as an open interdisciplinary space produced by the integration of these various approaches with evolutionary frameworks. Four ESB potentials are highlighted in that space: mechanistic, predictive, integrative and multilevel capacities. Credit: Arno Steinacher created this figure.

evolutionary and systems biological analyses reveals further dimensions of mutational processes and leads to better understanding of their evolutionary outcomes.

## ESB research questions

Despite the diversity of methods used in ESB, there is a broadly unifying motivation that drives the approach. The most general reason for why systems biology requires an evolutionary framework, and why evolutionary biology should concern itself with systems analyses, is as we have suggested above: each will be a more complete science with the other perspective taken into account. ESB integrates “how” and “why” questions about biological systems, and thus appeals to network biologists who want to understand not just how a system functions but why it functions in a particular way. This combination of predictive

and explanatory capacities is a distinct feature of ESB and we will elaborate on it below. But more specifically, ESB has a number of focused research objectives that set the agenda for different aspects of its practice, and it is these objectives that determine methodological combinations. We will outline a range of these questions in relation to how they can be answered.

### Generalizations of network properties beyond model species

Usually, any systems-biological understanding of network dynamics and structure is gained via modeling and experimentation with regard to a particular model system. For example, network modeling and experiments have shown that signaling networks enabling chemotaxis in *Escherichia coli* and osmoregulation in *Saccharomyces cerevisiae* display transient (i.e. adaptive) dynamics [10, 11]. While these and similar studies show how a particular physiological response is enabled by specific system dynamics and structure, they cannot easily be generalized to other organisms. A whole raft of questions remains unanswered: Are the observed system dynamics necessary for the physiological response? What are alternative molecular structures that have the same system dynamics? Could these alternatives be realized under different ecological and evolutionary pressures? Answering these questions requires going beyond organism-specific systems biology with the help of evolutionary analyses. The evolutionary methods that are used to answer such questions include experimental and mathematical comparison of network architecture and dynamics in different species (e.g. [12–14]), comparative genomic analyses at the network

level (e.g. [15, 16]), in silico evolution of network models (e.g. [17, 18]), and statistical inference and network reconstruction methods (e.g. [19]).

### Determining the evolutionary forces behind network-level patterns

Numerous studies have demonstrated global or recurring features in cellular networks (e.g. [20]). These organizational characteristics exist at different levels of the biological hierarchy and include specific connectivity distributions, the presence of modules and motifs, and biochemical features (e.g. cooperative binding, noise, and multifunctional enzymes). Understanding the significance of these features requires an account of the evolutionary forces that shaped the networks in which these patterns were found. In particular, the observed features can be adaptive (e.g. the result of specific selective pressures such as stabilizing or fluctuating selection [21]) or neutral (e.g. resulting from genetic drift [22]), or could have emerged due to a mix of these two forces (i.e. neutral mutations that become adaptive in new circumstances) [23]. Methods include in silico evolution of simple network models under adaptive and neutral forces, comparison of network features from different species (e.g. with different population sizes giving different roles to genetic drift, or different selection pressures producing different network features), and population genetics.

### Comprehensive understanding of mutational effects

Another theme in systems biology is predicting the effects of mutations on particular functions of cellular networks (e.g. large-scale mutation or deletion studies such as those undertaken in *S. cerevisiae* [24]). Evolutionary biology, on the other hand, aims to discern the role of mutational effects in any evolutionary process. In general, systems biology addresses this research theme within the model-experiment cycle, while evolutionary biology addresses it by creating simplified models and evaluating their ability to capture experimental observations [25]. ESB bridges these two methodologies to produce a more comprehensive understanding of mutational effects. For example, sensitivity analysis over different network models has been used to gain insight into the distribution of mutational effects on a specific property of system dynamics, such as the phase of oscillation in a circadian clock model [7, 26]. In the case of metabolism, mathematical models of optimal metabolite flows in metabolic networks under a steady state assumption (metabolic flux-balance analysis (FBA [27])), have been used to predict phenotypes produced by single-enzyme deletions [6]. Experimental evolution can be combined with next-generation sequencing to characterize both the mutations and their fitness effects (e.g. antibiotic resistance [28]). A valuable methodological development in ESB would be establishing a cycle between evolutionary experimentation and system-level modeling.

### Extending systems biology to co-evolved intercellular interactions

Contemporary molecular systems biology has until recently focused on intracellular networks. Examining intercellular network interactions in homogenous cultures of microorganisms (and between cells of multicellular organisms) has been carried out in the context of metabolism (via exchanged metabolites), quorum sensing, and the production of “public goods” such as scavenging enzymes or co-factors. In host-parasite interactions, heterogeneous cellular interactions can occur via parasite-encoded genes that function in the networks of the host. Systems-biological study of these cellular interactions does not necessarily require an evolutionary framework (e.g. applications of FBA to interactions between two species [29, 30]). However, cellular interactions have historically been studied to great effect in an evolutionary context, particularly with regard to the co-evolutionary dynamics and evolution of social behavior in microbes [31]. ESB is strategically well placed to combine network approaches with these broader evolutionary perspectives.

**To sum up,** the research questions driving ESB combine those asked separately in evolutionary biology and systems biology. In addressing these questions, ESB integrates the tools and approaches of evolutionary biology and systems biology to achieve results that are more than additive: systems biology gains a more multilevel, explanatory perspective with the incorporation of evolutionary timescales, while evolutionary biology is made more quantitative, predictive, and able to combine mathematical abstractions with mechanistic detail. ESB research can therefore deliver insights not possible from evolutionary or systems biology approaches applied in isolation. We will show more specifically how this works with two apt illustrations of ESB’s constructive dynamic.

### ESB exemplified: Robustness, epistasis, and pleiotropy

While there are large numbers of ESB studies to choose from (e.g. [5]), we will use as exemplars two areas of research that not only illuminate ESB-in-action, but also deal with some of the resistance to ESB we mentioned above. Robustness refers to a system’s ability to maintain its output in the face of perturbations. It has been demonstrated for several phenotypes via gene knockout studies [24] and metabolic flux measurements [32]. These elegant systems-biological studies were descriptive, however, and could not explain the evolutionary significance of robustness; nor could they make predictions about which evolved systems should be expected to display this property [33]. While several theoretical studies have attempted to address these questions (e.g. [34, 35]), more complete insights have come from recent ESB work that amalgamates population genetics, network-level dynamical models, genomic analyses, flux-balance models, and in silico evolution.

These studies point to plausible selective pressures that would generate robustness against gene deletion in gene-regulatory and metabolic networks [36–39]. Selection for

stable gene-expression patterns under recombination results in the evolution of robustness as a product of network connectivity [40] and the emergence of epistatic interactions among nodes [36, 37]. In the case of metabolic networks, selection for biomass production from variable resources (i.e. fluctuating selection) leads to the evolution of robust networks with overlapping metabolic routes and multifunctional enzymes [38]. These mechanistic determinants of robustness fit closely the findings of flux-redistribution experiments [32]. Additional ESB studies combining genomic data and FBA add support to claims that robustness depends on environmental conditions [39, 41].

These novel insights into the mechanistic determinants and evolutionary drivers of network-level robustness came about by synthesizing the methods of systems and evolutionary biology. Mechanistic findings complement systems-biology studies that aim to decipher the overall structure, dynamics, and function of metabolic networks (e.g. [32]). The improved understanding of the evolutionary forces involved has generated valuable predictions, including the hypotheses that gene-regulatory networks harbor genes able to buffer the effects of mutations [36], and that metabolic robustness in microbes parallels the evolutionary and ecological exposure of those organisms to fluctuating environments [38, 39, 41].

Epistasis (multiple genes or mutations producing phenotypic effects that deviate from a simple additive model) and pleiotropy (in which one gene or mutation affects multiple phenotypic traits) are crucial aspects of genotype-phenotype relationships [42, 43]. In ESB, the systematic application of FBA to genome-scale metabolic models of yeast reveals that epistasis commonly takes place between functional modules of genes [44]. This finding shifts the focus from interactions occurring between individual genes to those occurring between modules. Subsequent expansion of the same analysis to many metabolic environments and multiple phenotypes shows that gene pairs in yeast can interact “incoherently” relative to different phenotypes (i.e. negatively in relation to some phenotypes and positively in relation to others) [45]. These discoveries would have been invisible to any perspective based solely on individual interactions or simple models. Similarly, studies of system-level models of developmental regulatory networks indicate that both the overall function of a gene-regulatory network [46] and its gene-regulatory patterns [47] can give rise to specific signatures of statistical epistasis. The latter finding is potentially of high methodological importance to evolutionary biology because it shows that combining models of nonlinear network dynamics with commonly used genetic methods, such as quantitative trait loci mapping, can greatly improve the performance of each approach [47, 48].

Network-level research into epistasis often inevitably leads to insight into pleiotropic effects and interactions [49, 50]. Investigations of genome-wide pleiotropic effects in yeast, nematodes, and mice indicate that pleiotropy, while occurring at a low rate, is highly modular and has an important evolutionary role in the generation of adaptive complexity in organisms [51] (however, see [52]). Studies focused on single adaptive mutations with whole-network effects thus offer general insight into pleiotropy as a mechanism of evolution [53]. Ongoing ESB research seeks to clarify how

epistasis and pleiotropy enable or obstruct evolutionary change at the network level, and whether these processes are simply inevitable byproducts of network connectivity [45, 43].

## From old to new questions in ESB

ESB is able to elaborate on existing answers to old evolutionary questions as well as devise new system-level evolutionary questions. An example of the former, in which existing questions are addressed by new ESB approaches, is found in the fields of comparative genomics and evo-devo. An ESB extension of classic comparative genomics focuses on complete systems of interacting genes to decipher the variance in the structure of these networks (i.e. what interacts with what) and the evolutionary history of those interactions [54, 16]. These comparisons also take into account the evolutionary dynamics of both genomic and cellular network architecture [55–57].

The field of evo-devo has profited considerably from contemporary capacities for comparative network analysis, especially in regard to understanding the relationships between developmental processes and the evolution of transcriptional regulatory networks (notably *cis*-regulatory networks) [58–60]. The integration of new datasets now enables the incorporation of additional levels of regulation in the study of evolving developmental processes. In particular, it is becoming increasingly feasible to examine the roles of post-transcriptional and post-translational regulation in development [61]. Emerging ESB studies are already trying to study these processes and their connections [62, 63].

But perhaps a more thoroughgoing integration is happening in regard to the combination of comparative approaches with theoretical tools such as dynamical modeling, mathematical inference, and *in silico* evolution [64]. This amalgamation has already allowed the prediction of underlying regulatory networks from the analysis of phenotypic (i.e. gene expression) data in the model organism *Drosophila* [12]. Some evo-devo proponents have argued that modeling these networks dynamically merely confirms the results of existing qualitative experimental approaches but adds no additional insight (e.g. [65]). However, models that incorporate system dynamics, network evolution, and phenotypic datasets have the advantage of being able to examine directly genotype-phenotype mapping in diverse organisms [14, 66, 67], which is something qualitative experimentation cannot do. These models have revealed that due to nonlinear interactions, knowledge of network structure is often *not* sufficient for understanding the function and dynamics of any particular network (e.g. [68]). Such findings are emphasized by *in silico* evolution studies, which find network dynamics rather than structure to be the determining and conserved features of the cellular networks underlying complex developmental phenotypes (e.g. [17, 68–71]).

The capacity of ESB to reformulate or produce entirely new questions about evolutionary dynamics in light of systems biology is clearly illustrated by the study of noise in biological systems. Noisy dynamics arising at the cellular level can manifest as phenotypic variance at the population level [72]. This phenomenon raises questions about how noise might

affect evolutionary dynamics and which evolutionary conditions favor the selection of the underlying mechanisms. The latter question arises from the fact that noise levels themselves are subject to tuning via evolutionary processes such as point mutations in regulatory regions [73]. This interplay between noise, evolutionary dynamics, and systems behavior sets the stage perfectly for an integrative ESB approach. For example, there are indications from genomic analyses that certain functional gene classes exhibit significant increases in expression noise [74]. This correlation can be explained by system-dynamic and population-genetic models that explore the types of genotype-phenotype maps that could lead to the selection of increased expression noise [74, 75]. Further ESB studies are exploring the phenotypic consequences of noise and its impact on the evolutionary dynamics of metabolic networks [76].

In regulatory networks, noise combined with nonlinear gene-regulatory dynamics can lead to the bimodal distribution of phenotypes (i.e. phenotype switching) in otherwise isogenic populations [77, 78]. This is vividly illustrated in the case of antibiotic resistance. In a genetically identical population of bacteria, high doses of antibiotics kill most but not all organisms. Upon re-culturing, the surviving bacteria give rise to a population that is natively affected by antibiotics [78]. These resistant cells are thus phenotypic rather than genetic variants. Similar observations are made in the case of sporulation and DNA-uptake pathways, where pathway activity shows a significant variance across the population [77, 79]. While systems biology is still discovering the mechanistic bases of phenotypic variance [80], ESB provides a conceptual context in which to evaluate these findings [81–83]. In particular, a recent *in silico* study suggests that phenotypic switching in a gene-regulatory system can emerge as a byproduct of the evolution of elevated nonlinearity, which is itself selected for its impact on noise and evolvability [84]. Being able to examine phenotypic switching in laboratory-based experimental evolution [85] allows hypotheses about the evolution of noise to be tested (e.g. regarding the mutational steps toward the molecular mechanisms of noise).

## Broadening the scope of ESB

ESB is developing rapidly in several directions. While some of these initiatives advance established research areas, others set up new research agendas on the basis of methodological innovations.

## Expanding evolutionary and ecological frameworks

The cross-fertilization of different approaches in ESB is rebalancing the way in which evolution is understood to shape cellular networks. Rather than maintaining a division between adaptive and non-adaptive explanations, as often occurs in evolutionary research, ESB analyses are showing how both aspects are needed to explain evolved systems. Neutral processes can be understood to be co-dynamical with adaptive processes in network evolution [86]. For instance, in

the case of modularity, several studies indicate both adaptive and neutral scenarios to be of importance for modularity's emergence. These investigations indicate the environmental and evolutionary scenarios that can lead to the evolution of modularity: fluctuating selection in environments composed of modular tasks [21, 87], genetic drift [22], and stabilizing selection combined with selection for novel functions [88–90].

Although it will always be difficult to discern which plausible scenario has produced particular network features in specific organisms, *in silico* evolution and system-level mathematical models allow assessments of ecological and evolutionary conditions (such as fluctuating environments and population size) to be linked to network properties (and vice-versa). Rigorous evaluation of evolutionary scenarios can thereby incorporate comparative analyses of global network structures [15, 16, 54, 91], and together they can comprise system-based "reverse ecology" [92]. Reverse ecology in ESB aims to infer the historical ecologies of organisms from features of present-day cellular networks (rather than focusing on genetic loci as population genomics "reverse ecology" does). ESB aspires toward an increased ability to predict the future evolution of organisms (particularly single-celled organisms) in response to novel selective pressures and environments. This outcome is already within grasp in the case of metabolic networks [6].

## Encompassing intercellular network evolution

Selective pressures imposed by the abiotic environment on biological systems are only one component of ongoing evolutionary dynamics, which are always entwined with interspecies interactions (Fig. 1). While emerging systems-biological approaches are now beginning to scrutinize these interactions at the level of their underlying intracellular networks, one strand of ESB aims to understand more fully the effects of species interactions on network evolution. For example, the increasingly detailed characterization of parasite manipulation of host networks, at the level of single proteins and whole networks [93], has produced hypotheses about how such effects shape host network evolution [94]. Recent efforts to extend the application of FBA to species interactions do this by combining computational analysis with experiments [29, 30, 95]. These applications permit an examination of the environmental conditions and co-evolutionary dynamics that enable species to co-exist in syntropy.

Important open questions in this area of research include how to understand the evolution and dynamics of diverse microbial communities in the context of their intercellular networks, and how the network-level impact of host-parasite interactions can be captured more effectively. Taking an ESB approach, combined with the increasing ability to probe species interactions experimentally and genomically in natural contexts such as soil microcosms [96], expands the scope of how microbial communities are investigated. In addition, ESB studies that analyse cellular networks in light of social evolution can assess the impact of competitive or cooperative interactions on these networks, and in turn predict how they enable future competitive and cooperative dynamics [31].

## Re-engineering networks in light of evolution

A major future contributor to ESB is likely to be synthetic biology, which introduces an engineering approach to evolved systems. By accelerating the tinkering to which evolution has subjected biological systems, and incorporating mathematically redesigned features, synthetic biology can intervene in evolutionary processes and outcomes at the systems level. For example, synthetic circuits could be constructed to test hypotheses about intermediary states in the evolution of specific cellular networks, or to measure the effect of mutations in different genetic backgrounds.

Combining synthetic biology with experimental evolution could lead to an even greater capacity to test evolutionary hypotheses. Exposing synthetic circuits to further evolution under a variety of conditions (e.g. fluctuating versus stable environments), or developing high-throughput techniques to generate a diversity of circuits (e.g. [97]), allows the examination of both the effects of specific network features on future evolution and the role of environmental conditions on the outcome of network evolution. While recent experimental evolution has produced detailed characterizations of the molecular steps toward specific evolutionary trajectories and outcomes in natural systems (e.g. [98, 28]), combining experiments with synthetic biology and some of the theoretical approaches from ESB (e.g. dynamical network models, flux balance analysis) should increase research scope and depth.

## Integrating ESB and biomedicine

The improved ability to specify how and why networks have evolved will be particularly beneficial when the evolutionary outcomes relate to human health, as demonstrated by the investigation of the evolution of antibiotic resistance (e.g. [28]). So far, evolutionary medicine [99] and systems or network medicine [100] have interacted very little, and both are just nascent fields or even mere ambitions in the minds of some researchers. But it is very probable that many disease and treatment systems, from cancer to antibiotic resistance, are unlikely to be comprehended effectively or made tractable for intervention without being modeled as dynamic evolving systems [101, 102]. Using mathematical modeling and synthetic-biological constructions to predict, for example, how antibiotic resistance will evolve in relation to network vulnerabilities to specific molecular interactions, would have sustainable therapeutic effects [103]. Drug development, foundering in many traditional approaches (especially when restricted to the selection of target molecules), may thus yield to a synthesis of evolutionary and systems biological approaches because it steers drug development back to a focus on interactions within an evolving physiological context [104].

## Synthesizing explanation and prediction

We mentioned above that a central aim of ESB is to produce mechanistic and evolutionary understanding of genotype-

phenotype relations. Mechanistic explanations are causal accounts of how certain outcomes occur in specific circumstances. In the ESB case, the focus is on how particular genotypic features produce certain phenotypic properties. Historically, molecular biology has offered qualitative mechanistic explanations. One aspect of systems biology has worked to make these explanations quantitatively mechanistic rather than primarily descriptive-predictive. Dynamic mathematical modeling and time-series data are crucial to such explanations. ESB goes one step further and attempts to embed quantitative mechanistic explanation (which in Ernst Mayr's terminology [105] is proximate explanation), within ultimate or evolutionary explanation.

Prediction has sometimes been suggested to be an achievement beyond the capacity of evolutionary biology (e.g. [105, 106, 107]). Although evolutionary biology is accepted as explanatory, it has been placed at the ultimate end of the spectrum of explanation rather than the proximate, and the field has often been charged with untestability and panadaptationism. Systems biology on the other hand has focused largely on description and prediction (i.e. characterizing systems and predicting fluctuations of that system under specific perturbation regimes). The aim to achieve both ultimate and proximate explanation, combined with predictiveness, is a major motivation of ESB. This does not mean its practitioners expect to find universal laws. But they do expect to find regularities that hold across a wide range of networks and organisms as a result of common evolutionary processes, and there is good reason to think that finding these will help explain and predict system behavior. This is because evolutionary processes, despite being both stochastic and contingent, can leave stable traces of their impact on system architecture and function. These traces can then be linked back to evolutionary and ecological conditions. When this historical knowledge is combined with detailed mechanistic explanation, evolutionary trajectories can be projected forward and sought *in silico* or *in vitro*. Evolutionary predictions will never be easy to formulate and confirm, but ESB offers a platform for developing the predictive aspects of evolutionary biology and in the process, refining how evolutionary prediction is understood (because of the necessity of taking into account contingency and uncertainty).

Some of the basis for predictive ESB is already established in metabolic network research [6]. These predictions currently focus on detecting the effects of mutations on metabolic function, identifying environments from the structure and function of metabolic networks, and establishing causal connections between structural features of metabolic networks and ecological conditions. Extending this basic predictive capacity will mean going beyond metabolic networks to other cellular interactions, by building predictive models based on explanations of the origins and diversity of networks as they respond temporally to complex conditions.

## Conclusions and outlook

We have presented ESB as an emerging but loose synthesis of system approaches to biological phenomena

interpreted within an overarching evolutionary framework. However, there are numerous disciplinary and training barriers to be overcome to enable researchers to contribute to ESB. From the perspective of many quantitative and modeling-oriented biologists, evolutionary research consists of historical narratives that are not “hard” science. And from the evolutionary biologists’ perspective, their own field already has sufficiently rigorous methods at its disposal. ESB is in one sense a demonstration of the usefulness of thinking outside these disciplinary boxes – not to make researchers switch disciplinary allegiance but to enhance their own disciplinary achievements. For this reason, we advocate a more integrative systems biology curriculum – able to encompass evolutionary aspects of systems – at whatever educational levels system biology is taught.

Although we do not anticipate all evolutionary biologists becoming ESB advocates or all systems biologists signing up for evolutionary analyses, we do foresee that integrative ESB efforts will have increasing payoffs for difficult research questions. Describing, explaining, and predicting genotype-phenotype relationships will obviously need a great deal of experimental work to ground studies based on dynamic models and evolutionary simulations. However, if there is one thing molecular biology has learned in the last decade it is that integrative combinations of modeling and experimental approaches are crucial to more predictive and explanatory understandings of systems. All the caveats about modeling uncertainties, data shortcomings and experimental intractability that apply to systems biology apply to ESB. But as this review illustrates, many steps forward have been made and this is in part because the field is characterized more by research-driven inquiry than grand ambitions to create a new disciplinary entity.

We must emphasize that we are neither claiming that ESB is a panacea to disciplinary shortcomings, nor that it has all the right questions and answers. Instead, we are drawing attention to the broad phenomenon of ESB, rather than a specific strand here or there, because of the field’s general implications for the integration of certain modes of explanation and prediction. Even for scientists not interested in evolutionary or systems biology, these generalities are issues of scientific relevance. It may be the case that ESB becomes so broad in its problem orientation and methodology that it cannot be demarcated even as a field, let alone a discipline: it could simply become a general approach to the study of evolving biological systems. This, we believe, would be more valuable than if ESB were to form a tightly bounded field with disciplinary ambitions.

### Acknowledgements

We thank our referees and editor for very helpful comments, and several colleagues for comments on a much earlier version of this paper. MAO acknowledges support from the Australian Research Council and University of Sydney in the form of a Future Fellowship. We are grateful to Arno Steinacher for creating the two figures.

### References

1. Medina M. 2005. Genomes, phylogeny, and evolutionary systems biology. *Proc Natl Acad Sci USA* **102**: 6630–5.
2. Koonin EV, Wolf YI. 2006. Evolutionary systems biology: links between gene evolution and function. *Curr Opin Biotechnol* **17**: 481–7.
3. Dean AM, Thornton JW. 2007. Mechanistic approaches to the study of evolution: the functional synthesis. *Nat Rev Genet* **8**: 675–88.
4. O’Malley MA. 2012. Evolutionary systems biology: historical and philosophical perspectives on an emerging synthesis. In Soyer OS, ed; *Evolutionary Systems Biology*. NY: Springer. p. 1–28.
5. Soyer OS, ed. 2012. *Evolutionary Systems Biology*. NY: Springer.
6. Papp B, Notebaart RA, Pál C. 2011. Systems-biology approaches for predicting genomic evolution. *Nat Rev Genet* **12**: 591–602.
7. Loewe L. 2009. A framework for evolutionary systems biology. *BMC Syst Biol* **3**: 27.
8. Morange M. 1998. *A History of Molecular Biology*. Cambridge MA: Harvard University Press.
9. Boogerd FC, Bruggeman FJ, Hofmeyr J-HS, Westerhoff HV, eds; 2007. *Systems Biology: Philosophical Foundations*. Amsterdam: Elsevier.
10. Alon U, Surette MG, Barkai N, Leibler S. 1999. Robustness in bacterial chemotaxis. *Nature* **397**: 168–71.
11. Klipp E, Nordlander B, Krüger R, Gennemark P, et al. 2005. Integrative model of the response of yeast to osmotic shock. *Nat Biotechnol* **23**: 975–82.
12. Jaeger J, Surkova S, Blagov M, Janssens H, et al. 2004. Dynamic control of positional information in the early *Drosophila* embryo. *Nature* **430**: 368–71.
13. Gao F, Davidson EH. 2008. Transfer of a large gene regulatory apparatus to a new developmental address in echinoid evolution. *Proc Natl Acad Sci USA* **105**: 6091–6.
14. Hoyos E, Kim K, Milloz J, Barkoulas M, et al. 2011. Quantitative variation in autocrine signaling and pathway crosstalk in the *Caenorhabditis* vulval network. *Curr Biol* **21**: 527–38.
15. Nikolau E, Agrafioti I, Stumpf M, Quinn J, et al. 2009. Phylogenetic diversity of stress signalling pathways in fungi. *BMC Evol Biol* **9**: 44.
16. Wuichet K, Zhulin IB. 2010. Origins and diversification of a complex signal transduction system in prokaryotes. *Sci Signal* **3**: r50.
17. François P, Hakim V, Siggia ED. 2007. Deriving structure from evolution: metazoan segmentation. *Mol Syst Biol* **3**: 154.
18. Goldstein RA, Soyer OS. 2008. Evolution of taxis responses in virtual bacteria: non-adaptive dynamics. *PLoS Comput Biol* **4**: e1000084.
19. Thorne TW, Ho H-L, Huvet M, Haynes K, et al. 2011. Prediction of putative protein interactions through evolutionary analysis of osmotic stress response in the model yeast *Saccharomyces cerevisiae*. *Fungal Genet Biol* **48**: 504–11.
20. Alon U. 2007. Network motifs: theory and experimental approaches. *Nat Rev Genet* **8**: 450–61.
21. Kashtan N, Alon U. 2005. Spontaneous evolution of modularity and network motifs. *Proc Natl Acad Sci USA* **102**: 13773–8.
22. Lynch M. 2007. The evolution of genetic networks by non-adaptive processes. *Nat Rev Genet* **8**: 803–13.
23. Wagner A. 2008. Neutralism and selectionism: a network-based reconciliation. *Nat Rev Genet* **9**: 965–74.
24. Giaever G, Chu AM, Ni L, Connelly C, et al. 2002. Functional profiling of the *Saccharomyces cerevisiae* genome. *Nature* **418**: 387–91.
25. Orr HA. 2005. The genetic theory of adaptation: a brief history. *Nat Rev Genet* **6**: 119–27.
26. Akman OE, Rand DA, Brown PE, Millar AJ. 2010. Robustness from flexibility in the fungal circadian clock. *BMC Syst Biol* **4**: 88.
27. Feist AM, Herrgård MJ, Thiele I, Reed JL, et al. 2008. Reconstruction of biochemical networks in microorganisms. *Nat Rev Microbiol* **7**: 129–43.
28. Toprak E, Veres A, Michel J-B, Chait R, et al. 2012. Evolutionary paths to antibiotic resistance under dynamically sustained drug selection. *Nat Genet* **44**: 101–5.
29. Klitgord N, Segre D. 2010. Environments that induce synthetic microbial ecosystems. *PLoS Comput Biol* **6**: e1001002.
30. Bordbar A, Lewis NE, Schellenberger J, Palsson BØ, et al. 2010. Insight into human alveolar macrophage and *M. tuberculosis* interactions via metabolic reconstructions. *Mol Syst Biol* **6**: 422.
31. Foster KR. 2011. The sociobiology of molecular systems. *Nat Rev Genet* **12**: 193–203.
32. Emmerling M, Dauner M, Ponti A, Fiaux J, et al. 2002. Metabolic flux responses to pyruvate kinase knockout in *Escherichia coli*. *J Bacteriol* **184**: 152–64.

33. de Visser JA, Hermissen J, Wagner GP, Meyers LA, et al. 2003. Perspective: evolution and detection of genetic robustness. *Evolution* **57**: 1959–72.
34. Wagner A. 2000. The role of population size, pleiotropy and fitness effects of mutations in the evolution of overlapping gene functions. *Genetics* **154**: 1389–401.
35. Krakauer DC, Plotkin JB. 2002. Redundancy, antiredundancy, and the robustness of genomes. *Proc Natl Acad Sci USA* **99**: 1405–9.
36. Bergman A, Siegal ML. 2003. Evolutionary capacitance as a general feature of complex gene networks. *Nature* **424**: 549–52.
37. Azevedo RBR, Lohaus R, Srinivasan S, Dang KK, et al. 2006. Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. *Nature* **440**: 87–90.
38. Soyer OS, Pfeiffer T. 2010. Evolution under fluctuating environments explains observed robustness in metabolic networks. *PLoS Comput Biol* **6**: e1000907.
39. Harrison R, Papp B, Pál C, Oliver SG, et al. 2007. Plasticity of genetic interactions in metabolic networks of yeast. *Proc Natl Acad Sci USA* **104**: 2307–12.
40. Siegal ML, Bergman A. 2002. Waddington's canalization revisited: developmental stability and evolution. *Proc Natl Acad Sci USA* **99**: 10528–32.
41. Wang Z, Zhang J. 2009. Abundant indispensable redundancies in cellular metabolic networks. *Genome Biol Evol* **1**: 23–33.
42. Phillips PC. 2008. Epistasis – the essential role of gene interactions in the structure and evolution of genetic systems. *Nat Rev Genet* **9**: 855–67.
43. Wagner GP, Zhang J. 2011. The pleiotropic structure of the genotype-phenotype map: the evolvability of complex organisms. *Nat Rev Genet* **12**: 204–13.
44. Segre D, Deluna A, Church GM, Kishony R. 2005. Modular epistasis in yeast metabolism. *Nat Genet* **37**: 77–83.
45. Snitkin ES, Segre D. 2011. Epistatic interaction maps relative to multiple metabolic phenotypes. *PLoS Genet* **7**: e1001294.
46. Gutiérrez J. 2009. A developmental systems perspective on epistasis: computational exploration of mutational interactions in model developmental regulatory networks. *PLoS ONE* **4**: e6823.
47. Gjuvsland AB, Hayes BJ, Omholt SW, Carlberg O. 2007. Statistical epistasis is a generic feature of gene regulatory networks. *Genetics* **175**: 411–20.
48. Landry CA, Rifkin SA. 2012. The genotype-phenotype maps of systems biology and quantitative genetics: distinct but complementary. In Soyer OS, ed; *Evolutionary Systems Biology*. NY: Springer. p. 371–98.
49. Moore JH, Williams SM. 2005. Traversing the conceptual divide between biological and statistical epistasis: systems biology and a more modern synthesis. *BioEssays* **27**: 637–46.
50. Tyler AL, Asselbergs FW, Williams SM, Moore JH. 2009. Shadows of complexity: what biological networks reveal about epistasis and pleiotropy. *BioEssays* **31**: 220–7.
51. Wang Z, Liao BY, Zhang J. 2010. Genomic patterns of pleiotropy and the evolution of complexity. *Proc Natl Acad Sci USA* **107**: 18034–9.
52. Hill WG, Zhang ZS. 2012. On the pleiotropic structure of the genotype-phenotype map and the evolvability of complex organisms. *Genetics* **190**: 1131–7.
53. Knight CG, Zitzmann N, Prabhakar S, Antrobus R, et al. 2006. Unraveling adaptive evolution: how a single point mutation affects the protein coregulation network. *Nat Genet* **38**: 1015–22.
54. Huvet M, Toni T, Sheng X, Thorne T, et al. 2011. The evolution of the phage shock protein response system: interplay between protein function, genomic organization, and system function. *Mol Biol Evol* **28**: 1141–55.
55. Beltrao P, Trinidad JC, Fiedler D, Roguev A, et al. 2009. Evolution of phosphoregulation: comparison of phosphorylation patterns across yeast species. *PLoS Biol* **7**: e1000134.
56. Madan Babu M, Teichmann SA, Aravind L. 2006. Evolutionary dynamics of prokaryotic transcriptional regulatory networks. *J Mol Biol* **358**: 614–33.
57. Tan K, Shlomi T, Feizi H, Ideker T, et al. 2007. Transcriptional regulation of protein complexes within and across species. *Proc Natl Acad Sci USA* **104**: 1283–8.
58. Carroll SB. 2008. Evo-devo and an expanding evolutionary synthesis: a genetic theory of morphological evolution. *Cell* **134**: 25–36.
59. Peter IS, Davidson EH. 2011. Evolution of gene regulatory networks controlling body plan development. *Cell* **144**: 970–85.
60. Erwin DH, Davidson EH. 2009. The evolution of hierarchical gene regulatory networks. *Nat Rev Genet* **10**: 141–8.
61. Alonso CR, Wilkins AS. 2005. The molecular elements that underlie developmental evolution. *Nat Rev Genet* **6**: 709–15.
62. Jensen LJ, Jensen TS, de Lichtenberg U, Brunak S, et al. 2006. Co-evolution of transcriptional and post-translational cell-cycle regulation. *Nature* **443**: 594–7.
63. Chen K, Rajewsky N. 2007. The evolution of gene regulation by transcription factors and microRNAs. *Nat Rev Genet* **8**: 93–103.
64. Jaeger J, Irons D, Monk N. 2012. The inheritance of process: a dynamical systems approach. *J Exp Zool (Mol Dev Evol)* **318B**: 591–612.
65. Davidson EH. 2010. Emerging properties of animal gene regulatory networks. *Nature* **468**: 911–20.
66. Manu, Surkova S, Spirov AV, Gursky VV, et al. 2009. Canalization of gene expression in the *Drosophila* blastoderm by gap gene cross regulation. *PLoS Biol* **7**: e1000049.
67. Salazar-Ciudad I, Jernvall J. 2010. A computational model of teeth and the developmental origins of morphological variation. *Nature* **464**: 583–6.
68. Siegal ML, Promislow DE, Bergman A. 2007. Functional and evolutionary inference in gene networks: does topology matter? *Genetica* **129**: 83–103.
69. Fujimoto K, Ishihara S, Kaneko K. 2008. Network evolution of body plans. *PLoS ONE* **3**: e2772.
70. ten Tusscher KH, Hogeweg P. 2011. Evolution of networks for body plan patterning: interplay of modularity, robustness and evolvability. *PLoS Comput Biol* **7**: e1002208.
71. François P, Siggia ED. 2010. Predicting embryonic patterning using mutual entropy fitness and in silico evolution. *Development* **137**: 2385–95.
72. Eldar A, Elowitz MB. 2010. Functional roles for noise in genetic circuits. *Nature* **467**: 167–73.
73. Blake WJ, Balázsi G, Kohanski MA, Issacs FJ, et al. 2006. Phenotypic consequences of promoter-mediated transcriptional noise. *Mol Cell* **24**: 853–65.
74. Zhang Z, Qian W, Zhang J. 2009. Positive selection for elevated gene expression noise in yeast. *Mol Syst Biol* **5**: 299.
75. Tánase-Nicola S, ten Wolde PR. 2008. Regulatory control and the costs and benefits of biochemical noise. *PLoS Comput Biol* **4**: e1000125.
76. Wang Z, Zhang J. 2011. Impact of gene expression noise on organismal fitness and the efficacy of natural selection. *Proc Natl Acad Sci USA* **108**: E67–76.
77. Maamar H, Raj A, Dubnau D. 2007. Noise in gene expression determines cell fate in *Bacillus subtilis*. *Science* **317**: 526–7.
78. Balaban NQ, Merrin J, Chait R, Kowalik L, et al. 2004. Bacterial persistence as a phenotypic switch. *Science* **305**: 1622–5.
79. Veening JW, Stewart EJ, Berngruber TW, Taddei F, et al. 2008. Bet-hedging and epigenetic inheritance in bacterial cell development. *Proc Natl Acad Sci USA* **105**: 4393–8.
80. Levy SF, Ziv N, Siegal ML. 2012. Bet hedging in yeast by heterogeneous, age-correlated expression of a stress protectant. *PLoS Biol* **10**: e1001325.
81. Levy SF, Siegal ML. 2012. The robustness continuum. In Soyer OS, ed; *Evolutionary Systems Biology*. NY: Springer. p. 431–452.
82. Kussell E, Kishony R, Balaban NQ, Leibler S. 2005. Bacterial persistence: a model of survival in changing environments. *Genetics* **169**: 1807–14.
83. Salathé M, Van Cleve J, Feldman MW. 2009. Evolution of stochastic switching rates in asymmetric fitness landscapes. *Genetics* **182**: 1159–64.
84. Kuwahara H, Soyer OS. 2012. Bistability in feedback circuits as a byproduct of evolution of evolvability. *Mol Syst Biol* **8**: 564.
85. Beaumont HJ, Gallie J, Kost C, Ferguson GC, et al. 2009. Experimental evolution of bet hedging. *Nature* **462**: 90–3.
86. Wagner A. 2011. *The Origins of Evolutionary Innovations: A Theory of Transformative Change in Living Systems*. NY: Oxford University Press.
87. Parter M, Kashtan N, Alon U. 2008. Facilitated variation: how evolution learns from past environments to generalize to new environments. *PLoS Comput Biol* **4**: e1000206.
88. Soyer OS. 2007. Emergence and maintenance of functional modules in signaling pathways. *BMC Evol Biol* **7**: 205.
89. Hintze A, Adamcik C. 2008. Evolution of complex modular biological networks. *PLoS Comput Biol* **4**: e23.
90. Espinosa-Soto C, Wagner A. 2010. Specialization can drive the evolution of modularity. *PLoS Comput Biol* **6**: e1000719.
91. Borenstein E, Kupiec M, Feldman MW, Ruppin E. 2008. Large-scale reconstruction and phylogenetic analysis of metabolic environments. *Proc Natl Acad Sci USA* **105**: 14482–7.

92. Levy R, Borenstein E. 2012. Reverse ecology: from systems to environments and back. In Soyer OS, ed; *Evolutionary Systems Biology*. ed; NY: Springer. p. 329–45.
93. Bhavsar AP, Guttman JA, Finlay BB. 2007. Manipulation of host-cell pathways by bacterial pathogens. *Nature* **449**: 827–34.
94. Salathé M, Soyer OS. 2008. Parasites lead to evolution of robustness against gene loss in host signaling networks. *Mol Syst Biol* **4**: 202.
95. Hillesland KL, Stahl DA. 2010. Rapid evolution of stability and productivity at the origin of a microbial mutualism. *Proc Natl Acad Sci USA* **107**: 2124–9.
96. Gómez P, Buckling A. 2011. Bacteria-phage antagonistic coevolution in soil. *Science* **332**: 106–9.
97. Wang HH, Issacs FJ, Carr PA, Sun ZZ, et al. 2009. Programming cells by multiplex genome engineering and accelerated evolution. *Nature* **460**: 894–8.
98. Woods RJ, Barrick JE, Cooper TF, Shrestha U, et al. 2011. Second-order selection for evolvability in a large *Escherichia coli* population. *Science* **331**: 1433–6.
99. Williams GC, Nesse RM. 1991. The dawn of Darwinian medicine. *Q Rev Biol* **66**: 1–22.
100. Barabási AL, Gulbahce N, Loscalzo J. 2011. Network medicine: a network-based approach to human disease. *Nat Rev Genet* **12**: 56–68.
101. Gatenby RA, Gillies RJ, Brown JS. 2010. Evolutionary dynamics of cancer prevention. *Nat Rev Cancer* **10**: 526–7.
102. Baquero F, Coque TM, de la Cruz F. 2011. Ecology and evolution as targets: the need for novel eco-evo drugs and strategies to fight antibiotic resistance. *Antimicrob Agents Chemother* **55**: 3649–60.
103. Palmer AC, Kishony R. 2013. Understanding, predicting and manipulating the genotypic evolution of antibiotic resistance. *Nat Rev Genet* **14**: 243–8.
104. Pujol A, Mosca R, Farrés J, Aloy P. 2010. Unveiling the role of network and systems biology in drug discovery. *Trends Pharmacol Sci* **31**: 115–23.
105. Mayr E. 1961. Cause and effect in biology. *Science* **134**: 1501–6.
106. Peters RH. 1976. Tautology in evolution and ecology. *Am Nat* **110**: 1–12.
107. Gould SJ. 1990. *Wonderful Life: The Burgess Shale and the Nature of History*. London: Hutchinson Radius.

RESEARCH ARTICLE

Open Access

# Gene bionetworks involved in the epigenetic transgenerational inheritance of altered mate preference: environmental epigenetics and evolutionary biology

Michael K Skinner<sup>1\*</sup>, Marina I Savenkova<sup>1</sup>, Bin Zhang<sup>2</sup>, Andrea C Gore<sup>3</sup> and David Crews<sup>4</sup>

## Abstract

**Background:** Mate preference behavior is an essential first step in sexual selection and is a critical determinant in evolutionary biology. Previously an environmental compound (the fungicide vinclozolin) was found to promote the epigenetic transgenerational inheritance of an altered sperm epigenome and modified mate preference characteristics for three generations after exposure of a gestating female.

**Results:** The current study investigated gene networks involved in various regions of the brain that correlated with the altered mate preference behavior in the male and female. Statistically significant correlations of gene clusters and modules were identified to associate with specific mate preference behaviors. This novel systems biology approach identified gene networks (bionetworks) involved in sex-specific mate preference behavior. Observations demonstrate the ability of environmental factors to promote the epigenetic transgenerational inheritance of this altered evolutionary biology determinant.

**Conclusions:** Combined observations elucidate the potential molecular control of mate preference behavior and suggests environmental epigenetics can have a role in evolutionary biology.

**Keywords:** Epigenetics, Brain, Networks, Evolution, Behavior

## Background

The current molecular paradigm for neo-Darwinian evolutionary biology is that random DNA sequence mutations, gene flow, and genetic drift promote phenotype variation that allows an adaptation event to facilitate natural selection [1]. Although environment has an important role in natural selection, environmental factors generally do not have the capacity to alter DNA sequence or mutation rates. A small group of compounds can act as mutagens, but the vast majority of nutritionally derived and environmental toxicants do not alter DNA sequence [2]. The current genetic paradigm does not completely explain many observations such as rapid evolutionary events, environmental impacts on evolution, and the low frequency

of the occurrence of useful mutations [3,4]. The realization that epigenetics provides an additional molecular mechanism for the environment to influence genome activity and biology has suggested a potential role for environmental epigenetics in evolutionary biology [5-11]. Charles Darwin recognized sexual selection as one of two determinants in evolutionary biology, the other being natural selection [12]. The physical attributes and courtship rituals involved in mate preference are essential for reproductive fitness and propagation of a species. The current study examines how environmental factors can promote an epigenetic event to promote an alteration in mate preference behavior.

Previously we demonstrated that exposure of a gestating female rat to an environmental compound during fetal gonadal sex determination promoted epigenetic reprogramming of the male germline [13-15]. These reprogrammed differential DNA methylation regions (DMR) in the sperm epigenome have recently been shown to be induced by a

\* Correspondence: skinner@wsu.edu

<sup>1</sup>Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA

Full list of author information is available at the end of the article

variety of different environmental toxicants with exposure specific DMR [16,17]. The initial environmental compound used was the commonly used fungicide vinclozolin which is an anti-androgenic endocrine disruptor [18]. The primordial germ cells during migration down the genital ridge undergo a DNA methylation erasure that then upon gonadal sex determination the DNA re-methylation is initiated in a sex-specific manner [19]. Environmental exposures during this developmental stage modifies the epigenetic programming of the male germline that becomes re-programmed (imprinted-like) and promotes a transgenerational phenotypic variation and adult onset disease state in subsequent generations [13-16]. The epigenetic transgenerational inheritance of adult onset disease (i.e. after one year of age) in males includes infertility, prostate disease, kidney disease, immune abnormalities and spermatogenic defects [20,21], and in females includes mammary tumor development, kidney disease, reproductive tissue abnormalities and pregnancy abnormalities [22]. This germline mediated epigenetic transgenerational inheritance of adult onset disease is mediated in part through alterations in the sperm epigenome [13,14]. Since the germline establishes the base line epigenome of the organism, all tissues in both the female and male progeny including the brain appear to have altered tissue specific transgenerational epigenomes, transcriptomes and phenotypes [15,21,23-25].

Investigation of the epigenetic transgenerational inheritance of altered brain genome activity and behaviors previously demonstrated anxiety-like behavior increased in females and decreased in males, which correlate to alterations in specific brain region transcriptomes [23]. Altered stress responses are also detected in the transgenerational exposure lineage animals [26]. Interestingly, previous analysis of F3 generation control and vinclozolin lineage female and male rats (i.e. prior to the onset of disease) demonstrated an alteration in mate preference behavior [27]. The female rats, independent of control or vinclozolin lineage, prefer control lineage males if given a choice. This behavioral decision raises the possibility of an epigenetic contribution to mate preference and sexual selection. The current study was designed to directly correlate the altered mate preference behavior with gene networks in specific brain regions in both the females and males. Observations elucidate the potential molecular control of mate preference behavior and demonstrates environmental factors have the capacity to promote the epigenetic transgenerational inheritance of altered mate preference.

Systems biology analysis has allowed biological phenomena such as mate preference to be considered from the molecular to physiological level. The gene bionetwork [28] analysis previously developed to investigate the molecular basis of disease [29] was used in the current study.

This approach has been used to identify gene networks associated with disease, such as obesity and diabetes [30]. Recently, we have used this bionetwork analysis to study a normal developmental process of primordial follicle development in the ovary [31,32]. The gene networks identified were found to contain growth factors that are known to regulate the developmental process [31,32]. These bionetwork analyses use a large number of microarray transcriptome analyses under different perturbations to identify gene clusters and modules that are coordinately regulated [33,34]. The gene networks observed identify the genes with the highest level of integration and connection (i.e. connectivity) that associate with the phenotype [29,33-35]. This genomic approach was used in the current study to identify the gene bionetworks in various brain regions associated with mate preference.

Observations demonstrate an environmental compound exposure can induce an epigenetic reprogramming of the germline that promotes epigenetic transgenerational inheritance of altered mate preference behavior. Although no direct epigenetic modifications in the brain were examined, the environmentally induced epigenetic transgenerational model used indicates epigenetics can be involved in the induction of the altered behavioral phenotypes. Sex-specific effects were observed in both the male and female brain transcriptome and behavior correlations. The gene networks in specific brain regions that statistically correlate with various mate preference behaviors provides insight into this environmentally modified transgenerational behavior. This systems biology approach has elucidated novel mechanisms to be considered in mate preference biology.

## Results

The experimental design involved the development of transgenerational control and vinclozolin lineage animals for a mate preference behavioral analysis [27]. Subsequently, a transcriptome analysis was performed on 6 different brain regions from adult male and female F3 generation Sprague Dawley rats. These brain regions have previously been shown to be associated with mate preference behavior [27,36]. The transcriptome alterations were statistically correlated with changes in mate preference behaviors. As previously described [13,37], F0 generation gestating females were transiently exposed daily to vehicle control DMSO or vinclozolin from embryonic day 8–14 (E8-14) during fetal gonadal sex determination. The F1 generation offspring were bred at 90 days of age to generate F2 generation control and vinclozolin lineage progeny and then F2 generation animals were bred to generate the F3 generation control and vinclozolin lineage animals [13]. No sibling or cousin breeding was used to avoid any inbreeding artifacts. The F3 generation control and vinclozolin male and female rats were analyzed at 3–4 months

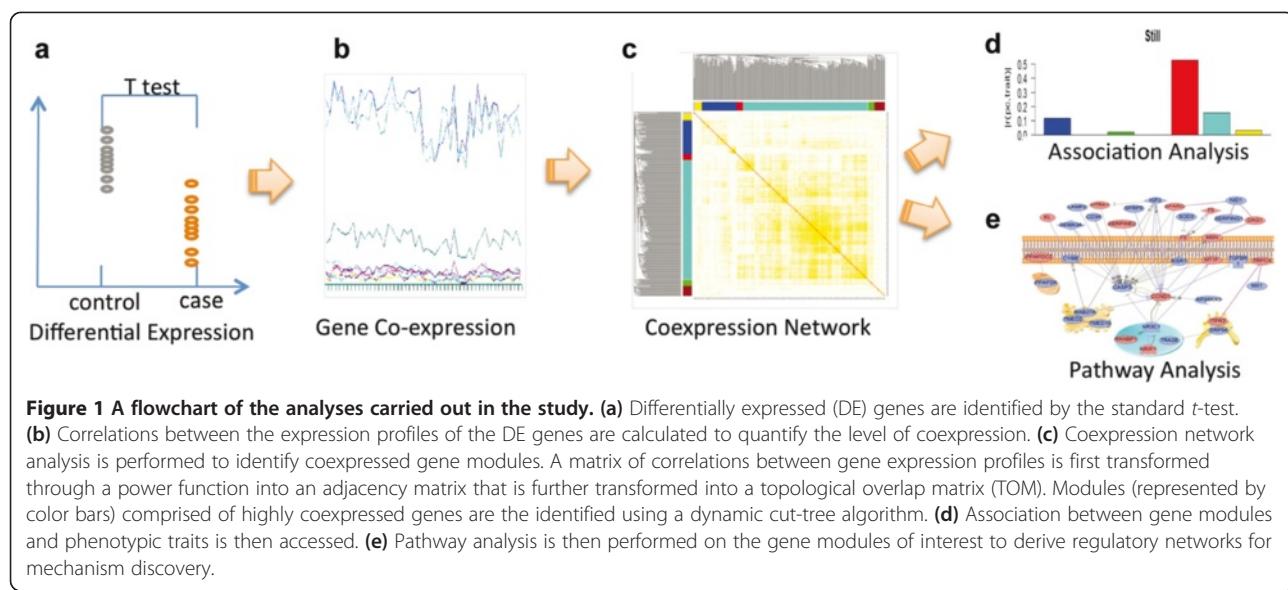
of age for mate preference behaviors, as previously described [27]. This is an age when no major adult onset disease has been detected or is anticipated [20]. Later at 11 month of age, animals were sacrificed and specific brain regions isolated and RNA collected for microarray transcriptome analysis. The differentially regulated gene sets ("Signature lists") for each brain region were identified. Subsequently a bioinformatics bionetwork analysis [23,31] was used to correlate gene modules and networks with mate preference behaviors observed (Figure 1).

Females of both control and vinclozolin lineages preferred control lineage males over vinclozolin lineage males [27]. The mate preference behaviors are described in detail in the Methods [27] (Additional file 1: Movie S1). The specific behavioral components associated with mate preference analysis include the following: "Wire Mesh" involved the experimental animal investigating the stimulus animals directly through the Wire Mesh; "Facial Investigation" entailed the actual nose-to-nose contact; "Plexiglas" refers to the experimental animal investigating the area immediately bordering the Wire Mesh that separated the experimental animal from the stimulus animal; "Walking" refers to general investigation of the central testing chamber as measured by undirected walking and sniffing; and "Still" in which the animal was stationary with minimal head movement. Additional file 2: Table S1A presents the values for each behavioral component associated with individual animals. After completion of the behavioral analysis the animals were sacrificed at 11 months of age and 6 different brain regions dissected including the amygdala (Amy), hippocampus (Hipp), olfactory bulb (OlfB), cingulate cortex (CngCtx), entorhinal cortex (EnCtx), and pre-optic area-anterior hypothalamus (POAH). The procedure to isolate the brain regions is described in the Materials

and Methods. The isolated tissue was immediately placed in Trizol reagent, frozen and stored. RNA was prepared for microarray transcriptome analysis from each animals brain regions independently.

For the microarray analysis each F3 generation control and vinclozolin lineage male and female animal had six different brain regions analyzed which totaled 134 different microarrays. The microarray data were pre-processed and demonstrated two abnormal arrays that were omitted for further analysis (Additional file 3: Figure S1B). Batch effect corrections were made for RNA preparation date and array scan date with no major batch effects detected. The array data were then processed as previously described [31] to identify the differentially expressed gene sets for each brain region (Table 1). The differentially expressed genes in the Signature lists required a greater than 1.2 fold change in expression and all changes in expression were statistically significant with  $p < 0.05$ , as described in the Methods. Since a 20% alteration in gene expression for many genes, such as transcription factors, can have dramatic cellular and biological responses [26,32], a more stringent cut off (e.g. 2 $\times$ ) was not used in the current study. In the current study the primary focus was on the coexpression patterns of the differentially expressed genes through the co-expression network analysis.

The number of control lineage versus vinclozolin lineage differentially expressed genes in the Signature lists ranged from 43 to 803 with both up-regulated and down-regulated genes (Table 1). The total number of control versus vinclozolin lineage differentially expressed genes for all brain regions combined was 1833 for females and 1693 for males. A list of all the genes separated by brain region, sex and functional gene categories is presented in Additional file 4: Table S2 A-I. The overlap and differences between the



**Table 1 Differentially expressed Signature genes and their overlap with modules generated in combined network**

Sex-region	Signature lists	Over-expressed	Under-expressed	Combined networks modules										Separate network modules					
				Number of modules	Turquoise	Blue	Brown	Yellow	Green	Red	Black	Pink	Magenta	Number of modules	Turquoise	Blue	Brown	Yellow	Green
Female regions	1833	939	894	4	1090	283	259	104											# genes in module
					# genes overlapped between module & Signature list														
F-Amy	139	38	101		81	19	15	0						3	71*	49	17		
F-CngCTX	803	481	322		640	16	57	83						4	444	183	82	74	
F-EnCTX	433	279	154		56	191	128	37						3	369*	35	10*		
F-Hipp	70	40	30		18	17	12	0						1	70				
F-OlfB	748	221	527		598	59	69	3						4	416	305	11	10*	
F-POAH	56	24	32		16	18	11	1						1	56*				
Male regions	1693	638	1055	9	505	287	222	155	88	66	50	40	36						# genes in module
					# genes overlapped between module & Signature list														
M-Amy	175	105	70		35	19	15	3	25	8	10	8	13	2	160*	10*			
M-CngCTX	785	189	596		354	79	193	39	56	8	0	5	0	1	780*				
M-EnCTX	385	210	175		87	133	26	8	13	3	0	5	22	1	378*				
M-Hipp	151	30	121		13	27	0	11	2	41	9	21	0	2	133*	13*			
M-OlfB	356	278	78		71	33	0	114	16	9	47	7	9	6	231	65*	20*	12*	11*
M-POAH	43	19	24		3	13	1	4	1	0	1	7	0	1	43*				11

\*- modules that showed statistically significant correlation with behavior.

Signature lists of each brain region for male and female is shown in Figure 2. The majority of genes were distinct to the different lists in a comparison of the brain regions. The one exception was an overlap between the cingulate cortex (CngCtx) and olfactory bulb (OlfB) in the female. Therefore, each brain region Signature list was distinct from each other and between the sexes.

Analysis of the cellular pathways and processes associated with the gene Signature lists for each brain region is shown in Figure 2 and Additional file 5. The top 36 pathways with the greatest combined number of genes associated are shown in Additional file 5. An extended list of pathways and processes with the associated genes from the different modules and tissues is presented in Additional file 6: Table S3. Several of the most highly represented pathways in the male and female were the MAPK signaling pathway, olfactory transduction, neuro-active ligand-receptor interactions and axon guidance. The Signature list genes distributed relatively evenly across the different pathways with no major over-representation identified. Most major cellular processes and pathways were represented with no major predominance of any individual specific regulatory mechanism observed (Additional file 6: Table S3).

A bionetwork cluster analysis was performed on the differentially expressed genes in the various brain regions as previously described [29,31] to identify gene modules and networks with coordinated and interconnected relationships (i.e. connectivity) [38] (Figure 1). Initially all the differentially expressed genes in the combined brain regions for male or female were analyzed, termed combined networks (Figure 3a). This combined analysis was performed to potentially identify common gene networks or

modules similar among all the brain regions that potentially correlate with the mate preference behavior parameters. The increased number of microarrays and data associated with the combined analysis also improves the power of the cluster and network analysis. The gene cluster analysis is shown and individual modules of genes identified are presented in different colors on the axis. The module colors represent increasing levels of connectivity [38] with white being negligible and red being highest. The combined Signature lists provided 4 modules in the female and 9 gene modules in the male. This can be seen as the blocked gene clusters designated as modules of different colors (Figure 3a). The number of genes in each module for male and female brain regions is shown in Table 1. This combined network analysis and modules were correlated with the mate preference behavior, but no significant correlations were found using this combined analysis (Additional file 7: Table S4).

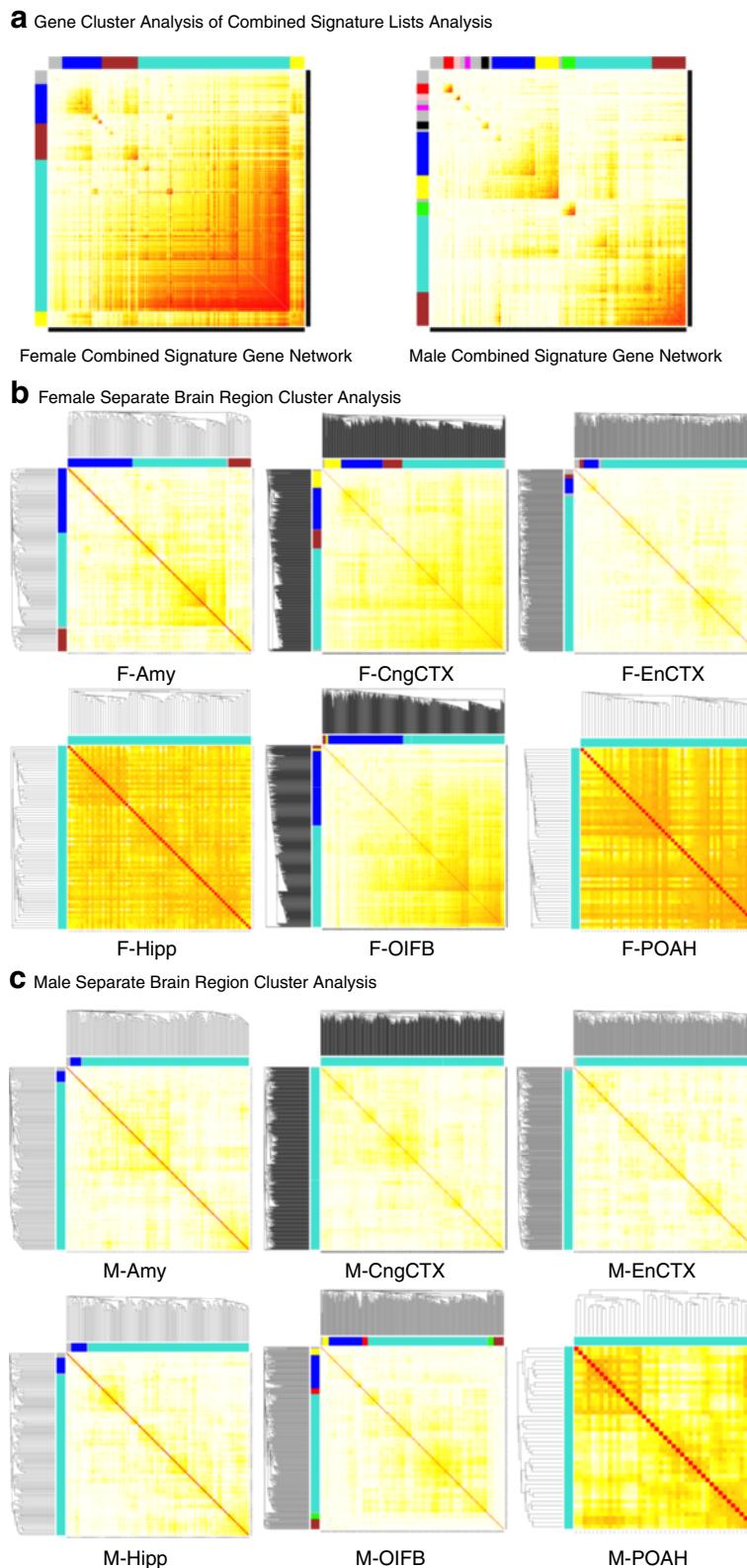
Due to the distinct functions of each brain region and distinct gene Signature lists, the combined analysis was found not to provide the specificity needed to identify the behavioral correlations with gene modules. Therefore, a more specific network analysis using the individual Signature lists for each brain region separately was performed, termed separate networks. Each brain region differentially expressed gene Signature list was used for separate network analysis. The cluster analysis identified specific gene modules for each brain region from the male and female gene Signature lists presented (Figure 3b, c). The modularity for the specific brain regions was not as strong as the combined region analysis. Each separate brain region is shown and the gene modules are identified by the different colors. The brain regions had 1–6 different modules

Brain Regions	Genes Number in Signature List	Number of KEGG Pathways	6 Female Regions	F-Amy	F-CngCTX	F-EnCTX	F-Hipp	F-OlfB	F-POAH	6 Male Regions	M-Amy	M-CngCTX	M-EnCTX	M-Hipp	M-OlfB	M-POAH
<b>6 Female Regions</b>	<b>1833</b>	<b>161</b>	<b>139</b>	<b>803</b>	<b>433</b>	<b>70</b>	<b>748</b>	<b>56</b>	<b>230</b>	<b>36</b>	<b>93</b>	<b>46</b>	<b>24</b>	<b>50</b>	<b>13</b>	
F-Amy	139	18	18		42	6	1	20	2	23	5	9	5	2	4	2
F-CngCTX	803	108	108	18		45	1	245	7	63	2	33	15	8	7	4
F-EnCTX	433	126	126	12	90		9	31	11	83	19	34	18	7	21	3
F-Hipp	70	34	34	9	20	28		9	3	11	2	3	0	1	6	1
F-OlfB	748	123	123	16	89	98	26		9	82	12	29	13	9	21	2
F-POAH	56	47	47	9	35	42	14	44		13	3	2	3	3	3	5
<b>6 Male Regions</b>	<b>1693</b>	<b>161</b>	<b>161</b>	<b>18</b>	<b>108</b>	<b>126</b>	<b>34</b>	<b>123</b>	<b>47</b>	<b>175</b>	<b>785</b>	<b>385</b>	<b>151</b>	<b>356</b>	<b>43</b>	
M-Amy	175	55	55	12	43	47	18	48	30	55		27	11	11	19	1
M-CngCTX	785	153	153	17	106	122	32	119	47	153	53		36	1	51	6
M-EnCTX	385	150	150	18	104	123	33	115	47	150	52	147		6	12	4
M-Hipp	151	121	121	16	86	100	28	95	45	121	47	118	119		22	3
M-OlfB	356	142	142	14	93	114	34	109	45	142	51	135	134	119		5
M-POAH	43	13	13	2	9	11	2	7	5	13	4	13	13	11	13	

Enriched KEGG Pathways Overlap

Signature Lists Genes Overlap

**Figure 2 Number of overlapped differentially expressed genes with pathways and Signature lists.** Number of genes overlapped between Signature lists is shown in regular font on grey background; number of affected KEGG pathways overlapped between Signature lists is shown in italicized font in white background; only KEGG pathways with 5 or more genes affected are counted.



**Figure 3** (See legend on next page.)

(See figure on previous page.)

**Figure 3 Male and female brain region bionetwork cluster analysis and corresponding gene modules.** Topological overlap matrixes of the gene co-expression network consisting of genes differentially expressed in F3 generation vinclozolin lineage rat brain as compared to F3 generation lineage Control animals. Genes in the rows and columns are sorted by an agglomerative hierarchical clustering algorithm. The different shades of color signify the strength of the connections between the nodes (from white signifying not significantly correlated to red signifying highly significantly correlated). Modules identified are colored along both column and row and are boxed. (a) Matrixes of combined network for male and female brain regions. (b) Matrixes of separate network for female amygdala (F-Amy), cingulate cortex (F-CngCTX), enterohinal cortex (F-EnCTX), hippocampus (F-Hipp), olfactory bulbs (F-OlfB), and preoptic area-anterior hypothalamus (F-POAH). (c) Matrixes of separate network for male amygdala (M-Amy), cingulate cortex (M-CngCTX), enterohinal cortex (M-EnCTX), hippocampus (M-Hipp), olfactory bulbs (M-OlfB), and preoptic area-anterior hypothalamus (M-POAH).

and associated gene networks (Table 1). The same differentially expressed gene Signature lists were used, but the network analysis was from the separate lists (Figure 3b and c). All subsequent analyses used the gene modules from this region specific network analyses.

The cluster analysis (Figure 3b and c) for each brain region provided modules of genes with coordinated gene expression and identified a connectivity index [29-31,33,38] for each associated gene. The connectivity index (k.in) for each of the differentially expressed genes in each region is presented in Additional file 4: Table S2. The top 10% of genes with the highest connectivity index for each of the gene modules was identified (Additional file 4: Table S2 as the genes in bold font). From this combined list of 185 genes for male and 225 genes for female, a gene sub-network analysis was performed. The most highly interconnected genes in all modules for female and male brain regions were used to identify the common direct connection interactions between genes in a gene sub-network (Figure 4). The female gene sub-network identified angiogenesis, growth and apoptosis as predominant cellular processes affected (Figure 4b). The male gene sub-network identified apoptosis as a predominant pathway affected (Figure 4a). These gene networks identify the common connections within the brain regions with the most highly interconnected genes differentially expressed between the control and vinclozolin F3 generation animals.

Analysis of the gene networks for each individual brain region gene module demonstrated that only males exhibited direct connection gene sub-networks for gene modules (male amygdala and cingulate cortex turquoise modules) (Figure 5). This region specific examination of gene network modules demonstrated most regions did not have direct connection sub-networks, but indirect interactions with various pathways and processes. An alternative analysis used the entire Signature list for each brain region to identify region specific gene sub-networks (Additional file 8: Figure S2 A-K). How these gene networks may correlate with the alterations in mate preference behavior required a statistical correlation of the gene sets with the behaviors (Figure 1).

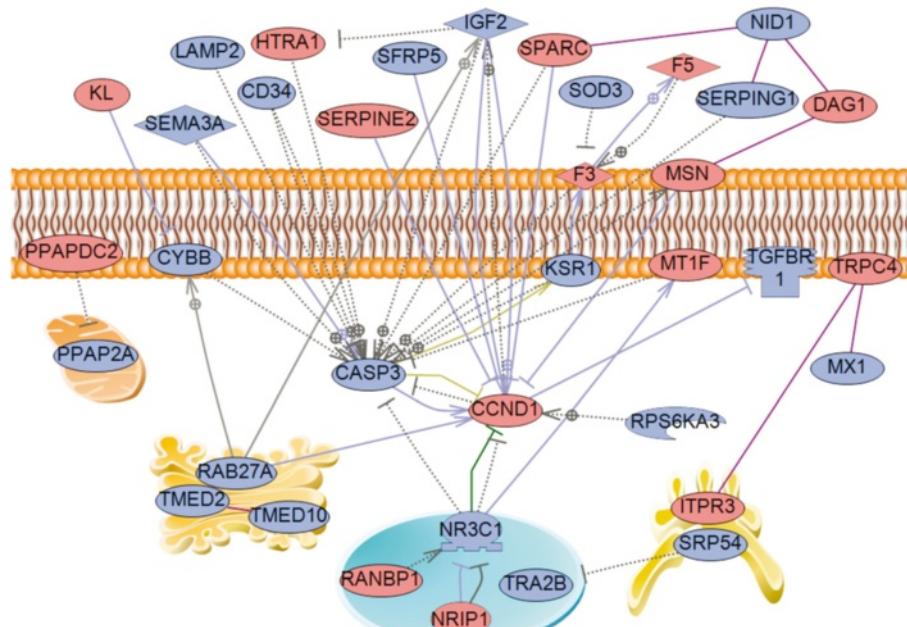
In considering the mate preference behaviors, the female is the discriminatory sex to choose a mate, while the male is non-discriminatory and has phenotypes and

behaviors to be selected [27,39]. The altered gene expression and correlations with behaviors needs to consider this in data interpretation. The behavioral parameters (Additional file 2: Table S1) for the mate preference analysis were statistically correlated to the separate network gene modules for the different brain regions, (Additional file 9: Table S5). The correlation and the p-values associated with the statistical correlation coefficients are presented. All correlations with a single or multiple principle component comparison are presented. Considering a  $p < 0.05$  or correlation coefficient  $> 0.5$  and  $p = 0.05-0.1$  between the gene module and behavior demonstrated correlations in four female brain regions and six modules with the female behavior (Table 2 and Additional file 9: Table S5). A summary of the statistically significant correlations and/or those with strong correlation coefficients is shown in Figure 6. Nearly all the female brain regions had statistically significant correlation with the Plexiglas behavioral parameter. The female amygdala (F-Amy) had a turquoise module with significant correlation with the Walking and Still parameters. The turquoise modules of female entorhinal cortex (F-EnCtx) had a strong correlation with the Wire Mesh and Plexiglas behavioral parameters (Figure 6).

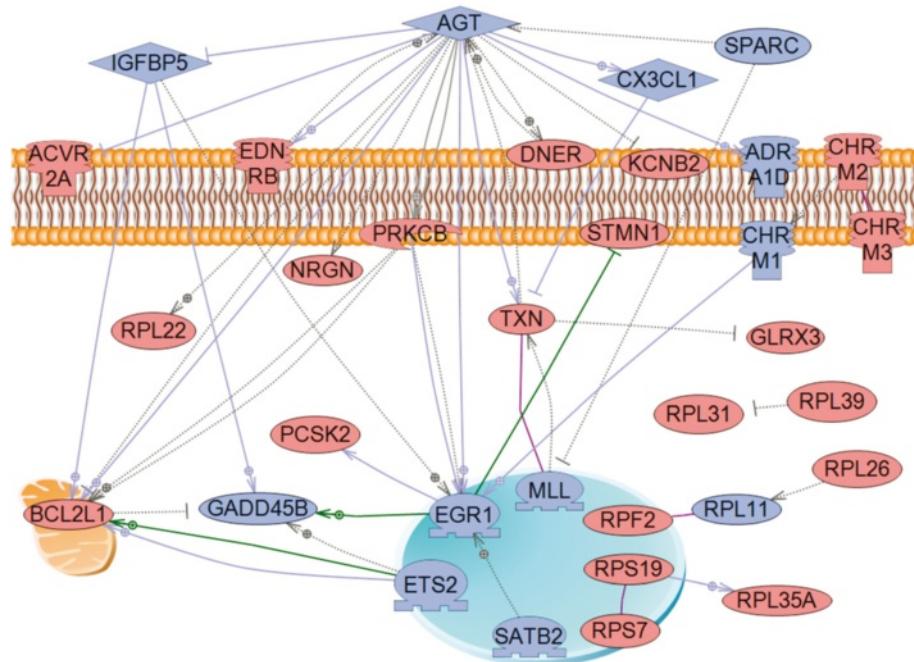
The six male brain regions and associated gene modules had a number of statistically significant correlations with the mate preference parameters (Table 2 and Additional file 9: Table S5). All the male brain regions had statistically significant correlation with at least one module and the Wire Mesh behavioral parameter. Amy and Hipp also had correlations with the behavioral Plexiglas parameter (Figure 6). Therefore, at least one gene module in nearly all brain regions statistically correlated to the mate preference parameters analyzed. These correlations can now be considered in regards to the regulatory roles of gene networks identified for mate preference behavior alterations for the female (chooser) versus the male (selected) (Figure 5 and Additional file 8: Figure S2).

The direct connection gene sub-networks for the critical male amygdala (M-Amy) and cingulate cortex (M-CngCtx) turquoise modules are shown in Figure 5. The Signature list for each brain region sub-networks demonstrate distinct networks for each region (Additional file 8: Figure S2). Since nearly all the brain regions and key modules (Figure 6) have a statistically significant correlation with the Wire

**a Male Direct Connection Gene Sub-Network**



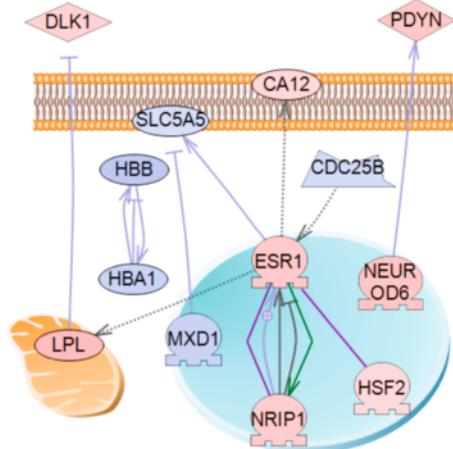
**b Female Direct Connection Gene Sub-Network**



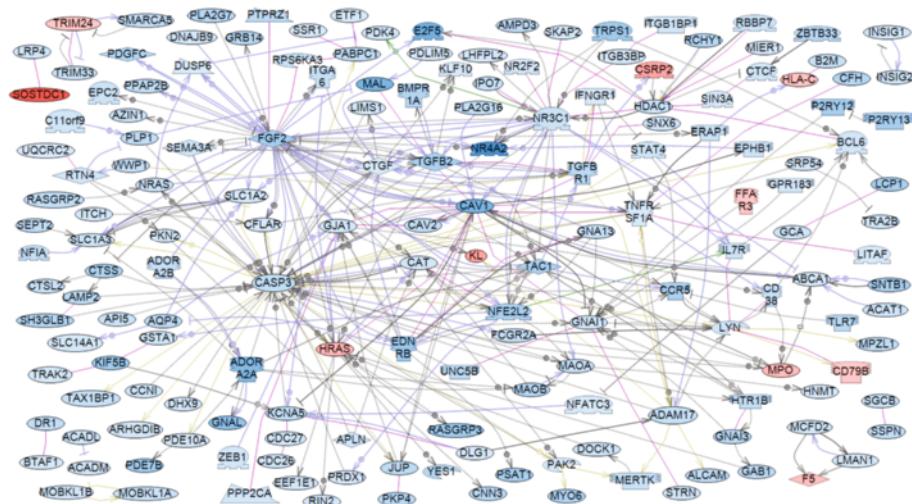
**Figure 4** Direct connection sub-networks for most highly top 10% connected genes from each module of separate network for male

**(a) or female (b).** Only directly connected genes are shown according to their location in the cell (on membrane, in Golgi apparatus, nucleus, cytoplasm or outside the cell). Node shapes and color code: oval and circle – protein; diamond – ligand; circle/oval on tripod platform – transcription factor; ice cream cone – receptor; crescent – kinase or protein kinase; irregular polygon – phosphatase; red color indicates up-regulated genes, blue – down-regulated. Arrows with plus sign show positive regulation/activation, arrows with minus sign – negative regulation/inhibition; grey arrows represent regulation, lilac - expression, purple – binding, green – promoter binding, and yellow – protein modification.

**a Male Amygdala Turquoise Module**



**b Male Cingulate Cortex Turquoise Module**



**Figure 5 Direct connection sub-networks for male amygdala turquoise module (a) and male cingulate cortex turquoise module (b).**  
 Only directly connected genes are shown. Node and arrows shapes and color code is the same as for Figure 4.

Mesh for male or Plexiglas for female mate preference behavior parameters, the combined gene sub-network (Figure 4) for all male or female brain regions identifies a potentially associated molecular control of behavior. Alternately, the analysis of separate regions differentially expressed gene sets (Signature lists) identified distinct gene sub-networks that associate with the different regions (Additional file 8: Figure S2). These potential gene sub-networks correlate and potentially regulate the mate choice behavior for the female and selection behavior/phenotype for the male. In addition to the gene networks, a correlation of critical cellular pathways in specific brain regions and modules (Additional file 5) that are associated with the epigenetic transgenerational inheritance of altered mate preference behavior.

The epigenetic transgenerational inheritance of the altered mate preference behavior requires the transmission of an altered epigenome in the germline (sperm) [5,13,15]. Previously the altered DNA methylation of the F3 generation sperm was characterized with 48 differentially DNA methylated regions (DMR) being identified in gene promoters [14]. These vinclozolin induced sperm DMR are in part what promotes an altered epigenome in the embryo and all developing tissues transgenerationally [15]. Although any developing tissue (e.g. brain) will have a dramatic cascade of epigenetic and genetic steps to achieve an adult fully differentiated state [40,41], the possibility that some of the original germline epigenetic marks (DMR) may persist was investigated. The genes associated with the 48 previously identified sperm DMR

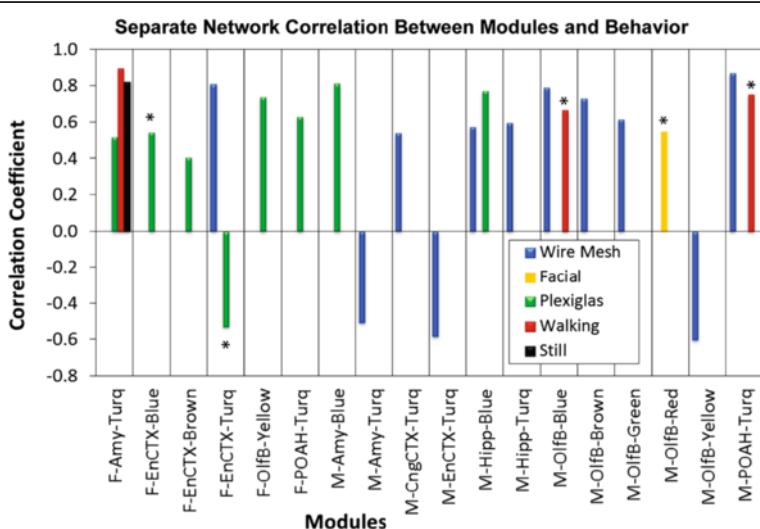
**Table 2 Gene modules highly correlated to different mate preference behavior parameters**

Sex-region	Behavior trait		Wire mesh		Facial		Plexiglas		Still		Walking	
	Module	# PC*	Correlation	p-value	Correlation	p-value	Correlation	p-value	Correlation	p-value	Correlation	p-value
F-Amy	Turquoise	2					0.52	<b>0.028</b>	0.82	<b>0.012</b>	0.90	<b>0.001</b>
F-EnCTX	Blue	1					0.55	<b>0.066</b>				
	Brown	1					0.41	<b>0.037</b>				
	Turquoise	3	0.81	<b>0.029</b>			-0.53	<b>0.083</b>				
F-OlfB	Yellow	2					0.74	<b>0.044</b>				
F-POAH	Turquoise	1					0.63	<b>0.0386</b>				
M-Amy	Blue	2					0.81	<b>0.008</b>				
	Turquoise	1	-0.51	<b>0.022</b>								
M-CngCTX	Turquoise	1	0.54	<b>0.048</b>								
M-EnCTX	Turquoise	1	-0.58	<b>0.037</b>								
M-Hipp	Blue	1	0.57	<b>0.023</b>			0.77	<b>0.017</b>				
	Turquoise	1	0.60	<b>0.034</b>								
M-OlfB	Blue	2	0.79	<b>0.012</b>							0.67	<b>0.072</b>
	Brown	1	0.73	<b>0.012</b>								
	Green	1	0.62	<b>0.020</b>								
	Red	1			0.55	<b>0.081</b>						
	Yellow	1	-0.60	<b>0.033</b>								
M-POAH	Turquoise	1	0.87	<b>0.002</b>							0.753	<b>0.091</b>

\*- number of principal components (PC) used to calculate correlation between modules and behavior.

were compared with the male and female brain region gene Signature lists. The comparison demonstrated the majority of the DMR did not correspond to differentially expressed genes in various brain regions. Only Rnase1 in the male Amy, Ig6-2a in the male EnCtx, Parp9 in the female CngCtx and Rp132 in the female OlfB

overlapped. Interestingly, a copy number variation (CNV) in the Fam111a site previously identified [14] was found in all brain regions identified with the epigenome analysis. This provides a positive control for the technology and ability to detect the DMR. Therefore, some of the original sperm DMR programmed sites may



**Figure 6 Gene module correlation with mate preference behavior parameters.** Separate network male and female modules highly correlated to types of behavior: Wire Mesh (blue bars), Facial Investigation (orange), Plexiglas (green), Walking (red) and Still (black). Bars not marked with asterisks have p-value < 0.05; bars marked with one asterisk have correlation coefficient >0.5 and > p-value = 0.1 - 0.05.

persist, but the vast majority of brain development and epigenetic programming, and potential distal regularity role of DMR in epigenetic control regions [25], is distinct from the original germline epigenetic marks.

## Discussion

A systems biology analysis of environmentally induced epigenetic transgenerational inheritance of altered mate preference behavior was performed to suggest a potential role for epigenetics in evolutionary biology. Previous research has demonstrated that environmental toxicants such as the fungicide vinclozolin can promote a reprogramming of the germline epigenome during fetal gonadal sex determination that then transmits altered phenotypes and adult onset disease states transgenerationally in the absence of future environmental exposure [41]. This is referred to as epigenetic transgenerational inheritance [5,15] and suggests a role for environmental epigenetics in the inheritance of phenotypic variation and disease, independent of classic genetic inheritance mechanisms. The basic molecular mechanism involved in this non-genetic form of inheritance is the ability of environmental factors to influence the epigenetic programming of the germline [15,19]. The primordial germ cells during migration down the genital ridge undergo an erasure of DNA methylation that then is initiated to re-methylate at the time of gonadal sex determination in a sex-specific manner [19]. An environmental toxicant such as vinclozolin appears to alter gonadal development to influence germline DNA methylation programming [42] and the differential DNA methylation regions (DMR) in the sperm become imprinted-like sites that appear to not get erased at fertilization so are transmitted to subsequent generations and male and female progeny [13-16]. In addition to vinclozolin, a number of other environmental toxicants such as the plastic compound bisphenol A (BPA) [16,43], dioxin [16,44], methoxychlor [13], phthalates [16], pesticides [16], hydrocarbons [16], and DDT [17] have been shown to induce transgenerational phenotypes. Other environmental factors such as nutrition and stress can also promote transgenerational phenotypes [45-48].

The vinclozolin induced epigenetic transgenerational phenotypes previously identified included adult onset rat disease after 12 months of age of male infertility, mammary tumors, prostate disease, kidney disease and immune abnormalities [20]. Therefore, the mate preference analysis was performed prior to adult onset disease to remove the disease as a confounding factor. A brain-behavior transgenerational phenotype observed was increased female anxiety and decreased male anxiety behaviors [23]. This transgenerational anxiety behavior was also examined on a molecular level to identify brain region specific changes in different gene expression and gene networks associated with the behavior [23]. Similar observations were made in

the analysis of transgenerational stress responses [26]. Interestingly, in a previous study we found that vinclozolin induced alterations in mate preference behavior [27]. Females from either control or vinclozolin F3 generation lineages prefer control lineage males over vinclozolin lineage males, whereas no altered mate preference in males was observed [27]. While the standard argument would be that the females are the discriminating sex and distinguish between males on as yet undetermined phenotype characteristic(s), it is important to realize that the absence of evidence (in the male) does not mean the evidence of absence of male involvement since preference is only the first step in a mating sequence. That is, under natural circumstances this is followed by a mutual decision. Mating in rodents involves pheromone and auditory cues produced by both sexes and evidence suggests (see below) that it is under such unfettered conditions that the complementarity of behavior and brain are expressed.

Observations from the current study need to consider the effects on the female brain as potentially altering female discrimination and preference. The effects on the male brain are presumed to be associated with the characteristics (e.g. auditory cues and pheromone production) being selected. This altered mate preference behavior suggests the existence of an environmentally altered epigenetic transgenerational inheritance of mate preference behavior [15]. The current study was designed to identify the gene bionetworks in various male and female brain regions that correlate with the behavior of the transgenerational inheritance model.

A novel gene bionetwork analysis was developed to identify gene networks correlated to disease [29]. The approach was to use a large number of microarrays to identify transcriptomes in specific tissues associated with control versus disease individuals in large cohorts. Differentially regulated genes that are coordinately regulated and having connectivity [38] are clustered in large gene sets to identify modules of genes that associate with the disease [29,30,33-35] (Figure 1). More recently, we have used a similar approach to investigate a normal development process to identify gene bionetworks associated with development [31,32]. The primordial follicle development in the ovary was investigated to identify a network of growth factors and associated signaling systems that regulate follicle development [31,32]. This bioinformatics approach to identify regulatory gene networks was used in the current study to correlate brain gene networks to mate preference behavior (Figure 1) in an epigenetic transgenerational model [27]. The six different brain regions isolated from F3 generation control and vinclozolin lineage females and males were used in a microarray analysis to determine the differential gene expression in each brain region. The region specific gene sets, "Signature list", and associated gene networks were investigated.

Analysis of significant pathways and cellular processes potentially influenced by the differentially expressed gene sets and networks did not identify predominant or over represented pathways. One pathway identified that previously has been shown to correlate with sexual selection is the olfactory transduction pathway [49-52]. However, most of the brain regions and specific networks or gene modules affected similar pathways with overlap between most. Therefore, no specific pathways were identified and most major pathways were influenced. A limitation in gene expression studies is that individual genes are assigned a specific function, but the gene may be involved in functional categories. This limitation needs to be considered in any gene expression data interpretation, but genome wide transcriptome analysis for gene sets has been shown to start to address this issue [53]. Combined observations suggest that the distinct differential expressed gene sets for the different brain regions appear to regulate common cellular processes and pathways among the brain regions and modules.

In contrast, analysis of gene networks identified unique gene sub-networks and gene modules associated with each brain region differentially expressed gene sets. The specific inter-connected genes were unique and overall networks of connected genes distinct. Therefore, the different functions associated with each brain region and associated with the altered mate preference behavior were identified. A statistical correlation of the gene modules for each brain region

with the different mate preference parameters measured identified a number of statistically significant correlations. All but one female brain region (Hipp) had statistically significant correlations with the Plexiglas behavior parameter. The female behavior directly associates with the discrimination and mate preference choice. Interestingly, all the male brain regions had some modules with statistically significant correlations with the Wire Mesh behavior parameter. The male behavior and/or phenotype (e.g. pheromonal production) associates with the selected behaviors and characteristics of the non-discriminant sex. Therefore, direct correlations with the gene modules, specific brain regions and mate preference behavior parameters were identified. The specific gene modules and behavioral parameters statistically correlated were distinct between the sexes and brain regions, but strong correlations of the gene networks to the mate preference behavior was established. Interestingly, both the Plexiglas and Wire Mesh are indicators of interest and assessment of the stimulus animal.

The gene bionetwork analysis and statistical correlation with the mate preference behavior provides insight into the molecular basis of how various male and female brain regions correlate and in part control the various behavioral parameters. Observations provide one of the first genomic and systems biology analysis of mate preference behavior (Figures 1 and 7). The experimental model used

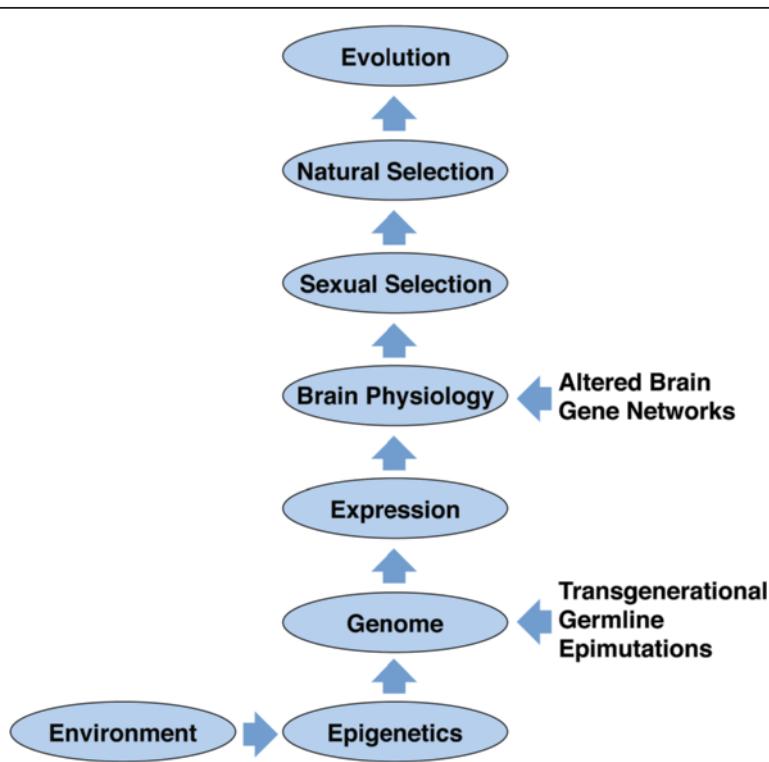


Figure 7 Schematic of role of epigenetics in evolution.

involved the ability of an environmental compound (vinclozolin) to induce an epigenetic modification of the germline (sperm) to promote epigenetic transgenerational inheritance of an altered mate preference behavior. This altered mate preference behavior was due to a baseline alteration in the epigenomes of all male and female tissues, including the brain, which are derived from the epigenetically altered germline [25]. The current study used a systems biology approach to help elucidate the molecular control of this process.

Darwin [12] considered natural selection and sexual selection as distinct processes driving the evolution of traits. Natural selection results in traits that are adaptive responses to changes in the environment. The resulting variation in traits between and within species is shaped by differential survivorship. In other words, animals that survive are those with traits that are adaptive to their environment [54]. Darwin [12] conceived of sexual selection as arising from aggressive interactions between males (male-male competition) and the female's selection of a mate (mate choice). Males compete amongst themselves for access to females. Aggression between males can have a direct effect on female reproduction by preventing other breeding males from having access to females or from harming the female. Importantly aggression amongst males can also have an indirect effect by inhibiting or suppressing the normal reproductive physiology of the female or even terminate a pregnancy [55].

Mate preference, in its simplest form, states that males compete for females and females choose between them. Although most research has focused on how females choose males, male choice of females is also important [56-58]. This point cannot be overemphasized. That is, in virtually all paradigms published to date, the choosing individual is the independent variable and the stimulus animal is the dependent variable. Although this study is similar to previous studies in that there are restrained stimulus animals and freely moving individuals that are the investigators, it differs in several important ways. First, the 'round robin' testing method insured that all males and all females served both as stimulus animals and experimental animals; thus, the "preferences" exhibited reflect both the males and the females. Second, this study extends to the molecular level events in the brain of the individuals, both of which have exhibited preferences, namely brain transgenerational transcriptome alterations that correlated to the opposing sex behaviors. In essence we are seeing the outcome of the complementary nature of mate choice. This has never been shown before.

Evolution favors reproductive success, and it is in the individual's interest to focus on selecting the best mate and to avoid mating with the wrong species [59,60]. Making the correct choice of a mate has a pronounced

impact on reproductive success of both partners. Except in unusual systems, in nature the mating partners choose one another [55,57,60]. Experiments with flies [61], birds [62], and rodents [63,64] indicate that individuals who are allowed to select, and be selected by, their mate enjoy greater reproductive success than force-paired animals. This consent is based not only on the internal milieu that motivates each individual to seek a partner, but also on the satisfactory nature of the phenotypic traits the potential mate displays.

There are a number of sexual selection hypotheses, all of which emphasize that females choosing optimally will produce young whose viability and survivorship are enhanced by the female's choice of mate [65]. The most attractive, and one that takes into account that mating is a cooperative act that involves both partners, is the sensory exploitation hypothesis [66]. This hypothesis postulates that males have evolved calls and/or pheromones to exploit the preexisting sensory biases in the female that themselves evolved for reasons independent of female choice. Male behavior then has changed to maximize stimulation of the female's sensory systems. Therefore, the coordination of the complementary signal and receiver, mounting and lordosis, coordination of egg and sperm maturation and release is required for successful completion for reproduction. These complementary processes are evident at all levels of biological organization [55,67] and we extend it here to the level of the genome and epigenome.

## Conclusions

The ability of an environmental factor to alter mate preference behavior suggests a critical role of environment in evolutionary biology. This is distinct from the generally accepted role of the environment in natural selection where environment is the active factor in the selection of an adaptive phenotype, but alternately here it involves the induction of phenotypes that can be acted on by natural selection. Since the majority of environmental factors can not alter DNA sequence or promote mutagenesis [2], an additional molecular mechanism to consider involves environmental epigenetics [15]. Many environmental compounds and factors such as nutrition can modify the epigenome to alter phenotypic variation. The role of epigenetics in evolutionary biology has been suggested previously [5,8-11,15,68], but no significant experimental evidence has been provided. The current study demonstrates an environmental factor can promote the epigenetic transgenerational inheritance of an altered mate preference behavior. The epigenetic modification of the germline (sperm) has been previously established [13,14,16] and will lead to epigenetic alterations in the brain transcriptomes of both females and males [23] to alter the mate preference behavior [27]. Therefore, the current study provides direct experimental evidence for a potential role of environmental

epigenetics in evolution by regulating a critical determinant such as mate preference on a molecular level (i.e. altered gene networks) in specific brain regions in a sex-specific manner (Figure 7). Although no direct epigenetic alterations were examined in the brain, the germline (sperm) epigenetic alterations that generate this altered male or female brain development have been documented [13,14,16]. This molecular mechanism does not suggest genetics will not have a critical role in evolutionary biology, but suggests environmental epigenetics will be an additional mechanism to consider. Epigenetics provides a mechanism for the environment to impact phenotypic variation and natural selection. Epigenetic and genetic mechanisms will cooperate to regulate on a molecular basis evolutionary biology. This appears to be a “neo-Lamarckian concept to facilitate neo-Darwinian evolution” [40,41].

The systems biology approach used in the current study links an environmental exposure, epigenetic transgenerational inheritance and molecular regulation of brain function to mate preference and evolutionary biology (Figure 7). Epigenetics will have a central role in how environmental factors influence how the gene networks emerge to induce phenotypic variation. Although genetics is critical for all aspects of biology, epigenetics provides the plasticity to allow the environment to alter biological events. This type of systems approach to understand complex biological traits, such as sexual selection, provides insights into how the various components (environment, phenotype and evolution) interact in a systems biology manner.

## Methods

### Animal housing protocol

Male and female rats of the F3 generation of Vinclozolin (Vinclozolin-Lineage) and DMSO Control (Control-Lineage) Lineages were selected out of litters from untreated F2 generation mothers in Dr. Michael Skinner's laboratory at Washington State University according to established protocols [13]. Briefly, gestating female F0 generation Sprague Dawley rats were injected with the fungicide vinclozolin (100 mg/kg) daily during fetal gonadal sex determination (E8-E14) and the F1 generation were bred to generate the F2 generation and then the F2 bred to generate the F3 generation [13]. At approximately PND 10 (before weaning), each animal was injected with a small microchip (AVID Identification system Inc. Norco, CA) subcutaneously between the shoulder blades. The animals were then shipped to the University of Texas from Washington State University on postnatal day (PND) 22, one day after weaning. Upon arriving at the University of Texas, one animal from each Lineage (Control and Vinclozolin) was pair-housed (one control and one vinclozolin animal) and remained in these dyads throughout the duration of the study. Because of the natural variation in dates of breeding, there was a 4-day spread of

birth date of animals in the first cohort but in the second cohort, all animals were born on the same day. However, all pair-housed animals were no more than one day apart in birth age and were paired randomly to prevent an age effect on cagemates.

Each dyad of animals was randomly placed in a six-wide, five-high metal housing rack in standard translucent polycarbonate rat cages (46 × 24 × 20.5 cm) with *ad libitum* access to tap water and standard rat chow (Purina rodent chow #5LL2 Prolab RMH 1800 diet). The animal room was on a 14:10 light/dark schedule. For environmental enrichment, a 7 cm diameter PVC pipe was placed in each cage.

### Ethics statement

All experimental protocols for the procedures with rats were pre-approved by the Washington State University Animal Care and Use Committee (IACUC approval # 02568-026) and by the University of Texas at Austin Animal Care and Use Committee (Public Health Service Animal Welfare Assurance Number A4107-01).

### Behavioral testing

After habituation to the testing arena, each individual was tested individually (when used as an experimental subject) or in pairs (when used as stimulus animals) with all individuals; the order of the testing was rotated during the course of both the male and female trials. All tests were conducted during the dark phase of the light cycle, beginning at 1200 h, 4 h after the progesterone injection, in a room illuminated with low levels of red light. Before trials, to confirm that females were receptive, each female was placed with a sexually experienced but otherwise experimentally naïve male; all females exhibited robust lordosis (arched back and lifted head posture) in response to mounting by the male.

Partner preference tests consisted of placing an individual (male or female) in the center of a large three-chamber glass-testing arena (122 × 46 × 54 cm). At either end was a small compartment (28 × 28 × 12.5 cm) containing the stimulus rats separated by a Wire-mesh barrier to allow exchange of olfactory, visual, and tactile cues. The area directly in front of the stimulus cage was marked by tape. Tests were conducted 2 h after the onset of the dark cycle under red-light illumination and lasted 10 min; all tests were videotaped for further review and analysis. At the end of each test, all animals were removed, and the entire testing arena was washed with a household cleaner and then wiped down with 70% ethanol to remove scent marks and residual odors. All males were tested with both types of females as stimulus animals (72 trials), and all females were tested with both types of males as stimulus animals (72 trials) (Movie S1).

The videotaped trials were analyzed by using JWwatcher v1.0 ([www.jwatcher.ucla.edu](http://www.jwatcher.ucla.edu)) computer software to quantify the behavior of each experimental animal. Time spent with a stimulus animal was recorded as soon as all four paws of the experimental animal crossed over the line of tape marking the boundary of that stimulus animal's compartment. As soon as one paw crossed over the tape back into the center compartment, the time recorded with the experimental animal was stopped. Preference behaviors were defined as those directed to the stimulus animal and included time spent in contact with the Wire Mesh separating the experimental and stimulus animal (Wire Mesh), during which the animals often touched noses through the Wire Mesh (facial investigation), and contacted the Plexiglas surface surrounding the front of the stimulus cage; the cumulative total time in these preference behaviors toward each stimulus animal was also calculated (Total). Other activity measured included undirected walking and sniffing (walking), standing still with minimal head movement (still). Videos demonstrating the test can be viewed as supporting information (Movie S1).

### Brain processing

The brain was removed in less than 1 minute and placed in crushed ice to chill. The brain was then cut in half in the sagittal plane along the midline. In all cases but one the right side was blocked and then 6 areas dissected (see list below) within 3–5 min. This procedure was done on iced tissue. The dissected brain areas were placed in chilled Trizol (150 l) in 1.5 ml Eppendorf tubes according to manufacturers specifications in each tube. No tissue fragment was more than 3 mm but in those instances, but where there were multiple fragments the amount of Trizol was doubled (approximately). After all animals were dissected, the Eppendorf tubes were vortexed (15 sec) and then frozen on dry ice. The brain regions collected were according to Paxinos & Watson [69]: olfactory bulbs (OlfB); cingulate cortex (CngCtx), anterior to POAH (Bregma 4.7 to 1.7); preoptic area-anterior hypothalamus (POAH), 4 mm rostral to anterior commissure (AC) (Bregma -0.26 to -1.40); amygdaloid nuclei (Amy), 3 mm caudal to AC (Bregma -2.3 to -3.6); hippocampus (Hipp), 6 mm caudal to AC (Bregma -2.12 to -4.52); entorhinal cortex (EnCtx) (Bregma -5.60 to -7.80).

### RNA preparation

Brain area samples from individual rats were homogenized in 150  $\mu$ l Trizol and then 600  $\mu$ l Trizol was added to final volume of 750 ml. Samples were stored at -80 or -20°C until RNA extraction. For microarray analysis, from 4 to 6 biological replicas (animals) were prepared as above for each brain area Control or Vinclozolin group depending on samples availability (Additional file 2: Table S1B). A total of 132 (67 Control and 65 Vinclozolin)

samples/chips were analyzed: (6 brain areas)  $\times$  (2 Male or Female)  $\times$  (2 Control or Vinclozolin)  $\times$  (4–6 biological replicas). RNA from individual animal brain area was extracted from Trizol samples according to standard Trizol extraction protocol (Invitrogen, USA) and stored in aqueous solution at -80°C until microarray analysis.

### Microarray analysis

The microarray analysis was performed by the Genomics Core Laboratory, Center for Reproductive Biology, Washington State University, Pullman, WA using standard Affymetrix reagents and protocol. Briefly, mRNA was transcribed into cDNA with random primers, cRNA was transcribed, and single-stranded sense DNA was synthesized which was fragmented and labeled with biotin. Biotin-labeled ssDNA was then hybridized to the Rat Gene 1.0 ST microarrays containing more than 30,000 transcripts (Affymetrix, Santa Clara, CA, USA). Hybridized chips were scanned on Affymetrix Scanner 3000. CEL files containing raw data were then pre-processed and analyzed with Partek Genomic Suite 6.5 software (Partek Incorporated, St. Louis, MO) using an RMA, GC-content adjusted algorithm. Raw data pre-processing was performed in 12 groups, one for each male or female brain area. Comparison of array sample histogram graphs for each group showed if data for all chips were similar and appropriate for further analysis (Additional file 3: Figure S1). By this criterion, 2 microarray samples (not counted in Additional file 4: Table S2B and not shown on Additional file 3: Figure S1) were omitted from repeated group pre-processing and further analysis.

The microarray quantitative data involves signals from an average of 28 different oligonucleotides (probes) arrayed for each transcript and many genes are represented on the chip by several transcripts. The hybridization to each probe must be consistent to allow a statistically significant quantitative measure of resulting gene expression signal. Therefore, the microarray provides an unbiased and highly stringent quantitative procedure compared to other protocols [70]. In contrast, a quantitative PCR procedure uses only two oligonucleotides and primer bias is a major factor in this type of analysis. Therefore, we did not attempt to use PCR based approaches as we feel the microarray analysis is more accurate and reproducible without primer bias such as PCR based approaches [31].

All microarray CEL files from this study have been deposited with the NCBI gene expression and hybridization array data repository GEO (GEO series accession number: GSE33830) and can be also accessed through [www.skinner.wsu.edu](http://www.skinner.wsu.edu). For gene annotation, Affymetrix annotation file RaGene1\_0stv1.na31.rn4.transcript.csv was used.

### Network analysis

The network analysis was restricted to genes differentially expressed between the control and the treatment

groups based on previously established criteria: (1) fold change of group means  $\geq 1.2$  or  $\leq 0.83$ ; (2)  $T$  test p-value  $\leq 0.05$ . The union of the differentially expressed genes from the different treatments resulted in 1,693 genes for males and 1833 for females being identified and used for constructing a weighted gene co-expression network [71,72]. Unlike traditional un-weighted gene co-expression networks in which two genes (nodes) are either connected or disconnected, the weighted gene co-expression network analysis assigns a connection weight to each gene pair using soft-thresholding and thus is robust to parameter selection. The weighted network analysis begins with a matrix of the Pearson correlations between all gene pairs, then converts the correlation matrix into an adjacency matrix using a power function  $f(x) = x^\beta$ . The parameter  $\beta$  of the power function is determined in such a way that the resulting adjacency matrix (i.e., the weighted co-expression network) is approximately scale-free. To measure how well a network satisfies a scale-free topology, we use the fitting index proposed by Zhang & Horvath [71] (i.e., the model fitting index  $R^2$  of the linear model that regresses  $\log(p(k))$  on  $\log(k)$  where  $k$  is connectivity and  $p(k)$  is the frequency distribution of connectivity). The fitting index of a perfect scale-free network is 1.

To explore the modular structures of the co-expression network, the adjacency matrix is further transformed into a topological overlap matrix [73]. As the topological overlap between two genes reflects not only their direct interaction, but also their indirect interactions through all the other genes in the network. Previous studies [71,73] have shown that topological overlap leads to more cohesive and biologically meaningful modules. To identify modules of highly co-regulated genes, we used average linkage hierarchical clustering to group genes based on the topological overlap of their connectivity, followed by a dynamic cut-tree algorithm to dynamically cut clustering dendrogram branches into gene modules [74]. Such networks were generated from all combined male or female differentially expressed genes (2 combined networks) or from each individual male or female brain region Signature lists (12 separate networks). From one to ten modules were identified in combined or separate networks and the module size was observed to range from 10 to 780 genes (Table 1).

To distinguish between modules, each module was assigned a unique color identifier, with the remaining, poorly connected genes colored grey. The hierarchical clustering over the topological overlap matrix (TOM) and the identified modules is shown (Figure 1). In this type of map, the rows and the columns represent genes in a symmetric fashion, and the color intensity represents the interaction strength between genes. This TOM heatmap highlights that genes in the transcriptional network fall into distinct network modules, where genes within a given module are more interconnected with each other

(blocks along the diagonal of the matrix) than with genes in other modules. Therefore, there are two types of global connectivity, adjacency-based one and TO based one. The adjacency-based connectivity ( $k.all$ ) is defined as the sum of the power-function transformed correlations between the gene  $g$  and all the other genes in the whole network while the TO-based connectivity ( $to.all$ ) is defined as the sum of the topological overlaps between the gene  $g$  and all the other genes. By default, connectivity used throughout the paper refers to TO-based connectivity  $to.all$ .

**Gene Co-expression Network Analysis Clarification:** Gene networks provide a convenient framework for exploring the context within which single genes operate. Networks are simply graphical models comprised of nodes and edges. For gene co-expression networks, an edge between two genes may indicate that the corresponding expression traits are correlated in a given population of interest. Depending on whether the interaction strength of two genes is considered, there are two different approaches for analyzing gene co-expression networks: 1) an unweighted network analysis that involves setting hard thresholds on the significance of the interactions, and 2) a weighted approach that avoids hard thresholds. Weighted gene co-expression networks preserve the continuous nature of gene-gene interactions at the transcriptional level and are robust to parameter selection. An important end product from the gene co-expression network analysis is a set of gene modules in which member genes are more highly correlated with each other than with genes outside a module. Most gene co-expression modules are enriched for GO functional annotations and are informative for identifying the functional components of the network that are associated with disease [75].

This gene co-expression network analysis (GCENA) has been increasingly used to identify gene sub-networks for prioritizing gene targets associated with a variety of common human diseases such as cancer and obesity [38,76-79]. One important end product of GCENA is the construction of gene modules comprised of highly interconnected genes. A number of studies have demonstrated that co-expression network modules are generally enriched for known biological pathways, for genes that are linked to common genetic loci and for genes associated with disease [33,38,71,75-78,80,81]. In this way, one can identify key groups of genes that are perturbed by genetic loci that lead to disease, and that define at the molecular level disease states. Furthermore, these studies have also shown the importance of the hub genes in the modules associated with various phenotypes. For example, GCENA identified ASPM, a hub gene in the cell cycle module, as a molecular target of glioblastoma [78] and MGC4504, a hub gene in the unfolded protein response module, as a target potentially involved in susceptibility to atherosclerosis [77].

## Pathway analysis

Resulting lists of differentially expressed genes for each male or female brain area as well as for each module generated in the combined network and some generated in separate networks analysis were analyzed for KEGG (Kyoto Encyclopedia for Genes and Genome, Kyoto University, Japan) pathway enrichment using Pathway-Express, a web-based tool freely available as part of the Onto-Tools (<http://vortex.cs.wayne.edu>) [82] as well as KEGG website ‘Search Pathway’ tool ([http://www.genome.jp/kegg/tool/search\\_pathway.html](http://www.genome.jp/kegg/tool/search_pathway.html)). Global literature analysis of various gene lists was performed using Pathway Studio 8.0 software (Ariadne Genomics, Inc., Rockville, MD).

## Additional files

**Additional file 1: Movie S1.** The first 19 sec of a 10-min mate-preference trial is shown. The trial is conducted under dim red light during the nocturnal (active) phase of the rats' light cycle. At the beginning of the video, the male is in the center of the chamber. The chamber is demarcated into thirds by tape on its floor. A stimulus female can be seen at the far end of the apparatus; the other stimulus female is not visible due to the position of camera. The stimulus females are free-moving in their chambers, but they are separated from the male by a wire mesh that is bounded by Plexiglas barrier. This enables the animals to communicate by olfactory, pheromonal, or behavioral cues, but physical interaction is limited to touching across the wire mesh. The trial begins with the removal of a holding box that confines the male. The male can be seen moving into the zone in front of one stimulus female and then moving across the central portion of the cage to the other stimulus female (out of sight). Several behaviors of the male can be seen on the video such as sniffing, facial investigation, walking, and standing of the female. The male is also seen investigating the various parts of the chamber, including the wire mesh, surrounding Plexiglas partition, and the glass walls of the chamber. Behaviors were scored for each male toward each pair of opposite lineage (Control- or Vinclozolin-Lineage) stimulus females.

**Additional file 2: Table S1.** Behavior and Sample Information.

**Additional file 3: Figure S1A.** Samples Histograms After Pre-processing (Male). **Figure S1B.** Samples Histograms After Pre-Processing (Female). Figure S1. Sample histograms and box plots for male (S1A) female (S1B) microarray signal values after pre-processing with RMA, GCcontent adjusted algorithm. Plots for F3 generation control (red) and F3 generation vinclozolin (blue) chips for female amygdala (A), cingulate cortex (B), enterorhinal cortex (C), hippocampus (D), olfactory bulbs (E), and preoptic areaanterior hypothalamus (F).

**Additional file 4: Table S2.** Genes Differentially Expressed in F3 Generation Vinclozolin Versus Control Lineage Male and Female Rat Brain Regions.

**Additional file 5: Top cellular pathways affected by signature gene lists and chosen modules from separate networks.**

**Additional file 6: Table S3.** Pathways Affected Male and Female Brain Region Signature Gene Lists and Chosen Modules from Separate Networks.

**Additional file 7: Table S4.** Correlation between combined network modules and behavior trait for F3-Vinclozolin rat brain regions.

**Additional file 8: Figure S2.** (Color) Brain Region Specific Signature List Direct Connection Gene Sub-Networks. Legend: Figure S2. Direct connection sub-networks for signature lists: female amygdala (A), female preoptic area-anterior hypothalamus (B), female hippocampus (C), female enterorhinal cortex (D), female cingulate cortex (E), female olfactory bulbs (F), male amygdala (G), male hippocampus (H), male cingulate cortex (I), male enterorhinal cortex (J), male olfactory bulbs (K) obtained by global literature analysis using Pathway Studio 8.0 software (Ariadne Genomics, Inc., Rockville, MD). Numbers in brackets on figures subtitles indicate number of genes in the list. Only directly connected genes are shown. Some sub-networks (G, H,

J) show gene location in the cell (on membrane, in Golgi apparatus, nucleus, cytoplasm or outside the cell). Node shapes and color code: oval and circle – protein; diamond – ligand; circle/oval on tripod platform – transcription factor; ice cream cone – receptor; crescent – kinase or protein kinase; irregular polygon – phosphatase; red color indicates up-regulated genes, blue – down-regulated. Arrows with plus sign show positive regulation/activation, arrows with minus sign – negative regulation/inhibition; grey arrows represent regulation, lilac – expression, purple – binding, green – promoter binding, and yellow – protein modification.

**Additional file 9: Table S5.** Correlation between separate network modules and behavior trait for F3-Vinclozolin rat brain regions.

## Abbreviations

DMR: Differential DNA methylation regions; E8-E14: Embryonic day 8–14; CngCtx: Cingulate cortex; OlfB: Olfactory bulb; k.in: Connectivity index; Amy: Amygdala; EnCtx: Entorhinal cortex; Hipp: Hippocampus; POAH: Preoptic area-anterior hypothalamus; BPA: Bisphenol A.

## Competing interests

The authors declare no conflicts of interest.

## Authors' contributions

MKS designed the study; MS BZ ACG DC performed the experiments; MKS and DC wrote the paper; MKS MS BZ ACG DC edited the manuscript. All authors read and approved the final manuscript.

## Acknowledgments

We thank Dr. M. Manikkam and Ms. R. Tracey for technical assistance and H. Johnson for assistance in preparation of the manuscript. We thank Drs. C. Guerrero-Bosagna and E. Nilsson for critical review of the manuscript. Research was supported in part by NIH grants ES 017538 (Crews) and ES 012974 (Skinner).

## Author details

<sup>1</sup>Center for Reproductive Biology, School of Biological Sciences, Washington State University, Pullman, WA 99164-4236, USA. <sup>2</sup>Department of Genetics & Genomic Sciences, Institute of Genomics and Multiscale Biology, Mount Sinai School of Medicine, New York, NY 10029, USA. <sup>3</sup>Pharmacology and Toxicology, Austin, Texas. <sup>4</sup>Section of Integrative Biology, University of Texas at Austin, Austin, TX 78712, USA.

Received: 24 October 2013 Accepted: 28 April 2014

Published: 16 May 2014

## References

1. Charlesworth B, Charlesworth D: **Darwin and genetics.** *Genetics* 2009, **183**(3):757–766.
2. Stein RA: **Epigenetics and environmental exposures.** *J Epidemiol Community Health* 2012, **66**(1):8–13.
3. Kalinowski ST: **Evolutionary and statistical properties of three genetic distances.** *Mol Ecol* 2002, **11**(8):1263–1273.
4. Wilkins AS: **Genetic networks as transmitting and amplifying devices for natural genetic tinkering.** *Novartis Found Symp* 2007, **284**:71–86. discussion 86–79, 110–115.
5. Jirtle RL, Skinner MK: **Environmental epigenomics and disease susceptibility.** *Nat Rev Genet* 2007, **8**(4):253–262.
6. Crews D, McLachlan JA: **Epigenetics, evolution, endocrine disruption, health, and disease.** *Endocrinology* 2006, **147**(6 Suppl):S4–S10.
7. Damiani G: **The Yin and Yang of anti-Darwinian epigenetics and Darwinian genetics.** *Riv Biol* 2007, **100**(3):361–402.
8. Day T, Bonduriansky R: **A unified approach to the evolutionary consequences of genetic and nongenetic inheritance.** *Am Nat* 2011, **178**(2):E18–E36.
9. Kuzawa CW, Thayer ZM: **Timescales of human adaptation: the role of epigenetic processes.** *Epigenomics* 2011, **3**(2):221–234.
10. Flatscher R, Frajman B, Schonswetter P, Paun O: **Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation?** *Genet Res Int* 2012, **2012**:698421.

11. Klimonos FD, Berg J, Collins S: How epigenetic mutations can affect genetic evolution: model and mechanism. *Bioessays* 2013, 35(6):571–578.
12. Darwin C: *The Descent of Man, and Selection in Relation to Sex*. London: T. Murray; 1871.
13. Anway MD, Cupp AS, Uzumcu M, Skinner MK: Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 2005, 308(5727):1466–1469.
14. Guerrero-Bosagna C, Settles M, Lucker BJ, Skinner MK: Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. *PLoS One* 2010, 5(9):e13100.
15. Skinner MK, Manikkam M, Guerrero-Bosagna C: Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab* 2010, 21(4):214–222.
16. Manikkam M, Guerrero-Bosagna C, Tracey R, Haque MM, Skinner MK: Transgenerational actions of environmental compounds on reproductive disease and epigenetic biomarkers of ancestral exposures. *PLoS One* 2012, 7(2):e31901.
17. Skinner MK, Manikkam M, Tracey R, Nilsson E, Haque MM, Guerrero-Bosagna C: Ancestral DDT exposures promote epigenetic transgenerational inheritance of obesity. *BMC Med* 2013, 11:228.
18. Kelce WR, Gray LE, Wilson EM: Antiestrogens as environmental endocrine disruptors. *Reprod Fertil Dev* 1998, 10(1):105–111.
19. Morgan HD, Santos F, Green K, Dean W, Reik W: Epigenetic reprogramming in mammals. *Hum Mol Genet* 2005, 14 Spec No 1:R47–R58.
20. Anway MD, Leathers C, Skinner MK: Endocrine disruptor vinclozolin induced epigenetic transgenerational adult-onset disease. *Endocrinology* 2006, 147(12):5515–5523.
21. Anway MD, Skinner MK: Transgenerational effects of the endocrine disruptor vinclozolin on the prostate transcriptome and adult onset disease. *Prostate* 2008, 68(5):517–529.
22. Nilsson EE, Anway MD, Stanfield J, Skinner MK: Transgenerational epigenetic effects of the endocrine disruptor vinclozolin on pregnancies and female adult onset disease. *Reproduction* 2008, 135(5):713–721.
23. Skinner MK, Anway MD, Savenkova MI, Gore AC, Crews D: Transgenerational epigenetic programming of the brain transcriptome and anxiety behavior. *PLoS One* 2008, 3(11):e3745.
24. Anway MD, Rekow SS, Skinner MK: Transgenerational epigenetic programming of the embryonic testis transcriptome. *Genomics* 2008, 91(1):30–40.
25. Skinner MK, Manikkam M, Haque MM, Zhang B, Savenkova M: Epigenetic transgenerational inheritance of somatic transcriptomes and epigenetic control regions. *Genome Biol* 2012, 13(10):R91.
26. Crews D, Gillette R, Scarpino SV, Manikkam M, Savenkova MI, Skinner MK: Epigenetic transgenerational inheritance of altered stress responses. *Proc Natl Acad Sci USA* 2012, 109(23):9143–9148.
27. Crews D, Gore AC, Hsu TS, Dangleben NL, Spinetta M, Schallert T, Anway MD, Skinner MK: Transgenerational epigenetic imprints on mate preference. *Proc Natl Acad Sci USA* 2007, 104(14):5942–5946.
28. Friend SH: The need for precompetitive integrative bionetwork disease model building. *Clin Pharmacol Ther* 2010, 87(5):536–539.
29. Schadt EE, Lamb J, Yang X, Zhu J, Edwards S, Guhathakurta D, Sieberts SK, Monks S, Reitman M, Zhang C, Lum PY, Leonardson A, Thieringer R, Metzger JM, Yang L, Castle J, Zhu H, Kash SF, Drake TA, Sachs A, Lusis AJ: An integrative genomics approach to infer causal associations between gene expression and disease. *Nat Genet* 2005, 37:710–717.
30. Yang X, Deignan JL, Qi H, Zhu J, Qian S, Zhong J, Torosyan G, Majid S, Falkard B, Kleinhanz RR, Karlsson J, Castellani LW, Mumick S, Wang K, Xie T, Coon M, Zhang C, Estrada-Smith D, Farber CR, Wang SS, van Nas A, Ghazalpour A, Zhang B, Macneil DJ, Lamb JR, Dipple KM, Reitman ML, Mehrabian M, Lum PY, Schadt EE, et al: Validation of candidate causal genes for obesity that affect shared metabolic pathways and networks. *Nat Genet* 2009, 41:415–423.
31. Nilsson EE, Savenkova MI, Schindler R, Zhang B, Schadt EE, Skinner MK: Gene bionetwork analysis of ovarian primordial follicle development. *PLoS One* 2010, 5(7):e11637.
32. Nilsson E, Zhang B, Skinner MK: Gene bionetworks that regulate ovarian primordial follicle assembly. *BMC Genomics* 2013, 14(1):496.
33. Zhu J, Zhang B, Smith EN, Drees B, Brem RB, Kruglyak L, Bumgarner RE, Schadt EE: Integrating large-scale functional genomic data to dissect the complexity of yeast regulatory networks. *Nat Genet* 2008, 40(7):854–861.
34. Millstein J, Zhang B, Zhu J, Schadt EE: Disentangling molecular relationships with a causal inference test. *BMC Genet* 2009, 10:23.
35. Pandey G, Zhang B, Chang AN, Myers CL, Zhu J, Kumar V, Schadt EE: An integrative multi-network and multi-classifier approach to predict genetic interactions. *PLoS Comput Biol* 2010, 6(9):e1000928.
36. DiBenedictis BT, Ingraham KL, Baum MJ, Cherry JA: Disruption of urinary odor preference and lordosis behavior in female mice given lesions of the medial amygdala. *Physiol Behav* 2012, 105(2):554–559.
37. Anway MD, Memon MA, Uzumcu M, Skinner MK: Transgenerational effect of the endocrine disruptor vinclozolin on male spermatogenesis. *J Androl* 2006, 27(6):868–879.
38. Ghazalpour A, Doss S, Zhang B, Wang S, Plaisier C, Castellanos R, Brozell A, Schadt EE, Drake TA, Lusis AJ, Horvath S: Integrating genetic and network analysis to characterize genes related to mouse weight. *PLoS Genet* 2006, 2:e130.
39. Sakata JT, Crews D: Developmental sculpting of social phenotype and plasticity. *Neurosci Biobehav Rev* 2004, 28(2):95–112.
40. Skinner MK: Role of epigenetics in developmental biology and transgenerational inheritance. *Birth Defects Res C Embryo Today* 2011, 93(1):51–55.
41. Skinner MK: Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics* 2011, 6(7):838–842.
42. Skinner M, Guerrero-Bosagna C, Haque MM, Nilsson E, Bhandari R, McCarrey J: Environmentally induced transgenerational epigenetic reprogramming of primordial germ cells and subsequent germline. *PLoS One* 2013, 8(7):e66318.
43. Salian S, Doshi T, Vanage G: Impairment in protein expression profile of testicular steroid receptor coregulators in male rat offspring perinatally exposed to Bisphenol A. *Life Sci* 2009, 85(1–2):11–18.
44. Bruner-Tran KL, Osteen KG: Developmental exposure to TCDD reduces fertility and negatively affects pregnancy outcomes across multiple generations. *Reprod Toxicol* 2011, 31(3):344–350.
45. Waterland RA, Travisano M, Tahiliani KG, Rached MT, Mirza S: Methyl donor supplementation prevents transgenerational amplification of obesity. *Int J Obes (Lond)* 2008, 32(9):1373–1379.
46. Painter RC, Osmond C, Gluckman P, Hanson M, Phillips DL, Roseboom TJ: Transgenerational effects of prenatal exposure to the Dutch famine on neonatal adiposity and health in later life. *BJOG* 2008, 115(10):1243–1249.
47. Pembrey ME, Bygren LO, Kaati G, Edvinsson S, Northstone K, Sjöström M, Golding J: Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet* 2006, 14(2):159–166.
48. Dias BG, Ressler KJ: Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* 2014, 17(1):89–96.
49. Buck LB: Olfactory receptors and odor coding in mammals. *Nutr Rev* 2004, 62(11 Pt 2):S184–S188. discussion S224–141.
50. Brennan PA, Zufall F: Pheromonal communication in vertebrates. *Nature* 2006, 444(7117):308–315.
51. Dulac C, Wagner S: Genetic analysis of brain circuits underlying pheromone signaling. *Annu Rev Genet* 2006, 40:449–467.
52. Green RE, Krause J, Briggs AW, Maricic T, Stenzel U, Kircher M, Patterson N, Li H, Zhai W, Fritz MH, Hansen NF, Durand EY, Malaspina AS, Jensen JD, Marques-Bonet T, Alkan C, Prüfer K, Meyer M, Burbano HA, Good JM, Schultz R, Aximu-Petri A, Butthof A, Hobert B, Hoffner B, Siegmund M, Weihmann A, Nusbaum C, Lander ES, Russ, et al: A draft sequence of the Neandertal genome. *Science* 2010, 328:710–722.
53. Kim SY, Kim YS: A gene sets approach for identifying prognostic gene signatures for outcome prediction. *BMC Genomics* 2008, 9:177.
54. Williams G: *Adaptation and Natural Selection: A Critique of Some Current Evolutionary Thought*. Princeton, NJ: Princeton Univ. Press; 1966.
55. Crews D: Diversity of hormone-behavior relations in reproductive behavior. In *Introduction to Behavioral Endocrinology*. 2nd edition. Edited by Becker JBM, McCarthy M, Crews D: MIT Press/Bradford Books; 1992:143–186.
56. Crews D: The evolutionary antecedents to love. *Psychoneuroendocrinology* 1998, 23(8):751–764.
57. Gowaty PA, Anderson WW, Bluhm CK, Drickamer LC, Kim YK, Moore AJ: The hypothesis of reproductive compensation and its assumptions about mate preferences and offspring viability. *Proc Natl Acad Sci USA* 2007, 104(38):15023–15027.
58. Matile B, Wilson AB: Body size preferences in the pot-bellied seahorse *Hippocampus abdominalis*: choosy males and indiscriminate females. *Behav Ecol Sociobiol* 2009, 63(10):1403–1410.

59. Carson HL: The contribution of sexual behavior to Darwinian fitness. *Behav Genet* 1987, 17(6):597–611.
60. Carson HL: Mate choice theory and the mode of selection in sexual populations. *Proc Natl Acad Sci U S A* 2003, 100(11):6584–6587.
61. Gowaty PA, Steinichen R, Anderson WW: Mutual interest between the sexes and reproductive success in *Drosophila pseudoobscura*. *Evolution* 2002, 56(12):2537–2540.
62. Stuenden CE, Bluhm CK, Cheng KM, Rajamahendran R: Factors affecting reproductive performance in captive Mallard ducks. *Theriogenology* 1999, 52(3):435–446.
63. Drickamer LC, Gowaty PA, Holmes CM: Free female mate choice in house mice affects reproductive success and offspring viability and performance. *Anim Behav* 2000, 59(2):371–378.
64. Drickamer LCP, Wagner DM: Free mutual mate preferences in house mice affect reproductive success and offspring performance. *Animal Behav* 2003, 65:105–114.
65. Kirkpatrick M, Ryan MJ: The evolution of mating preferences and the paradox of the lek. *Nature* 1991, 350:33–38.
66. Ryan M: Sexual selection, sensory systems, and sensory exploitation. *Oxford Survey Evol Biol* 1990, 7:157–196.
67. Beach F: Animal models for human sexuality. In *Sex, Hormones and Behaviour*. Edited by Potter RWJ. Amsterdam: Ciba Foundation Symposium 62, Excerpta Medica; 1979:113–143.
68. Crews D: Evolution of neuroendocrine mechanisms that regulate sexual behavior. *Trends Endocrinol Metab* 2005, 16(8):354–361.
69. Paxinos G, Watson C: *The Rat Brain in Stereotaxic Coordinates*. New York: Academic; 2007.
70. Bosotti R, Locatelli G, Healy S, Scacheri E, Sartori L, Mercurio C, Calogero R, Isacchi A: Cross platform microarray analysis for robust identification of differentially expressed genes. *BMC Bioinforma* 2007, 8(Suppl 1):S5.
71. Zhang B, Horvath S: A general framework for weighted gene co-expression network analysis. *Stat Appl Genet Mol Biol* 2005, 4: Article17.
72. Zhu J, Wiener MC, Zhang C, Fridman A, Minch E, Lum PY, Sachs JR, Schadt EE: Increasing the power to detect causal associations by combining genotypic and expression data in segregating populations. *PLoS Comput Biol* 2007, 3(4):e69.
73. Ravasz E, Somera AL, Mongru DA, Oltvai ZN, Barabasi AL: Hierarchical organization of modularity in metabolic networks. *Science* 2002, 297(5586):1551–1555.
74. Langfelder P, Zhang B, Horvath S: Defining clusters from a hierarchical cluster tree: the dynamic tree cut package for r. *Bioinformatics* 2008, 24(5):719–720.
75. Lum PY, Chen Y, Zhu J, Lamb J, Melmed S, Wang S, Drake TA, Lusis AJ, Schadt EE: Elucidating the murine brain transcriptional network in a segregating mouse population to identify core functional modules for obesity and diabetes. *J Neurochem* 2006, 97(Suppl 1):50–62.
76. Chen Y, Zhu J, Lum PY, Yang X, Pinto S, MacNeil DJ, Zhang C, Lamb J, Edwards S, Sieberts SK, Leonardson A, Castellini LW, Wang S, Champy MF, Zhang B, Emilsson V, Doss S, Ghazalpour A, Horvath S, Drake TA, Lusis AJ, Schadt EE: Variations in DNA elucidate molecular networks that cause disease. *Nature* 2008, 452:429–435.
77. Gargalovic PS, Imura M, Zhang B, Gharavi NM, Clark MJ, Pagnon J, Yang WP, He A, Truong A, Patel S, Nelson SF, Horvath S, Berliner JA, Kirchgessner TG, Lusis AJ: Identification of inflammatory gene modules based on variations of human endothelial cell responses to oxidized lipids. *Proc Natl Acad Sci U S A* 2006, 103:12741–12746.
78. Horvath S, Zhang B, Carlson M, Lu KV, Zhu S, Feliciano RM, Laurance MF, Zhao W, Qi S, Chen Z, Lee Y, Scheck AC, Liao LM, Wu H, Geschwind DH, Febbo PG, Kornblum HI, Cloughesy TF, Nelson SF, Mischel PS: Analysis of oncogenic signaling networks in glioblastoma identifies ASPM as a molecular target. *Proc Natl Acad Sci U S A* 2006, 103:17402–17407.
79. Emilsson V, Thorleifsson G, Zhang B, Leonardson AS, Zink F, Zhu J, Carlson S, Helgason A, Walters GB, Gunnarsdottir S, Mouy M, Steinhorsdottir V, Eiriksdottir GH, Bjornsdottir G, Reynisdottir I, Gudbjartsson D, Helgadottir A, Jonasdottir A, Styrkarsdottir U, Gretarsdottir S, Magnusson KP, Stefansson H, Fossdal R, Kristjansson K, Gislason HG, Stefansson T, Leifsson BG, Thorsteinsdottir U, Lamb JR, Gulcher, et al: Genetics of gene expression and its effect on disease. *Nature* 2008, 452:423–428.
80. Schadt EE, Molony C, Chudin E, Hao K, Yang X, Lum PY, Kasarskis A, Zhang B, Wang S, Suver C, Zhu J, Millstein J, Sieberts S, Lamb J, GuhaThakurta D, Derry J, Storey JD, Avila-Campillo I, Kruger MJ, Johnson JM, Rohl CA, van Nas A, Mehrabian M, Drake TA, Lusis AJ, Smith RC, Guengerich FP, Strom SC, Schuetz E, Rushmore TH, et al: Mapping the genetic architecture of gene expression in human liver. *PLoS Biol* 2008, 6:e107.

81. Zhu J, Zhang B, Schadt EE: A systems biology approach to drug discovery. *Adv Genet* 2008, 60:603–635.
82. Draghici S, Khatri P, Tarca AL, Amin K, Done A, Voichita C, Georgescu C, Romero R: A systems biology approach for pathway level analysis. *Genome Res* 2007, 17(10):1537–1545.

doi:10.1186/1471-2164-15-377

Cite this article as: Skinner et al.: Gene bionetworks involved in the epigenetic transgenerational inheritance of altered mate preference: environmental epigenetics and evolutionary biology. *BMC Genomics* 2014 15:377.

Submit your next manuscript to BioMed Central and take full advantage of:

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)



# Environmental Epigenetics and a Unified Theory of the Molecular Aspects of Evolution: A Neo-Lamarckian Concept that Facilitates Neo-Darwinian Evolution

Michael K. Skinner\*

Center for Reproductive Biology, School of Biological Sciences, Washington State University

\*Corresponding author: E-mail: skinner@wsu.edu.

Accepted: April 17, 2015

## Abstract

Environment has a critical role in the natural selection process for Darwinian evolution. The primary molecular component currently considered for neo-Darwinian evolution involves genetic alterations and random mutations that generate the phenotypic variation required for natural selection to act. The vast majority of environmental factors cannot directly alter DNA sequence. Epigenetic mechanisms directly regulate genetic processes and can be dramatically altered by environmental factors. Therefore, environmental epigenetics provides a molecular mechanism to directly alter phenotypic variation generationally. Lamarck proposed in 1802 the concept that environment can directly alter phenotype in a heritable manner. Environmental epigenetics and epigenetic transgenerational inheritance provide molecular mechanisms for this process. Therefore, environment can on a molecular level influence the phenotypic variation directly. The ability of environmental epigenetics to alter phenotypic and genotypic variation directly can significantly impact natural selection. Neo-Lamarckian concept can facilitate neo-Darwinian evolution. A unified theory of evolution is presented to describe the integration of environmental epigenetic and genetic aspects of evolution.

**Key words:** epigenetics, Lamarck, Darwin, natural selection, environment, review.

## Introduction

Charles Darwin's concept of evolution by natural selection is the unifying theme for much of modern biology (Darwin 1859). Remarkably, Darwin had no understanding of the molecular mechanisms involved in this process. Integration of Darwin's thinking with advances in genetic and molecular sciences over the past century facilitated the development of a well supported neo-Darwinian theory of evolution (Olson-Manning et al. 2012). The current primary concept for the molecular basis of evolution involves genetics and mutations, such that random DNA sequence and chromosomal alterations create a genetic variation that directly impacts phenotype and phenotypic variation. The majority of models in evolutionary biology involves DNA sequence mutations as the primary molecular mechanism underlying heritable phenotypic variation (Laland et al. 2014). A conundrum in evolutionary theory is that the frequency of potentially advantageous genetic mutations is extremely low (Jablonka and Raz 2009; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Nei and Nozawa 2011; Laland et al. 2014). Although recent studies with organisms such as microbes

demonstrate genotypic variation are sufficient (Levy and Siegal 2008; Avelar et al. 2013; Ho and Zhang 2014) and additional mechanisms such as random genetic drift, genetic assimilation, directed mutations and epistasis also play important roles, genetic theory alone has difficulty explaining some aspects of evolution (Laland et al. 2014). For example, phenotypic mutation rates and genotypic mutation rates are dramatically different and genetics has been the primary molecular mechanism considered (Burger et al. 2006), but the inclusion of an additional mechanism such as epigenetics can help explain this discordance. Understanding the origins of genotypic variation and rapid evolutionary phenomenon under environmental pressure is difficult to explain with only classic genetics considered. Opposing groups of evolutionary biologists are now debating the need to "rethink" the theory (Laland et al. 2014). Genetics is the primary molecular mechanism considered in classic neo-Darwinian evolution theory (Olson-Manning et al. 2012) (table 1 and fig. 1).

In addition to evolution considerations, a large number of biological phenomena have been observed that cannot be

**Table 1**

Evolution Theory Components

## Neo-Lamarckian concept

Environment directly alters phenotype generationally

## Darwinian evolution theory

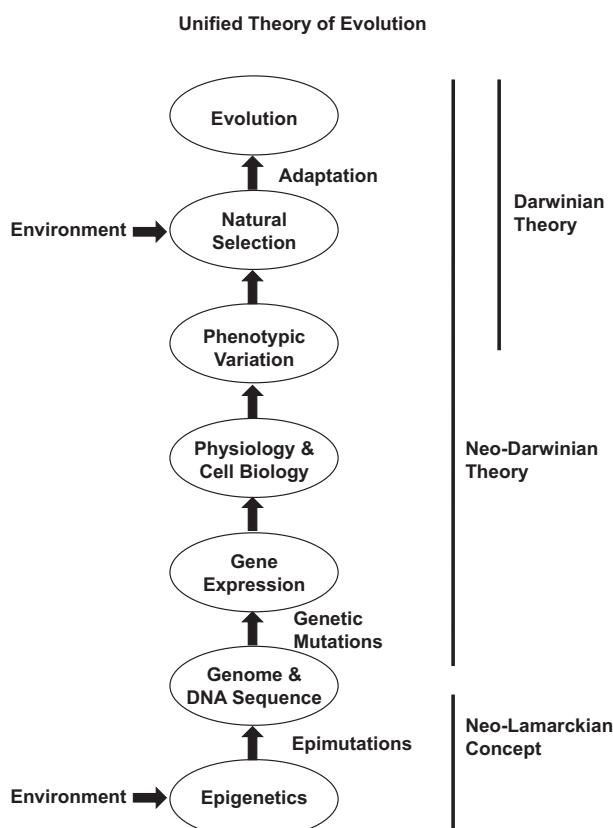
Natural selection acts on phenotypic variation

## Neo-Darwinian evolution theory

Genetic mutations promote phenotypic variation on which natural selection acts

## Unified evolution theory

Environmental epigenetic alterations promote genetic mutations to alter genotypic variation Environmental epigenetics and genetic mutations both promote phenotypic variation on which natural selection acts



**Fig. 1.**—Schematic of the unified theory of evolution. No dominance is suggested by the appearance of specific circles (e.g., epimutations vs. genetics) such that all are equally important components.

easily explained by genetics alone. These include the fact that identical twins with similar genetics generally have discordant disease (Zwijnenburg et al. 2010; Kratz et al. 2014; Tan et al. 2015), or the fact that generally only a small percentage of a disease population has been found to have a correlated genetic mutation, or the fact that many diseases have increased

in frequency an order of magnitude in only a couple decades, or the fact that hundreds of environmental contaminants not able to alter DNA sequence have been shown to alter disease or phenotype later in life (Skinner 2014a). Many biological observations do not follow normal Mendelian genetic rules and are difficult to explain with classic genetic processes or mechanisms (McClintock 1984). An example in evolution is that the rates of molecular and morphological evolution are largely decoupled and these patterns of phenotypic divergence are regulatory and not classic genetic mutations (Janecka et al. 2012). Epigenetic resolution of the “curse of complexity” in adaptive evolution of complex traits has been suggested (Badyaev 2014).

Recently documented molecular mechanisms that can dramatically influence genome activity and contribute to phenotypic variation involve epigenetics (Skinner et al. 2010). Many of the above phenomenon when epigenetics is considered as an additional molecular mechanism can be more easily understood, such as the discordance of identical twins (Zwijnenburg et al. 2010; Kratz et al. 2014; Tan et al. 2015). Waddington (1953) coined the term epigenetics and the classic epigenetic definitions of Waddington (1953) and others (Skinner 2011) are descriptive, without an understanding of the molecular elements (Skinner 2011). Considering our current molecular understanding, epigenetics is defined as “molecular processes around DNA that regulate genome activity independent of DNA sequence and are mitotically stable” (Skinner et al. 2010). These epigenetic mechanisms include DNA methylation, histone modifications, chromatin structure, and selected noncoding RNA (ncRNA) (Skinner 2014a). Epigenetic processes such as DNA methylation can become programmed (e.g., imprinted) and be inherited over generations (Skinner 2014a). Environmental factors have been shown to promote the epigenetic transgenerational inheritance of phenotypic variation. Several examples of environmentally induced epigenetic transgenerational inheritance of phenotypic change have been shown to be inherited for hundreds of generations (Cubas et al. 1999). Therefore, like genetic changes, epigenetic changes can have an important role in short-term microevolution (Day and Bonduriansky 2011) and contribute to macroevolutionary (i.e., at or above the level of species) processes, such as speciation and adaptive radiation (Rebolledo et al. 2010; Flatscher et al. 2012). A number of insightful reviews have proposed a role for epigenetics in evolution, primarily as a responsive molecular mechanism in natural selection (Jablonska et al. 1998; Pigliucci 2007; Laland et al. 2014).

## Environment and Evolution

A variety of environmental factors can influence evolution and general biology. These range from ecological parameters such as temperature and light to nutritional parameters such as caloric restriction or high fat diets. A host of environmental chemicals from phytochemicals to toxicants can also influence

phenotype and health (Skinner 2014a). Environment has a critical role in natural selection and Darwinian evolution (Darwin 1859). Natural selection is a process in which environmental factors influence the survival or reproductive success of individuals bearing different phenotypes. The current paradigm in evolutionary biology holds that changes in DNA sequence underlie the variation that can evolve in response to natural selection (Laland et al. 2014) (table 1). Although James Baldwin in 1896 suggested environment through sociobiology type mechanisms (i.e., behavior) could alter phenotypic variation, these are thought to be due to genetic changes and considered a neo-Darwinian process (Baldwin 1896; Paenke et al. 2007). Therefore, in neo-Darwinian evolution the primary link between the environment and evolution is to mediate the natural selection process (Olson-Manning et al. 2012; Laland et al. 2014).

In contrast, Lamarck proposed one of the early evolutionary theories in 1802 in that environment promotes the phenotypic alterations associated with evolution (Lamarck 1802; Calabi 2001). This is distinct to the role of environment providing selective pressure in natural selection, such that environment directly alters the phenotype to influence evolution. This theory was seen as conflicting with Darwin's natural selection evolutionary theory and so was discounted and today is not seriously considered in modern evolutionary theory or neo-Darwinian evolution (Day and Bonduriansky 2011). However, if there was a molecular mechanism that generationaly could facilitate the ability of the environment to alter genotypic and phenotypic variation, such a neo-Lamarckian concept may facilitate evolution (table 1 and fig. 1).

Interestingly, Darwin (1868) himself was a strong proponent of the inheritance of acquired characteristics. The blending of inheritance and evolution by natural selection appeared to be a fundamentally flawed concept that would require an untenably high mutation rate in order to maintain the trait variation required for selection (Jenkins 1867). To address this, Darwin (1868) proposed pangenesis, a complex theory of environmentally responsive somatic cell transmittance to offspring. Therefore, Darwin conceptually supported Lamarck's theory of the inheritance of acquired characteristics, but until the last 30 years the potential molecular mechanism was unclear.

## Environmental Epigenetics

Epigenetics provides molecular mechanisms for the environment to directly alter phenotypic variation and its subsequent inheritance (Crews et al. 2007; Skinner, Gurerrero-Bosagna, Haque, et al. 2014). A variety of epigenetic mechanisms have been identified including DNA methylation, histone modifications, chromatin structure, and selected ncRNA. All these mechanisms have the ability to program and alter gene expression and have been shown to have a critical role in normal development and biological processes (Skinner et al. 2010;

Skinner 2014a). For example, the ability to generate an embryonic stem cell requires the erasure of DNA methylation such that the cell becomes pluripotent (Seisenberger et al. 2013). Although the vast majority of environmental factors cannot alter DNA sequence, epigenetic processes can be dramatically altered in response to environmental factors from nutrition to temperature (Skinner 2014a). All organisms that have been investigated contain highly conserved epigenetic processes (e.g., DNA methylation) that can be environmentally modified (Skinner 2014a). Epigenetics provides an additional molecular mechanism, integrated with genetics, to regulate biology.

The ability of environment to directly alter the development and function of cells and tissues is critical for the health and phenotype of the individual. This direct environmental epigenetic effect on the individual would likely have a limited impact on evolution, unless the epigenetic changes could be transmitted between generations. A large number of environmental factors from nutrition to toxicants have been shown to induce the epigenetic transgenerational inheritance of disease and phenotypic variation (Skinner 2014a). Epigenetic transgenerational inheritance is defined as the germline transmission of epigenetic information between generations in the absence of direct exposure (Skinner et al. 2010). Environmental exposures during a critical period of germline development, fetal gonadal sex determination or gametogenesis, have been shown to permanently program epigenetic marks such as DNA methylation (Skinner 2014a). Nutrition (Pembrey et al. 2006; Burdge et al. 2011), temperature (Song et al. 2013), stress (Skinner 2014b), and toxicants (Anway et al. 2005; Skinner 2014a) have all been shown to promote the epigenetic transgenerational inheritance of phenotypic variation (Skinner 2014a). The phenomenon has been observed in plants, insects, fish, rodents, pigs, and humans (Skinner 2014a). In mammals the altered transgenerational phenotypes have been observed for generations (Skinner 2014a), with environmentally induced epigenetic transgenerational inheritance of phenotypic variation in plants being transmitted for hundreds of generations (Cubas et al. 1999). Therefore, environment can promote the epigenetic transgenerational inheritance of phenotypic variation. The ability of environment to alter phenotype and alter phenotypic variation, independent of genetics, through this epigenetic mechanism is proposed to be important for evolution (Anway et al. 2005; Jablonka and Raz 2009; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Skinner 2014a).

Darwin proposed that one of the critical determinants of evolution was sexual selection (Darwin 1859). A previous study investigated the ability of an environmental factor (toxicant) to promote the epigenetic transgenerational inheritance of an alteration in mate preference associated with sexual selection (Crews et al. 2007). An F0 generation gestating female rat was exposed to the agricultural fungicide vinclozolin transiently and then the F3 generation animals

(great-grand-offspring) were obtained to assess alterations in mate preference behavior (Anway et al. 2005). A dramatic alteration in mate preference was observed (Crews et al. 2007) along with epigenetic alterations (termed epimutations) in the germline (sperm) (Guerrero-Bosagna et al. 2010). Transgenerational transcriptome changes in the brain regions correlated with the alterations in mate preference behavior (Skinner et al. 2008). Therefore, an environmental factor that altered sexual selection was found to promote a permanent alteration in the sperm epigenome in an imprinted-like manner that was inherited for multiple generations (Crews et al. 2007; Skinner et al. 2010). These studies suggest that environmental epigenetics may play an important role in evolutionary change. The role of epigenetics in mate choice and evolution has been further discussed (Zeh JA and Zeh DW 2008; Bonduriansky and Day 2013). Indeed, several recent reviews have suggested a role for epigenetics in microevolution and macroevolution (Jablonska and Raz 2009; Rebolledo et al. 2010; Skinner et al. 2010; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Flatscher et al. 2012; Klironomos et al. 2013; Badyaev 2014; Jaeger and Monk 2014; Skinner 2014a).

## Unified Theory

Environmental epigenetics and epigenetic transgenerational inheritance provide a molecular mechanism for the neo-Lamarckian concept that environmental factors directly alter phenotype (table 1). The ability of environmental epigenetics to alter phenotypic variation provides an initial element for evolution where environment can directly establish the variation and phenotype in a population (fig. 1). Although aspects of the original Lamarckian evolution theory were not accurate (Lamarck 1802), such as having "directed" phenotypes within a generation (Koonin and Wolf 2009; Koonin 2014), the concept that environment can directly impact phenotype is supported by environmental and transgenerational epigenetic studies (Crews et al. 2007; Koonin and Wolf 2009; Koonin 2014; Skinner, Guerrero-Bosagna, Haque, et al. 2014). Therefore, the first aspect of the unified theory involves the ability of environment to impact epigenetic programming generationally to alter phenotypic variation (fig. 1).

The well-established aspect of Darwinian evolution is the ability of environment through natural selection to act on phenotypic variation within an evolutionary event (Darwin 1859; Olson-Manning et al. 2012). The classic neo-Darwinian view is that genetic mutations and genetic variation are the primary molecular mechanism involved in generating the phenotypic variation (Nei and Nozawa 2011; Olson-Manning et al. 2012) (table 1). Although epigenetics can also have a critical role in the establishment and maintenance of phenotypic variation, the genetic mutations and genotype of the phenotype will be critical. This neo-Darwinian natural

selection event for evolution is the other component of the unified theory (fig. 1).

A combination of environmental epigenetic impacts on phenotypic variation and the ability of environment to mediate natural selection will both be important for evolution. Therefore, this neo-Lamarckian concept facilitates neo-Darwinian evolution (fig. 1). This unified theory provides an expanded understanding of the molecular aspects of evolution and solutions for issues such as the mechanisms for rapid evolutionary phenomenon. The mechanisms that environment can impact evolution are also expanded. An integration of epigenetics and genetics will be essential to consider in our future understanding of the molecular aspects of evolution (Jablonska and Raz 2009; Day and Bonduriansky 2011; Laland et al. 2014; Skinner 2014a).

An additional important consideration involves the ability of epigenetic processes to promote genetic mutations (table 1). In cancer biology, altered epigenetics has been shown to promote genome instability and formation of genetic mutations (Feinberg 2004). Nearly all genetic mutations can be directly influenced by epigenetic processes. The most frequent point mutation (single nucleotide polymorphism) is a C to T conversion that is facilitated by CpG DNA methylation (Jones et al. 1992). Repeat elements in the genome when expanded create copy number variations (CNV) that are controlled by hypermethylation of DNA (Liu et al. 2012). Transposable elements are also silenced by hypermethylation of DNA (Yagi et al. 2012). Translocation events and inversions are also influenced by histone modifications, DNA methylation, and ncRNA (Solary et al. 2014). Therefore, epigenetics can directly influence genetic mutations and the origin of genotypic variation is influenced by environmental epigenetic alterations (table 1). In contrast, genetic mutations have been shown to influence epigenetics (Furey and Sethupathy 2013). Recently, we have found that environmentally induced epigenetic transgenerational inheritance of disease and phenotypic variation can promote genetic mutations (i.e., CNV) in later generations (Skinner MK, Guerrero-Bosagna C, Haque MM, unpublished data). Therefore, environmental epigenetics may not only promote increased phenotypic variation, but epigenetics can also drive genetic change and increase genotypic variation. This also needs to be considered in the unified evolution theory (fig. 1).

## Discussion

Environmental epigenetics and epigenetic transgenerational inheritance alter phenotypic variation which can be acted on by natural selection. Therefore, environmental epigenetics can directly influence phenotype and this neo-Lamarckian concept can facilitate natural selection and neo-Darwinian evolution. These different aspects of evolution should not be seen as conflicting, but instead can form a unified theory for evolution (fig. 1). This expanded understanding of the molecular aspects of evolution provides novel insights into the mechanism for

rapid evolutionary events. An expanded understanding of how environment impacts evolution is also provided. The unified theory provides novel considerations that environment can both act to directly influence phenotypic variation and directly facilitate natural selection (fig. 1). Previous evolutionary models have primarily considered genetics and mutations as the primary molecular driver for evolution (Nei and Nozawa 2011; Olson-Manning et al. 2012; Laland et al. 2014). More recently, a number of models have started to consider epigenetics in these evolution models as well (Rebollo et al. 2010; Skinner et al. 2010; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Flatscher et al. 2012; Klironomos et al. 2013; Badyaev 2014; Jablonka and Lamb 2014; Jaeger and Monk 2014). For example, consideration of epigenetics as an additional molecular mechanism has assisted in the understanding of genetic drift (Gordon et al. 2012), genetic assimilation (Zuckerkandl and Cavalli 2007), and directed mutation (Jablonka and Lamb 2007; Kryazhimskiy et al. 2014). The consideration of epigenetics can also be used to better understand neutral evolution (Kimura 1989) through mechanisms, such as robustness (Ohta 2011). The unified theory suggests additional variables that should be considered are the multiple roles of environment and the integration of epigenetics into future evolution models.

Epigenetic transgenerational inheritance of phenotypic variation will have an important role in microevolutionary and macroevolutionary changes, including speciation. A recent study was designed to investigate the epigenetic changes associated with phylogenetic distance in Darwin's finches (Skinner, Gurerrero-Bosagna, Haque, et al. 2014), a well-known example of adaptive radiation (Darwin 1859; Lack 1947; Burns et al. 2002; Grant and Grant 2008; Huber et al. 2010; Donohue 2011). Erythrocyte DNA was obtained from five species of sympatric Darwin's finches that vary in phylogenetic relatedness. Genome-wide alterations in genetic mutations, using CNV, were compared with epigenetic alterations associated with differential DNA methylation regions (epimutations) (Skinner, Gurerrero-Bosagna, Haque, et al. 2014). A greater number of epimutations than genetic mutations were observed among the different species, with the number of epimutations increasing with phylogenetic distance. The number, chromosomal locations, regional clustering, and overlap of epimutations suggest that epigenetic change has likely had a role in the speciation and evolution of Darwin's finches (Skinner, Gurerrero-Bosagna, Haque, et al. 2014). A number of additional observations also support a role of epigenetics and speciation. Using *Drosophila* and maternally inherited ncRNA silencing of transposons a role for epigenetics and speciation was discussed (Brennecke et al. 2008). The role of epigenetics and a punctuated equilibrium in the mobilization of transposable elements was also suggested (Zeh et al. 2009). An interesting study comparing Neanderthal and human DNA methylation maps also supports

a role for epigenetics in speciation (Gokhman et al. 2014) and evolution.

Although the causal role of epimutations was not established in the Darwin's finch adaptive radiation (Skinner, Gurerrero-Bosagna, Haque, et al. 2014) or other models (Brennecke et al. 2008; Zeh et al. 2009; Gokhman et al. 2014), the causal role of genome-wide genetic mutations has also not been established (Laland et al. 2011). Future studies need to focus on the causal relationship of epigenetic alterations in relation to phenotypic variation that is acted on by natural selection. Genetics and genetic mutations are critical for evolution, but they are not the only molecular factors to consider. Although the major paradigm in the biological sciences is genetic determinism, this paradigm is limited in its ability to explain biological phenomenon ranging from the molecular basis of disease etiology (Skinner 2014a) to certain aspects of evolution by natural selection (Skinner et al. 2010; Day and Bonduriansky 2011; Longo et al. 2012). As Thomas Kuhn suggested during a scientific revolution when the current paradigm reveals anomalies then new science needs to be considered (Kuhn 1962). This type of challenge to current paradigms is also supported by other scientific philosophy, such as Popper (Rieppel 2008) and MacIntyre (MacIntyre 1977). A paradigm shift is required to explain how genetics and epigenetics integrate to regulate genome activity and evolution, and these advances will need to be incorporated into future evolutionary biology modeling (Rebollo et al. 2010; Skinner et al. 2010; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Flatscher et al. 2012; Klironomos et al. 2013; Badyaev 2014; Jablonka and Lamb 2014; Jaeger and Monk 2014; Skinner 2014a) and theory.

## Summary

The integration of environmental epigenetics into the molecular aspects of evolution theory suggests a neo-Lamarckian concept that facilitates neo-Darwinian evolution. Several of the novel factors to be considered are summarized below. In regards to the neo-Lamarckian concept:

1. Environmental epigenetics provides a molecular mechanism for Lamarck's proposal that environment can directly alter phenotype in a heritable manner.
2. Environmental exposures at critical developmental windows promote the epigenetic transgenerational inheritance of germline (e.g., sperm) epimutations that alter phenotypic variation.
3. Direct environmental exposures of developing somatic tissue can alter somatic epigenomes and phenotype in the individual exposed, but this will not be heritable and the phenotypes will often be distinct to transgenerational phenotypes.
4. In regards to novel aspects of neo-Darwinian evolution:
5. Transgenerational germline epimutations alter genome stability to promote genetic mutations and genotypic variation in subsequent generations.

6. Phenotypic variation is derived from a combination of integrated genetic and epigenetic processes on which natural selection acts.
7. Environment has a critical role in natural selection, as well as in the induction of heritable adaptive phenotypic variation.

As shown in figure 1, these concepts and components contribute to a unified theory that integrates environmental epigenetics into the molecular aspects of evolution. It is important to note that there is not a dominance of genetics or epigenetics, but the two molecular processes integrate to regulate biology.

Previously, an environmental exposure was found to promote the epigenetic transgenerational inheritance of phenotypic traits such as mate preference, which can play an important role in evolution (Crews et al. 2007; Skinner 2014a). Several reviews have subsequently suggested a role for epigenetics in evolution (Jablonka and Raz 2009; Rebollo et al. 2010; Skinner et al. 2010; Day and Bonduriansky 2011; Kuzawa and Thayer 2011; Flatscher et al. 2012) and experimental models have shown the importance of epigenetic associated genes (Mihola et al. 2009) and molecular elements (Long et al. 2013; Skinner, Gurerrero-Bosagna, Haque, et al. 2014) in evolution. The current report extends these studies to present a unified theory that combines both neo-Lamarckian and neo-Darwinian aspects and expands our understanding of how environment impacts evolution. The integration of epigenetics and genetics will be critical for all areas of biology including evolution.

## Acknowledgments

The authors acknowledge the advice and critical reviews of Dr Richard Gomulkiewicz and Eric Nilsson (Washington State University), and Dr Carlos Guerrero-Bosagna (Linköping University, Sweden). The helpful comments of the reviewers of this article are also very much appreciated. They thank Ms Heather Johnson for assistance in preparation of the manuscript. The research was supported by a John Templeton Foundation grant to M.K.S. The author declares no competing financial interests.

## Literature Cited

- Anway MD, Cupp AS, Uzumcu M, Skinner MK. 2005. Epigenetic transgenerational actions of endocrine disruptors and male fertility. *Science* 308:1466–1469.
- Avelar AT, Perfeito L, Gordo I, Ferreira MG. 2013. Genome architecture is a selectable trait that can be maintained by antagonistic pleiotropy. *Nat Commun.* 4:2235.
- Badyaev AV. 2014. Epigenetic resolution of the “curse of complexity” in adaptive evolution of complex traits. *J Physiol.* 592:2251–2260.
- Baldwin J. 1896. A new factor in evolution. *Am Nat.* 30:441–451.
- Bonduriansky R, Day T. 2013. Nongenetic inheritance and the evolution of costly female preference. *J Evol Biol.* 26:76–87.
- Brennecke J, et al. 2008. An epigenetic role for maternally inherited piRNAs in transposon silencing. *Science* 322:1387–1392.
- Burdge GC, et al. 2011. Progressive, transgenerational changes in offspring phenotype and epigenotype following nutritional transition. *PLoS One* 6:e28282.
- Burger R, Willensdorfer M, Nowak MA. 2006. Why are phenotypic mutation rates much higher than genotypic mutation rates? *Genetics* 172:197–206.
- Burns KJ, Hackett SJ, Klein NK. 2002. Phylogenetic relationships and morphological diversity in Darwin’s finches and their relatives. *Evolution* 56:1240–1252.
- Calabi L. 2001. On Darwin’s “metaphysical notebooks.” I: Teleology and the project of a theory. *Riv Biol.* 94:123–159.
- Crews D, et al. 2007. Transgenerational epigenetic imprints on mate preference. *Proc Natl Acad Sci U S A.* 104:5942–5946.
- Cubas P, Vincent C, Coen E. 1999. An epigenetic mutation responsible for natural variation in floral symmetry. *Nature* 401:157–161.
- Darwin C. 1859. *On the origin of species*. London: John Murray.
- Darwin C. 1868. *The variation of animals and plants under domestication*. London: John Murray.
- Day T, Bonduriansky R. 2011. A unified approach to the evolutionary consequences of genetic and nongenetic inheritance. *Am Nat.* 178: E18–E36.
- Donohue K. 2011. Darwin’s finches: readings in the evolution of a scientific paradigm. Chicago (IL): University of Chicago Press.
- Feinberg AP. 2004. The epigenetics of cancer etiology. *Semin Cancer Biol.* 14:427–432.
- Flatscher R, Frajman B, Schonswetter P, Paun O. 2012. Environmental heterogeneity and phenotypic divergence: can heritable epigenetic variation aid speciation? *Genet Res Int.* 2012: 698421.
- Furey TS, Sethupathy P. 2013. Genetics. Genetics driving epigenetics. *Science* 342:705–706.
- Gokhman D, et al. 2014. Reconstructing the DNA methylation maps of the Neandertal and the Denisovan. *Science* 344:523–527.
- Gordon L, et al. 2012. Neonatal DNA methylation profile in human twins is specified by a complex interplay between intrauterine environmental and genetic factors, subject to tissue-specific influence. *Genome Res.* 22:1395–1406.
- Grant P, Grant R. 2008. How and why species multiply: the radiation of Darwin’s finches. Princeton (NJ): Princeton University Press.
- Guerrero-Bosagna C, Settles M, Lucker B, Skinner M. 2010. Epigenetic transgenerational actions of vinclozolin on promoter regions of the sperm epigenome. *PLoS One* 5:e13100.
- Ho WC, Zhang J. 2014. The genotype-phenotype map of yeast complex traits: basic parameters and the role of natural selection. *Mol Biol Evol.* 31:1568–1580.
- Huber SK, et al. 2010. Ecoimmunity in Darwin’s finches: invasive parasites trigger acquired immunity in the medium ground finch (*Geospiza fortis*). *PLoS One* 5:e8605.
- Jablonka E, Lamb MJ. 2007. Precis of evolution in four dimensions. *Behav Brain Sci.* 30:353–365; discussion: 365–389.
- Jablonka E, Lamb MJ. 2014. Evolution in four dimensions, revised edition. Cambridge: MIT Press.
- Jablonka E, Lamb MJ, Avital E. 1998. “Lamarckian” mechanisms in darwinian evolution. *Trends Ecol Evol.* 13:206–210.
- Jablonka E, Raz G. 2009. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. *Q Rev Biol.* 84:131–176.
- Jaeger J, Monk N. 2014. Bioattractors: dynamical systems theory and the evolution of regulatory processes. *J Physiol.* 592:2267–2281.
- Janecka J, Chowdhary B, Murphy W. 2012. Exploring the correlations between sequence evolution rate and phenotypic divergence across the Mammalian tree provides insights into adaptive evolution. *J Biosci.* 37:897–909.
- Jenkins F. 1867. The origins of species. *North Br Rev.* 46:277–318.

- Jones PA, Rideout WM 3rd, Shen JC, Spruck CH, Tsai YC. 1992. Methylation, mutation and cancer. *Bioessays* 14:33–36.
- Kimura M. 1989. The neutral theory of molecular evolution and the world view of the neutralists. *Genome* 31:24–31.
- Klironomos FD, Berg J, Collins S. 2013. How epigenetic mutations can affect genetic evolution: model and mechanism. *Bioessays* 35: 571–578.
- Koonin EV. 2014. Calorie restriction a Lamarck. *Cell* 158:237–238.
- Koonin EV, Wolf YI. 2009. Is evolution Darwinian or/and Lamarckian? *Biol Direct.* 4:42.
- Kratz CP, Edelman DC, Wang Y, Meltzer PS, Greene MH. 2014. Genetic and epigenetic analysis of monozygotic twins discordant for testicular cancer. *Int J Mol Epidemiol Genet.* 5:135–139.
- Kryazhimskiy S, Rice DP, Jerison ER, Desai MM. 2014. Microbial evolution. Global epistasis makes adaptation predictable despite sequence-level stochasticity. *Science* 344:1519–1522.
- Kuhn TS. 1962. The structure of scientific revolutions. Chicago (IL): University of Chicago Press.
- Kuzawa CW, Thayer ZM. 2011. Timescales of human adaptation: the role of epigenetic processes. *Epigenomics* 3:221–234.
- Lack D. 1947. Darwin's finches. New York: Cambridge University Press.
- Laland K, et al. 2014. Does evolutionary theory need a rethink? *Nature* 514:161–164.
- Laland KN, Sterelny K, Odling-Smee J, Hoppitt W, Uller T. 2011. Cause and effect in biology revisited: is Mayr's proximate-ultimate dichotomy still useful? *Science* 334:1512–1516.
- Lamarck J. 1802. Recherches sur l'organisation des corps vivans. Paris: Chez L'auteur, Maillard.
- Levy SF, Siegal ML. 2008. Network hubs buffer environmental variation in *Saccharomyces cerevisiae*. *PLoS Biol.* 6:e264.
- Liu MM, Chan CC, Tuo J. 2012. Genetic mechanisms and age-related macular degeneration: common variants, rare variants, copy number variations, epigenetics, and mitochondrial genetics. *Hum Genomics.* 6: 13.
- Long HK, et al. 2013. Epigenetic conservation at gene regulatory elements revealed by non-methylated DNA profiling in seven vertebrates. *Elife* 2: e00348.
- Longo G, Miquel PA, Sonnenschein C, Soto AM. 2012. Is information a proper observable for biological organization? *Prog Biophys Mol Biol.* 109:108–114.
- MacIntyre A. 1977. Epistemological crises, dramatic narrative and the philosophy of science. *Monist* 60:453–472.
- McClintock B. 1984. The significance of responses of the genome to challenge. *Science* 226:792–801.
- Mihola O, Trachtulec Z, Vlcek C, Schimenti JC, Forejt J. 2009. A mouse speciation gene encodes a meiotic histone H3 methyltransferase. *Science* 323:373–375.
- Nei M, Nozawa M. 2011. Roles of mutation and selection in speciation: from Hugo de Vries to the modern genomic era. *Genome Biol Evol.* 3: 812–829.
- Ohta T. 2011. Near-neutrality, robustness, and epigenetics. *Genome Biol Evol.* 3:1034–1038.
- Olson-Manning CF, Wagner MR, Mitchell-Olds T. 2012. Adaptive evolution: evaluating empirical support for theoretical predictions. *Nat Rev Genet.* 13:867–877.
- Paenke I, Sendhoff B, Kawecki TJ. 2007. Influence of plasticity and learning on evolution under directional selection. *Am Nat.* 170:E47–E58.
- Pembrey ME, et al. 2006. ALSPAC Study Team. 2006. Sex-specific, male-line transgenerational responses in humans. *Eur J Hum Genet.* 14: 159–166.
- Pigliucci M. 2007. Do we need an extended evolutionary synthesis? *Evolution* 61:2743–2749.
- Rebollo R, Horard B, Hubert B, Vieira C. 2010. Jumping genes and epigenetics: towards new species. *Gene* 454:1–7.
- Rieppel O. 2008. Re-writing Popper's philosophy of science for systematics. *Hist Philos Life Sci.* 30:293–316.
- Seisenberger S, Peat JR, Reik W. 2013. Conceptual links between DNA methylation reprogramming in the early embryo and primordial germ cells. *Curr Opin Cell Biol.* 25:281–288.
- Skinner MK. 2011. Environmental epigenetic transgenerational inheritance and somatic epigenetic mitotic stability. *Epigenetics* 6:838–842.
- Skinner MK. 2014a. Endocrine disruptor induction of epigenetic transgenerational inheritance of disease. *Mol Cell Endocrinol.* 398: 4–12.
- Skinner MK. 2014b. Environmental stress and epigenetic transgenerational inheritance. *BMC Med.* 12–153.
- Skinner MK, Anway MD, Savenkova MI, Gore AC, Crews D. 2008. Transgenerational epigenetic programming of the brain transcriptome and anxiety behavior. *PLoS One* 3:e3745.
- Skinner MK, Gurerrero-Bosagna C, Haque MM, Nilsson EE, et al. 2014. Epigenetics and the evolution of Darwin's finches. *Genome Biol Evol.* 6:1972–1989.
- Skinner MK, Manikkam M, Guerrero-Bosagna C. 2015. Epigenetic transgenerational actions of environmental factors in disease etiology. *Trends Endocrinol Metab.* 21:214–222.
- Solary E, Bernard OA, Tefferi A, Fuks F, Vainchenker W. 2014. The Ten-Eleven Translocation-2 (TET2) gene in hematopoiesis and hematopoietic diseases. *Leukemia* 28:485–496.
- Song J, Irwin J, Dean C. 2013. Remembering the prolonged cold of winter. *Curr Biol.* 23:R807–R811.
- Tan Q, Christiansen L, von Bornemann Hjelmborg J, Christensen K. 2015. Twin methodology in epigenetic studies. *J Exp Biol.* 218:134–139.
- Waddington CH. 1953. Epigenetics and evolution. *Symp Soc Exp Biol.* 7: 186–199.
- Yagi S, Hirosawa M, Shiota K. 2012. DNA methylation profile: a composer-, conductor-, and player-orchestrated Mammalian genome consisting of genes and transposable genetic elements. *J Reprod Dev.* 58:265–273.
- Zeh DW, Zeh JA, Ishida Y. 2009. Transposable elements and an epigenetic basis for punctuated equilibria. *Bioessays* 31:715–726.
- Zeh JA, Zeh DW. 2008. Maternal inheritance, epigenetics and the evolution of polyandry. *Genetica* 134:45–54.
- Zuckerlandl E, Cavalli G. 2007. Combinatorial epigenetics, "junk DNA," and the evolution of complex organisms. *Gene* 390:232–242.
- Zwijnenburg PJ, Meijers-Heijboer H, Boomsma DI. 2010. Identical but not the same: the value of discordant monozygotic twins in genetic research. *Am J Med Genet B Neuropsychiatr Genet.* 153B:1134–1149.

Associate editor: Dan Graur