MOLECULAR ECOLOGY

Molecular Ecology (2011) doi: 10.1111/j.1365-294X.2011.05116.x

INVITED REVIEW

Beyond the phenotypic gambit: molecular behavioural ecology and the evolution of genetic architecture

STEVAN A. SPRINGER,* BERNARD J. CRESPI† and WILLIE J. SWANSON‡

Department of Biology, University of Washington, Seattle, WA 98195, USA, †Department of Biological Sciences, Simon Fraser University, Burnaby, BC, Canada V5A 1S6, ‡Department of Genome Sciences, University of Washington, Seattle, WA 98195, USA

Abstract

Most studies of behaviour examine traits whose proximate causes include sensory input and neural decision-making, but conflict and collaboration in biological systems began long before brains or sensory systems evolved. Many behaviours result from non-neural mechanisms such as direct physical contact between recognition proteins or modifications of development that coincide with altered behaviour. These simple molecular mechanisms form the basis of important biological functions and can enact organismal interactions that are as subtle, strategic and interesting as any. The genetic changes that underlie divergent molecular behaviours are often targets of selection, indicating that their functional variation has important fitness consequences. These behaviours evolve by discrete units of quantifiable phenotypic effect (amino acid and regulatory mutations, often by successive mutations of the same gene), so the role of selection in shaping evolutionary change can be evaluated on the scale at which heritable phenotypic variation originates. We describe experimental strategies for finding genes that underlie biochemical and developmental alterations of behaviour, survey the existing literature highlighting cases where the simplicity of molecular behaviours has allowed insight to the evolutionary process and discuss the utility of a genetic knowledge of the sources and spectrum of phenotypic variation for a deeper understanding of how genetic and phenotypic architectures evolve.

Keywords: adaptation, behaviour/social evolution, evolutionary conflict, genetic architecture

Received 11 June 2010; revision received 23 September 2010; accepted 28 September 2010

Introduction: the goals of evolutionary research

A major goal of evolutionary research is to describe how processes such as natural selection and genetic drift act on heritable polymorphisms within populations and result in phenotypic evolution. To reach this goal, biologists must understand the sources of phenotypic variation and their consequences for organismal fitness. Niko Tinbergen first outlined the four distinct lines of inquiry that biologists use to approach the problem of connecting phenotypic variation and fitness, studies of: the developmental history, phylogenetic history, mechanistic function and adaptive function of organismal traits (Fig. 1). Tinbergen proposed this framework of studying the causes and effects of phenotypic variation with reference to behaviour, and although each question can in principle be asked of any trait, the difficulty of achieving all of these goals at once is often considerable. Very few studies have yet integrated answers to all four of Tinbergen's modes of inquiry into a complete picture of the causes and effects of phenotypic variation on fitness. Most studies operate on one side or another of a divide: analysing either the genetic causes of phenotypic variation or the effects of phenotypic variation on fitness. Rarely can both of these perspectives be fully integrated.

The difficulties in undertaking research on the genetics of phenotypic evolution do not arise from

Correspondence: Stevan A. Springer, Fax: (206) 543 3041; E-mail: stevan@u.washington.edu

the demands of asking each of Tinbergen's questions individually—certainly, some are more readily answered than others, but the tools to answer each type of question exist. Genetics and molecular biology have uncovered a common set of regulatory schema underpinning morphological development (Carroll 2005; Stern & Orgogozo 2008). Molecular phylogenetics has eased the study of evolutionary origin and history (Zuckerkandl & Pauling 1965; Felsenstein 1985). Genomic information has enabled genetic mapping experiments to pinpoint small chromosomal regions, sometimes the exact causal mutations, which underlie phenotypic differences between individuals or species (Bradshaw & Schemske 2003; Colosimo *et al.* 2005; Hoekstra *et al.* 2006). Contemporary and historical selection pressures can be estimated against the neutral expectations of quantitative genetic and population genetic models (Tajima 1989; McDonald & Kreitman 1991; Orr 1998; Yang & Bielawski 2000) and against the predicted equilibria of models of phenotype optimization by selection (Maynard Smith 1978).

Unfortunately, systems that are ideal for one type of inquiry are often poorly suited to another. Only a few of the organisms with mature genetic tools and life histories amenable to crossing experiments have well-understood natural ecology, and it has thus proven difficult in any given study system to combine answers to Tinbergen's questions into a complete picture of the effects of naturally occurring genetic variation on organismal phenotype and fitness. This is perhaps the most important and challenging goal of modern evolutionary biology: knowledge of the relationship between genetic variation, phenotypic variation, and fitness is essential if we wish to understand the spectrum of variation available to the evolutionary process. It is worth asking then: what are the practical barriers that limit genetic investigations into phenotypic change in nature? What types of traits are tractable to the methods required to overcome these barriers? Such traits could provide us not only a comprehensive look at genetic and phenotypic evolution, but may open new general approaches to understanding the evolutionary process itself. What existing questions can we answer and what new questions will we ask once we know the historical causes of phenotypic variation and their functional effects on fitness?

Here, we describe a class of phenotypes with features amenable to all four of Tinbergen's questions. Molecular behaviours are traits that sense aspects of the environment or mediate recognition between biological units (primarily: genes, chromosomes, organelles, cells and organisms) by direct physical interaction or developmental modifications of form rather than neural processing of sensory inputs (Box 1). These are attractive evolutionary research systems, because they are jointly tractable to all of Tinbergen's questions. Their genetic constituents can be found by mapping to physical locations in the genome or by biochemical and genetic methods of assessing function. Importantly, the genotype–phenotype relationship is often simple enough to analyse meaningfully: measurable portions of phenotypic change are accomplished by regulatory mutations and amino acid changes to individual proteins. Additionally, the current and historical action of selection can be inferred using models of phenotype optimization and by applying neutral models of molecular evolution to DNA sequence variation.

Behavioural traits are the front line of interaction between parties with conflicting evolutionary interests and are thus prime candidates for studies of adaptation. The prevalence of positive Darwinian selection driving sequence divergence in molecular recognition proteins suggests that new mutational variants often have large effects on fitness (Hughes 2002; Swanson & Vacquier 2002). Selection is also evident in the genomic architecture of the traits themselves: diversity-generating gene structures, and patterns of linkage between functionally related genes are often too complex to explain by neutral processes. We argue that the genetic constituents of molecular behaviours have patterns of regulation and diversification that are consistent with their involvement in evolutionary conflict. Understanding the genetic causes of phenotypic variation is thus key to understanding the evolution of these simple systems. To the extent that adaptive molecular and developmental changes contribute to recognizable behavioural phenotypes, they represent interesting research subjects for molecular and organismal biologists alike.

We begin by summarizing methods of finding the genetic causes of phenotypic variation (connecting genotype and phenotype) and of inferring the importance of selection in shaping the evolution of genetic or phenotypic variation (connecting genotype and fitness, or phenotype and fitness). We then review some of the early empirical and theoretical results of this research programme: simple molecular mechanisms that are now known to underlie adaptive variation in many important organismal phenotypes. We conclude by considering how different forms of evolutionary conflict within and between alleles and individuals shape the evolution of genetic architectures. Our aim is not to describe the genetics of behaviour in general, rather we focus on describing research on behavioural phenotypes whose proximate causes are mechanistically simple modifications of physical form: molecular and developmental.

Box 1 Behaviour without brains

Behaviours are inherently interesting to biologists, because they are among the most variable and multifaceted of all phenotypic traits. Studies of behaviour therefore typically presume a complex genetic basis (Grafen 1984), but in fact some of the most intricate neural behaviours have analogs at the molecular level. Direct molecular interactions can produce behaviours with all the sophistication of their most complex counterparts, without the vast physiological machinery required of sensory and neural mechanisms.

The strategies used by athletes during competition give a useful example of a complex behavioural conflict that could be negotiated either by skilled competitors or by simple molecular mechanisms. Enthusiasts of oval-track bicycle racing celebrate the sport as a test of skill, in which strategy wins more races than swiftness. Groups of riders cut through wind more efficiently than lone individuals, and racers must therefore negotiate membership in a group or be left behind. One especially tactical contest, the 'Miss-and-Out' race, begins with a group of riders circling the track, driven by the removal of the last-place rider each lap. Single riders cannot distance themselves from the pack, and so alliances form and break as groups attempt to lead the charge and evade the chopping block. Eventually, the race is whittled to only two contestants, former friends and now bitter enemies, who must sprint the final lap to victory or defeat.

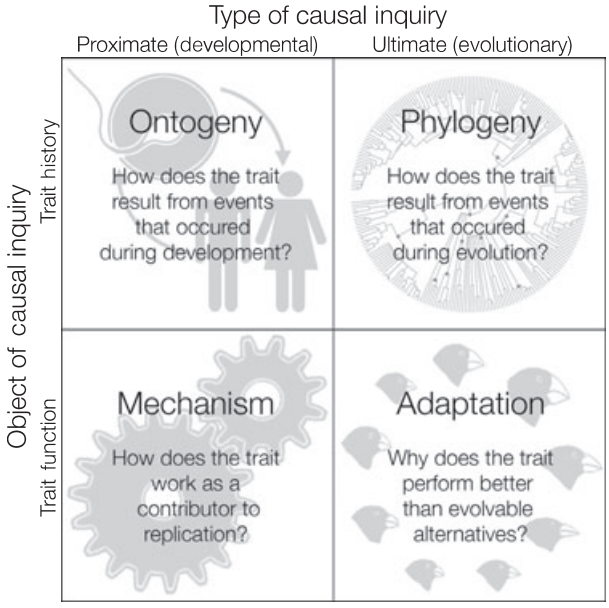

Fig. 1 Tinbergen's four modes of biological inquiry.

The sperm of possum and silverfish face a similar dilemma. In these taxa, sperm pairs swim more efficiently than loners, but only one sperm can successfully fertilize an egg (Moore & Moore 2002). Sperm that do not find a partner find themselves out of the running. Groups of sperm compete with each other in a race towards the egg, but individual sperm must decide when to abandon their partners for the final sprint to the fertilization line. Getting to the winner's circle requires that sperm enact a number of complex behaviours: settling on a suitable partner, deciding when to defect, and how to best stab the estranged ally in the back. What is most remarkable about these strategies, lauded for their complexity and subtlety by cycling devotees, is that they are employed by sperm—without the benefit of a brain or complex neuro-sensory system. Many interactions between genes, organelles, cells, tissues and individuals involve negotiation, but not all of these debates are under neural control. Indeed, in much of biological communication, it is direct molecular interactions that do all of the talking.

Molecular behaviours share more with their neural counterparts than such a simple analogy can relate: not only are the conflicts and strategies superficially similar, but their evolution can be described by identical models. Opti-

Box 1 *(Continued)*

mization models are agnostic to the cause of variation in a trait and concern themselves only with the range of heritable strategies (Maynard Smith 1978). The mechanism of inheritance is assumed to be unimportant in terms of its effect on the end point of selection. Behaviours with a simple mechanistic basis are tractable to genetic analysis, and thus allow us to test the optimality model's assumption that genetic constraints do not influence the outcome of evolution by selection. New phenotypic variants of molecular behaviours arise by mutation, in discrete units of quantifiable phenotypic effect, allowing experimental exploration of the spectrum of phenotypic variation available to evolution (Weinreich *et al.* 2006; Bridgham *et al.* 2009; Field & Matz 2010). Despite their simple nature, molecular behaviours contribute to many interesting organismal phenotypes, and this combination of simplicity and obvious organismal relevance is ideal if one wishes to study how genetic constraints on phenotypic variation shape the process of adaptation.

Methods of connecting variation in genotype, phenotype and fitness

Connecting genotype and phenotype: the genetic bases of phenotypic variation

Finding the material differences that encode and cause variation in a focal trait is not a trivial problem. Phenotypic variation is often the product of contributions from the environment and from many genetic loci that can be physically located anywhere in the genome. Tracing the source of natural phenotypic variation to specific genetic differences involves first finding candidate genes—chromosomal regions whose different alleles are responsible for differences in phenotype—and second evaluating the phenotypes and fitness effects associated with natural allelic variants of these genes. There are two general approaches to choosing, from the tens of thousands of genes in a genome, candidates whose allelic variants might create variation in phenotype. Top-down (or forward) approaches start with a phenotype of interest and seek its molecular underpinnings. Bottom-up (or reverse) approaches begin with a molecule or polymorphism of interest and seek its role in creating phenotypic variation (Box 2).

Top-down and bottom-up methods are complementary approaches to finding the genetic differences that cause phenotypic variation, each with its own advantages and disadvantages. Top-down genetic mapping allows one to study the genetics of nearly any specifiable trait, especially if the organism can be crossed and raised in large numbers and controlled environmental conditions. Importantly, because these methods rely solely on physical linkage, they are agnostic to the mechanism that creates phenotypic variation: regulatory, structural and epigenetic polymorphisms can all be located. Top-down methods also allow one to quantify the relative contributions of multiple causal variants to the total phenotypic variation, and targeted crossing designs can uncover the role of nonadditive genetic interactions such as epistasis and dominance (Glazier *et al.* 2002). Bottom-up candidate gene approaches circumvent the technical difficulties of pure physical mapping at the cost of a rigorous quantitative knowledge of the myriad genetic contributions to phenotypic variation. Because pure bottom-up approaches begin without an independent confirmation of linkage to phenotype, candidates must be chosen and evaluated carefully to avoid investing effort into a gene that does not create variation in the trait of interest.

Top-down and bottom-up methods yield candidates for functional investigation, but both methods must eventually look up towards phenotype, to experimentally confirm that polymorphisms in the candidate gene actually have the expected phenotypic effects. Bottom-up approaches reap an advantage in this latter step because candidates can be chosen specifically for their functional tractability or because they contain genetic variation that is likely to be evolutionarily interesting. With top-down approaches, one rarely knows in advance what the genetic basis of natural phenotypic variation will look like: if the contributing genes have been subject to selection or if their individual influences on phenotypic variation are large enough to be tractable to functional analysis. Establishing the relationship between genetic and phenotypic variation gives clues to the mechanism of phenotype production, but not to the causes of evolutionary change. These insights must be gained by determining the relationship between natural genetic or phenotypic variation and fitness.

Box 2 Top-down and bottom-up methods of connecting genotype and phenotype

Top-down: choosing candidates by their chromosomal location

Many studies have used physical linkage to find genomic regions with polymorphisms that contribute to phenotypic variation and to place these regions on a genetic map of markers located at intervals across the genome (Fig. 2). These linkage-mapping studies begin with a natural pedigree of familial relationships over a number of generations, or experimentally produce such relationships by breeding individuals that differ in a phenotype of interest. To ensure that the differences being studied are heritable, each individual must retain its diagnostic phenotype when grown in a shared set of environmental circumstances, often referred to as a common garden.

Classically, linkage-mapping experiments breed parents (P), and then offspring from the first (F1) generation are crossed with one another to create a second (F2) generation. Recombination during meiosis in the F1 generation creates F2 individuals with chromosomes that are mosaics of physical segments derived from each of the original parents (P). Phenotype–marker associations occur when the opportunity for recombination, between markers that identify each parent and polymorphisms that cause phenotypic variation in the focal trait, is limited by their close physical proximity on the chromosome. When physically linked, the marker variants and phenotypes that characterize a given parent are inherited together in the same chromosome segment and thus found together in the majority of F2 individuals. Chromosomal segments that are physically associated with phenotypic variation in a focal trait despite the opportunity for recombination are called quantitative trait loci (QTL).

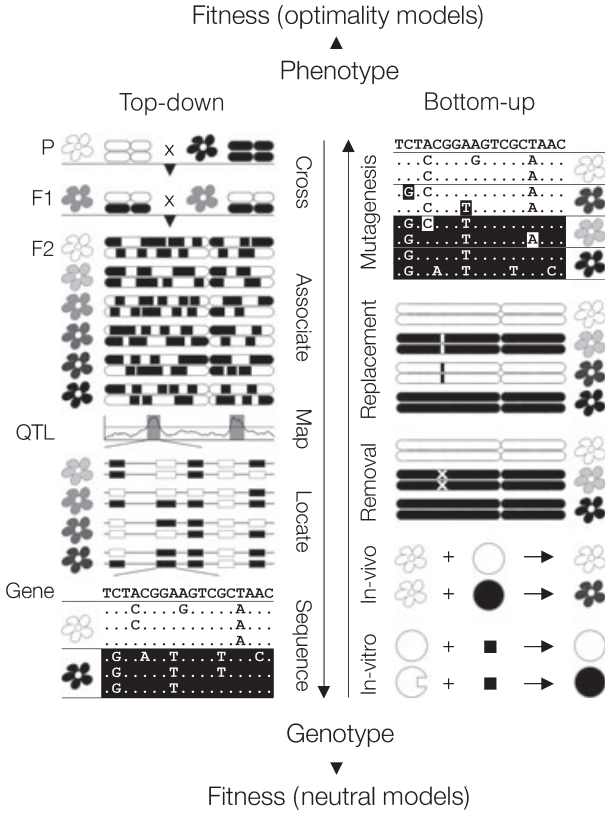

Fig. 2 Top-down and bottom-up methods of connecting genotype and phenotype.

A number of studies have mapped QTL that contribute to natural phenotypic variation: by transplanting experimentally created F2 individuals into natural ecological circumstances (Bradshaw & Schemske 2003), or by using recombination in natural hybrid zones to find markers that remain associated with the phenotypes of each parental taxon despite generations of crossing (Teeter *et al*. 2008). In some cases, appreciable portions of the total phenotypic variation can be attributed to a handful of QTL with individually large effect on phenotype. In practice, these

Box 2 *(Continued)*

QTL can contain hundreds or thousands of gene regions each of which is a candidate, a potential cause of the observed phenotypic difference between parental individuals (Glazier *et al.* 2002).

Closing the gap between markers and causal variants by physical mapping requires increasing both the number of markers and the number of recombination events surveyed. Finding markers is no longer a limitation, as the resolution of physical mapping is sufficient to bracket chromosomal regions small enough to sequence in their entirety. However, physical mapping can rarely achieve the resolution required to link just one causal variant to phenotypic differences among individuals, because doing so requires surveying an impractical number of recombination events. Despite the difficulty of this approach, several studies have been able to map QTL associated with natural phenotypic variation down to their constituent genetic polymorphisms (Chouard 2010). Most of these successful physical mapping efforts take assistance from complementary techniques to pinpoint the exact genetic polymorphisms that cause phenotypic variation. In systems with well-developed genetic tools, QTL can be chopped into manageable pieces by deleting or introgressing short chromosomal segments to test their phenotypic effect directly (Bradshaw & Schemske 2003; Presgraves *et al.* 2003). In organisms with a sequenced genome, known genes within a QTL region can be targeted for further functional evaluation based on their biological properties (Shapiro *et al.* 2004; Hoekstra *et al.* 2006).

Bottom-up: choosing candidates by their biological properties

Bottom-up methods are sometimes directly referred to as candidate gene approaches. They choose genes for evaluation based on attributes other than (or in addition to) their physical position on a chromosome: by the location, timing, or amount of gene expression, or by experimentally established functional properties. Bottom-up studies often survey discrete organismal structures (organelles, cells, tissues, organs and secretions) to identify genes of potential functional interest. For example, screens of sperm protein content have identified candidate genes involved in sperm–egg interaction (Swanson & Vacquier 2002). A similar logic underlies studies that identify candidate genes expressed abundantly in specified organs, physiological states, developmental stages, sexes or species. Proteomic techniques can identify the protein content of a candidate structure directly, by matching the properties of its resident proteins to those predicted by a reference genome or cDNA library. Proteomic discovery studies can also target specific biological functions using isotopic labels to hide all but the desired set of proteins (Findlay *et al.* 2008). Because of their rapidity, scale and specificity, proteomic methods could also be used to narrow the search for candidates following physical mapping by finding only those proteins whose biological features suggest a role in phenotype production and whose chromosomal positions lie within a given QTL.

The functional properties of a desired class of proteins can also be used to narrow a list of candidate genes. For example, binding partners that interact during fertilization have been identified by isolating proteins from the outer membranes of eggs and retaining those with a biochemical affinity for a given sperm protein (Swanson & Vacquier 1997). The biological function of altruistic greenbeard genes necessitates self-interaction, and these can be identified by functional screens for proteins with self-affinity (Benabentos *et al.* 2009). Methods of identifying protein interactions such as yeast two-hybrid systems, which measure biochemical affinity between two proteins by fusing them to the translation apparatus of a reporter gene and measuring its expression, can quickly identify interactions between many pairs of proteins. Chromatin immuno-precipitation experiments identify the affinity of DNA-binding regulatory proteins by allowing them to bind their targets and isolating the DNA–protein complex. Protein–protein and protein–DNA affinity experiments can now be deployed on a whole-genome scale, measuring biochemical affinities between all pairs of proteins (Schwikowski *et al.* 2000), or the complete DNA-binding repertoire of a given regulatory protein (Buck & Lieb 2004).

Functional properties can also be established by direct experimental manipulation of a trait's genetic underpinnings. Gene knockouts remove the presence of a candidate gene to examine its involvement in phenotype production. In some cases, the phenotypic effects of knockouts in a model system resemble those of naturally occurring variation (Shapiro *et al.* 2004). Mutagenesis experiments examine the spectrum of phenotypic variation produced by artificial mutations to a candidate gene or to the genome as a whole (Weinreich *et al.* 2006). Interestingly, these too can create organismal phenotypes that resemble those of closely related natural species, suggesting that genetic mechanisms may profoundly influence the direction of phenotypic change during some instances of natural evolution (Koufopanou & Bell 1991).

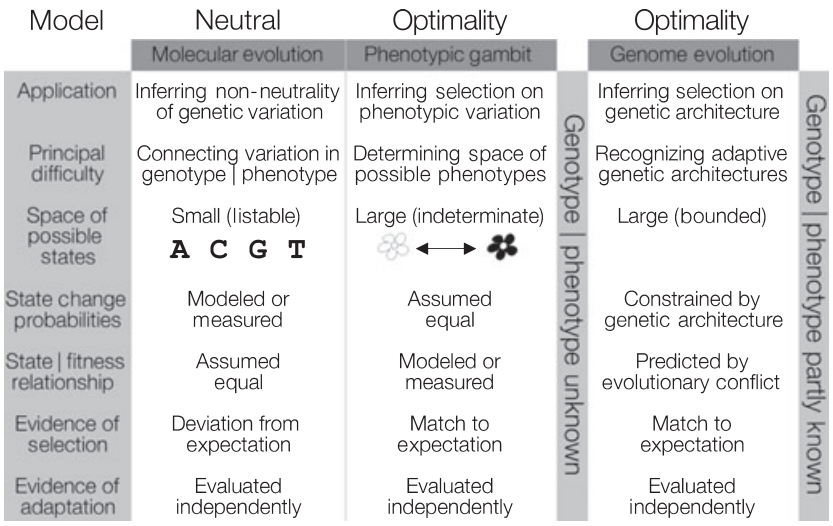

Fig. 3 Neutral and optimality methods of associating genotype, phenotype and fitness. Neutral models assume all phenotypic states have equal fitness to predict patterns of genetic evolution in the absence of selection. Optimality models assume all small phenotypic changes are equally likely to predict the phenotypic outcome of a specified selection pressure. Many genes have features (patterns of regulation, physical linkage to other genes, or combinatorial mechanisms of generating diversity) that appear to be the outcome of selection on genetic architecture itself. We suggest that many such genomic features are recognizable outcomes of evolutionary conflict.

Connecting genotype and fitness: neutral models of genetic evolution

Another way to choose candidate genes is to estimate the evolutionary pressures that cause them to change over time. A history of selection in the evolution of a gene or chromosomal region can be inferred using DNA sequence alone, by comparing the aspects of sequence evolution to the expectations of population genetic models. These models describe expected patterns of DNA sequence polymorphism or divergence under neutrality, as though selection had no effect on sequence evolution. Neutral expectations can be compared with naturally occurring patterns of sequence evolution to determine whether change in a candidate gene can be explained without invoking non-neutral processes: violations of the null model suggest that patterns of genetic change were influenced by selection via their effects of phenotype (Fig. 3).

Tests of neutrality take a number of general forms: among others, those that examine the distribution of sequence polymorphism in multiple individuals of the same population or species (Tajima 1989), those that use patterns of sequence variation within species to predict divergence between species (McDonald & Kreitman 1991) and those that compare the rates of amino acid changing and silent mutations in protein-coding regions (Yang & Bielawski 2000). Methods of detecting selection can be adapted to identify pairs of proteins that interact over evolutionary time. If change in one co-evolving locus directly selects for a compensatory change in its partner, then the evolutionary rates of functional portions of the two genes are expected to show higher correlation than between two functionally unrelated genes. Interlocus co-evolution, and an implied history of functionally important molecular interaction, should therefore be evident from phylogenetic analyses of co-evolutionary rates (Clark *et al.* 2009). Computational methods of detecting selection are targeted enough to assess selection on single genetic polymorphisms and can examine selection in un-manipulated natural populations; these too are now being implemented on whole genomes (Andolfatto 2005; Nielsen *et al.* 2005; Tang *et al.* 2007; Akey 2009).

Selection is often assumed to be the primary cause of phenotypic evolution, but for most genes selection operates primarily as a stabilizing force (Nielsen *et al.* 2007). Proteins that control molecular behaviours are a major exception to this rule: pervasive adaptive evolution is often found at genes that mediate direct biological interactions (Hughes 2002; Swanson & Vacquier 2002). The striking prevalence and intensity of selection on genes involved in molecular behavioural interactions may exist either because tests of selection are not powerful enough to detect selection spread across the many loci that contribute to divergence in complex quantitative traits, or because the simple nature of the genotype–phenotype relationship and the strong and continued selection pressure exerted by evolutionary conflict concentrates the signal of selection on proteins at the interface of interaction between individuals. Regardless, the genetic constituents of biological interactions are excellent systems to study how selection shapes the evolution of traits whose phenotypic effects are relevant to fitness.

Of course, connections between genotype and fitness alone are not sufficient to test adaptive hypotheses. Comparisons to population genetic models allow us to ask *how* a gene evolves—*by what evolutionary processes?*, but not *why* it evolves—*for what adaptive purpose?* Questions about adaptation cannot be sensibly asked without information relating the phenotypic effects of genetic

changes with their effects on survival or reproduction, because these are what selection actually acts on.

Connecting phenotype and fitness: the phenotypic gambit and optimality models of phenotypic evolution

The difficulty of finding and understanding the genetic and developmental mechanisms that create phenotypic variation caused many evolutionary biologists to temporarily abandon this problem and focus their effort solely on the relationship between phenotype and fitness (Grafen 1984). This approach, termed the phenotypic gambit, posits that genetic constraints do not need to be explicitly incorporated into models of phenotypic evolution: any phenotype that can be reached by mutation can be acted on additively by selection, and the effects of neutral evolutionary processes are negligible. Using these assumptions, models of phenotypic change can be built which describe the dynamics of evolution that result from applying a specified selection pressure—formally, a series of state equations that describe the relationship between phenotype and fitness— to a given suite of phenotypic variation (the strategy set).

The assumptions of the phenotypic gambit, if considered strictly, are usually not met: it is a simplifying assumption meant to focus attention on the selective pressures that optimize phenotype to environment, and away from genetic complexities. In contrast to neutral models that ask *what would happen without selection?*, models of adaptive evolution ask *what would happen by selection?* (Fig. 3). Optimality methods do not test whether a trait is an adaptation or not: they lack the ability to distinguish a direct adaptation from a pleiotropic by-product of selection on another aspect of form (Lewontin 1978). Rather, optimality methods ask whether a particular set of assumptions about the fitness consequences of variation in a trait are sufficient to predict an observed phenotype (Maynard Smith 1978). Modelling phenotypic evolution with this approach has lead to the concept of an evolutionary stable strategy (ESS: an evolutionary equilibrium where selection cannot replace the existing phenotype with any new alternative) and to deep symmetries between evolutionary and economic dynamics.

Given the nature of its assumptions, one might expect that the phenotypic gambit would not work well, but in fact many phenotypes have been found to match the optima predicted by models built without explicit genetic constraints: (Parker & Maynard Smith 1990). Paradoxically, the success of the approach could lie in the sheer complexity of the genetic systems that it assumes away. Many genes of small effect can underpin variation in complex phenotypes, and in this circumstance, the genetic mechanisms that could cause the gambit to fail bear little individual weight (Grafen 1984). If this is the case, identifying genetic variants underlying complex phenotypes may tell us little more about the constraints on their evolution than classical quantitative genetic approaches of measuring genetic co-variances (Arnold 1992).

The effort required to uncover the genetic basis of natural variation in a phenotypic trait is considerable. We now have the tools to relate genotype and phenotype, but from a practical standpoint, genetic studies of phenotypic evolution are only worthwhile if they can tell us more about the evolutionary process than a purely phenotypic approach can. A primary focus of genetic studies of adaptation should therefore be systems where the phenotypic gambit's assumptions cannot provide a realistic approximation of phenotypic evolution. Behavioural interactions such as those resulting from direct contact between proteins, noncoding RNAs or DNA sequences minimally involve single genes or pairs of genes. Their genetic basis is less complex than that of most neural behavioural traits, and it is thus more likely that the evolutionary dynamics of individual genes will result in violations of the phenotypic gambit's assumptions. For example, heterozygote advantage is an example where the phenotypic gambit's assumptions fail because of the details of the underlying genetic mechanism: one cannot build a model of a haploid system that accurately predicts the evolutionary dynamics of overdominant selection on a single diploid locus (Grafen 1984). Constraints on the suite of phenotypic variation available by mutation may also influence phenotypic evolution: though the form of such constraints is currently less well understood (Haldane 1933; Olson 1999; Weinreich *et al.* 2006). Molecular behaviours are therefore an excellent means to study the genetics of adaptive phenotypic evolution: not only are they often subject to strong selection and tractable to functional analysis, but also it is likely that genetic constraints will tangibly influence their evolution.

Results of molecular behavioural ecology

Early empirical and theoretical results indicate that fascinating evolutionary dynamics can occur in systems where the phenotypic gambit fails (Hayashi *et al.* 2007). Next, we describe some of the early results of molecular behavioural ecology, focusing on areas where a genetic understanding of the causes of phenotypic variation has yielded insight on the process of evolution. We survey existing research on molecular behaviours, highlighting examples where the study of simple elements engaged in evolutionary conflict has contributed to our understanding of six fundamentally important biological phe-

nomena: sex, reproduction, speciation, sociality, development and disease. We hope to show that natural genetic variation in the components of these simple phenotypes has tangible effects on their evolution (and our ability to understand their evolution) and that they are thus worthwhile studying from an explicit genetic perspective despite the considerable difficulty of doing so.

Sex

One of the truly surprising results to come from comparisons to neutral models of sequence evolution is the rapid evolution of proteins that function in central biological processes (Nielsen *et al.* 2005, 2007). Some of these examples were expected: Hamilton's interpretation of the Red Queen hypothesis proposed that host–pathogen conflict could drive a co-evolutionary arms race that continually favours the rare host genotypes produced by sexual recombination (Hamilton *et al.* 1990). Direct molecular recognition events are key components of host–pathogen interactions and, as expected, their constituents evolve rapidly in response to conflict (Paterson *et al.* 2010). The first functionally characterized example of positive selection came from the adaptive evolution of vertebrate MHC proteins, which alert the immune system to the presence of a pathogen by presenting its peptides on the surface of an infected cell (Hughes 2002). Host–parasite conflict also drives the evolution of genetic architectures, such as the VDJ recombination mechanism, that create the molecular diversity that allows slowly evolving hosts to recognize quickly evolving pathogens (Litman *et al.* 2007).

Counter-evolution allows pathogens to evade detection, as evident from comparisons of parasitic and mutualistic endosymbionts: the surface proteins of parasitic Wolbachia evolve by positive selection, homologous proteins in mutualistic strains do not (Jiggins *et al.* 2002). Although mechanistically simple, the means by which parasites evade detection can be remarkably sophisticated. Some have proteins that engineer their host's surveillance system, mimicking a protein that sends an 'all clear' signal to the immune system (Elde *et al.* 2009); others camouflage themselves by decorating their surfaces with host antigens, or deliberately modify their own surface antigens to avoid detection (Marsh & Helenius 2006; Mercer & Helenius 2008). It has been proposed that sex itself evolved in response to pressures placed on hosts by their parasites—be they discrete pathogenic organisms, or ultra-selfish alleles replicating within the host genome (Hamilton *et al.* 1990; Haig & Grafen 1991; Hurst *et al.* 1996). Molecular evidence of rapid host–parasite co-evolution shows that these predicted conflicts are real and that simple molecular mechanisms can create variation with tangible fitness consequences for both interacting parties.

Reproduction

Examples of positive selection also come from classes of proteins that might otherwise have been expected to be functionally conserved, for example, reproductive proteins often evolve rapidly despite their critical importance to gamete production, transport, storage and fertilization (Swanson & Vacquier 2002). The exact form of selection is not known, but rapid reproductive protein evolution is thought to be a consequence of two forms of sexual selection: competition, such as male–male competition for fertilizations in polygynous mating systems, and conflict, such as male–female antagonism over fertilization rate (Hayashi *et al.* 2007). These two forms of sexual selection have markedly different consequences for phenotypic evolution. Sexual competition leads to symmetric arms races that continually exaggerate the phenotype under selection, whereas sexual conflict leads to asymmetric arms races and Red Queen dynamics—phenotypic stasis in the face of the rapid evolution of each counter-evolving party. These two types of arms race could be potent motivators of evolutionary change (Dawkins & Krebs 1979), but determining the importance and form of sexual selection in reproductive protein evolution requires the exclusion of alternative selective explanations, which has thus far proven difficult.

In the rare cases where they are known, the phenotypic effects of natural reproductive protein variants within populations are consistent with sexual selection. For example, abalone sperm protein lysin and its egg receptor VERL influence sperm–egg compatibility and both change rapidly by continually co-evolving; their evolutionary rates are correlated along phylogenetic lineages. Patterns of polymorphism in these proteins within abalone species resemble a unique prediction of sexual conflict: lysin is monomorphic within the pink abalone, but VERL has diversified into two clades, possibly to prevent sperm from specializing on either type of egg (Clark *et al.* 2009). In sea urchins, evidence from spawning experiments in natural conditions suggests that fertilization ecology has important consequences for patterns of conflict and collaboration during mating. Molecular polymorphism in bindin (a sperm protein that influences sperm–egg compatibility) is maintained by frequency and density-dependent sexual selection on males and females (Levitan & Ferrell 2006). In both of these systems, linkage disequilibrium has been reported between sperm–egg recognition proteins suggesting that these genes influence sperm–egg compatibility in natu-

ral populations (Palumbi 1999; Clark *et al.* 2009). Coevolution driven by sexual selection may thus explain the remarkably rapid evolution of reproductive proteins, the maintenance of reproductive compatibility within species despite this divergence and perhaps even the evolution of reproductive barriers between species. In internally fertilizing organisms, molecular interactions between seminal fluid proteins and the female reproductive tract can also influence reproductive compatibility (Findlay *et al.* 2008; Matute 2010).

Speciation

Prezygotic isolation because of differences in postmating reproductive behaviour or sperm–egg interaction is often a primary barrier to gene exchange between closely related sympatric species. Studies of sperm–egg recognition proteins across species yield a number of consistent patterns: they are often subject to diversifying selection that causes rapid divergence between species, and selection is particularly strong in taxa with broadly sympatric ranges (Yang *et al.* 2000). Studies of sperm–egg recognition protein variation within species have shown character displacement in sympatry (Geyer & Palumbi 2003), and positively selected divergence in association with geographic areas of secondary contact and hybridization (Springer & Crespi 2007), suggesting a direct role for the rapid divergence of reproductive proteins in speciation. In *Drosophila yakuba*, the adaptive evolution of prezygotic isolation in response to reduced hybrid fitness (reinforcement) depends not only on the neural behaviours that govern mate choice, but on the physical and chemical interactions that take place after mating and before fertilization (Matute 2010). Direct molecular interactions are fundamental to many of the barriers that reproductively isolate species; studying the molecular and phenotypic evolution of these traits has the potential to establish the role of selection in the origin of species in general.

The major class of postzygotic isolating barrier, Dobzhansky–Muller incompatibility causing hybrid inviability or sterility, can also result from direct interaction, or lack thereof, between gene products co-existing in the hybrids of incipient species. Recent studies have uncovered the genetic basis of postzygotic barriers to gene exchange, and in many of the currently identified cases, reduced hybrid fitness is a consequence of disrupted protein–protein or protein–DNA interactions (Presgraves 2010). By their nature, Dobzhansky–Muller interactions do not easily occur within populations to initiate the evolution of reproductive isolation in an otherwise freely mating population. As a result, these reproductive barriers may evolve most often between populations that do not exchange genes, and their effects on postzygotic isolation are by-products, not direct targets of selection.

Interestingly, the genes underlying Dobzhansky–Muller interactions are often evolving rapidly, but most are not obviously associated with different environmental adaptations; rather, these incompatibilities appear to be pleiotropic consequences of evolutionary conflicts that have resolved differently in independent populations. It is perhaps not surprising then that these traits are sometimes associated with the fastest evolving interactions in the genome, evolutionary conflicts between hosts and parasitic organisms or ultraselfish alleles (Presgraves 2010). The genetic basis of postzygotic isolation shows that not every important evolutionary change is a direct consequence of selection. When reproductive compatibility is not directly maintained by selection, it can quickly be lost as a pleiotropic by-product of rapid evolution caused by evolutionary conflicts operating on other components of phenotypic variation.

Sociality

Another of Hamilton's many contributions to evolutionary thought was his recognition that cooperation among relatives can evolve if the indirect fitness benefits of a behaviour exceed the direct fitness costs (Hamilton 1964). For costly helping behaviours to evolve, their benefits must be directed preferentially to other individuals that carry the same helpful allele. Organisms solve this identity crisis one of two ways: probabilistically, by helping kin, and mechanistically, by recognizing the presence of a helpful allele in other individuals and directing benefits to them specifically (Gardner & West 2010).

Although once thought to be only a theoretical curiosity, genes causing mechanistic (greenbeard) recognition have now been found in many different contexts that often involve direct physical interactions between proteins and influence processes as diverse as survival in harsh environmental conditions and maternal–foetal conflict (Haig 1996; Smukalla *et al.* 2008). In cooperatively reproducing social microorganisms like *Dictyostelium* slime moulds and *Myxococcus* bacteria, cheaters benefit by increasing their representation in reproductive spores and shirking their contribution to nonreproductive tissues. Selection to associate with cooperators, and to cheat, results in sophisticated recognition behaviours from simple mechanisms: molecular crime and punishment. (Fiegna & Velicer 2005; Benabentos *et al.* 2009). Simple molecular recognition mechanisms allow unicellular organisms to engage in one of biology's most sophisticated negotiations, distinguishing helpful individuals from cheaters to allow the evolution of

stable social systems of cooperation (Crespi 2001). Solutions to this conflict during transitions to multicellularity may have had a similarly simple mechanistic basis.

Development

For multicellular organisms, some of the most important and persistent evolutionary conflicts are a consequence of the asymmetric fates of somatic and germ cells. Vertically transmitted elements (alleles, sex chromosomes and intracellular pathogens) can gain a selective advantage by altering host development: sending somatic cells to the germline or biasing the offspring sex ratio, to increase the proportion of the host's gametes that contain the ultra-selfish element (Hurst *et al.* 1996; Jiggins *et al.* 2002). The conflict of interest caused by these genomic 'outlaws' can be clearly seen by considering instances where the element over-represents itself by selectively killing gametes that do not contain copies of it. The benefit to the selfish element is an increased proportional representation in the next generation; the cost to the host, and to other nonselfish elements in the genome, is the material investment in gametes lost to the actions of the selfish element. The rapid evolution sometimes observed at proteins involved in gametogenesis and germline specification may be a response to such conflict between hosts and their vertically transmitted parasites (Eckmann *et al.* 2004; Bauer DuMont *et al.* 2007).

Other aspects of development are also associated with positive diversifying selection and genetic mechanisms of generating diversity. Sensory receptor proteins that physically receive information from the environment can evolve rapidly or diversify in response to changing conditions. Opsins are proteins expressed in the retina that are receptive to specific wavelengths of light. In the Coelocanth, historical movement to deep water is associated with selection to tune the receptivity of opsins to the longer wavelengths of light that predominate at depth (Yokoyama *et al.* 1999). Olfactory genes diversify into families of duplicate genes presumably in response to selection for recognizing volatile compounds in the environment (Gilad *et al.* 2003). Diversifying selection in one context can also be co-opted into other developmental roles. The myriad connections of neurons made during insect brain development are, in part, a consequence of self-adhesion avoidance behaviour caused by Dscam, a protein whose genetic architecture allows it to generate thousands of variants within a single individual (Hattori *et al.* 2009). Dscam variants also function in innate immunity, although it is not known which of these two recognition functions is the original (Watson *et al.* 2005).

Disease

To explain how genetic constraints can violate the phenotypic gambit's assumptions, Grafen (1984) described a disease phenotype that persists because its simple genetic basis prevents absolute optimization. In human populations suffering from malaria, individuals have a selective advantage when they are heterozygous at the gene that causes malaria resistance and sickle-cell anaemia. Mendelian segregation recreates homozygous genotypes each generation and prevents all individuals in a population from attaining the optimal heterozygous phenotype. This is an evolutionary conflict because of antagonistic pleiotropy: the effects of the disease allele on fitness reverse in different contexts (depending on which allele it is paired with) and so no optimum can exist.

Evolutionary conflicts because of simple genetic interactions happen in many different biological contexts (Hurst *et al.* 1996). In some cases, rapid protein evolution in response to selection on one aspect of form compromises other phenotypic features that happen to share a common genetic basis, and positive selection has been found to act on a number of so-called disease genes (Nielsen *et al.* 2005). In others, simple molecular players mediate chronic conflicts between interacting parties. Maternal–foetal conflict over resources provides a particularly clear example. In the foetuses of some mammals, the paternally expressed gene IGF2 stimulates growth, while the maternally expressed M6P/IGF2R gene codes for a 'decoy' receptor that degrades IGF2, reducing its growth stimulatory effects. The foetus is thus an intermediary in a conflict between mother and father, whose actions are set by the appropriate epigenetic regulation of alleles acting on each parent's behalf (Moore & Haig 1991).

Grafen suggested that the evolution of genetic architecture can eliminate constraints on phenotypic optimization: if the sickle-cell gene was duplicated, each daughter-locus could fix for the alternative allele and all individuals would effectively be heterozygous. Mechanisms that overcome pleiotropic constraints on evolutionary change can be favoured when those individuals whose phenotypes are constrained from reaching a particular value have a lower fitness than those whose are not. When phenotypes cannot reach optima because genetic architectures and regulatory mechanisms of overcoming pleiotropic constraints have not evolved, or when evolved mechanisms become dysregulated, we recognize the negative fitness consequences of these genetic trade-offs as disease. Understanding the circumstances that cause evolutionary constraints and that favour particular mechanisms of release is thus very important. We suggest that these problems can be

approached: by studying the elements in conflict, and the arena in which these conflicts play out, and by determining how these two features influence the corrective mechanisms that evolve in response to evolutionary conflict.

Beyond the phenotypic gambit: applying optimality models of selection to the evolution of genetic architecture

Early studies of the phenotypic effects of genetic mutations examined experimental mutants of laboratory organisms. The negative fitness consequences of these artificial variants persuaded researchers that most mutations with phenotypic effects large enough to measure are straightforwardly disadvantageous and of little interest to those studying adaptation (Grafen 1984). However, in the absence of information about the phenotypic effects of natural mutations—the genetic constituents of actual instances of phenotypic adaptation—universal assertions about the nature of mutations of measurable effect are difficult to justify. Not all mutations of measurable phenotypic effect or fitness effect are straightforwardly disadvantageous. There is now abundant evidence of positive selection in nature, much of it acting on molecular recognition proteins. Finding genetic variants that fixed during bouts of natural phenotypic evolution and evaluating their effects on phenotype and fitness is possible, but only a few studies have yet confirmed that such changes have functional effects that match hypotheses proposing which components of phenotypic variation are under selection (Bishop 2005; Zhang 2006; Yokoyama *et al.* 2008). Making these connections, between selection inferred from natural genetic variation and its natural phenotypic consequences, will be a primary task of modern evolutionary biology (Nielsen 2009).

It is now also clear that many genes have evolved genetic architectures that violate the assumptions of the phenotypic gambit. A clear example is the VDJ recombination system of vertebrate adaptive immunity, with its combinatorial recombination mechanism that creates massive amounts of molecular diversity each generation (Litman *et al.* 2007). As noted above, similar mechanisms of combinatorial diversity generation have been found in recognition proteins involved in everything from sperm–egg recognition to neuron self-avoidance during brain development (Moy *et al.* 2008; Hattori *et al.* 2009). In these cases, the molecular diversity itself is not heritable, but the mechanism of producing it generation after generation is. Such mechanisms are simply too complex to have evolved by chance: they demand a selective explanation. One can imagine applying an optimization approach to genetic evolution that describes the fitness pay-offs of genetic variation to predict which mechanisms of gene regulation, physical association or diversification are expected to evolve under a given selection pressure. However, we are currently unable to apply such models to the evolution of genetic architecture because we lack a framework for assigning fitness pay-offs to different architectural variants.

The difficulty of understanding how selection shapes the evolution of genetic architecture is twofold. First, it is often hard to make advance predictions about which genetic changes would be favoured by a particular selection pressure. For instance, it would currently be difficult to predict with any certainty, a plausible genetic response to ecological selection. Second, it can be hard to recognize the adaptive value of an unusual genetic architecture, and instances of adaptive genetic evolution may therefore go unnoticed. Applying optimality models of selection to genetic systems is not impossible, but we lack intuition about how to model the fitness consequences of genetic variation: translations between the selection applied to a phenotype and the genetic mechanisms that are expected to evolve in response. We propose that these translations emerge when the evolutionary conflicts that influence phenotypic evolution can be appreciated from the perspective of the elements in conflict. The adaptive importance of a given genetic architecture can be understood by asking what genetic limitations on phenotypic change favour each of the conflicting parties. The evolutionary conflicts described in Box 3 exert specific constraints on the evolution of their components, allowing predictions to be made concerning which genetic responses would be favoured by a given evolutionary conflict.

Take for example morphological development, which is guided by a small number of elements that act in many different spatial and temporal contexts within an organism (Carroll 2005). Because of this genetic basis, the evolution of developmental phenotypes is often constrained by pleiotropy because of single-allele, single-individual conflict: the exaggeration of a phenotypic change which is adaptive in one organismal context is limited by effects which are deleterious in another because the conflicting phenotypes share a common causal element. This constraint predicts that morphological evolution should favour the evolution of genetic mechanisms that minimize the scope of developmentally antagonistic pleiotropy. Combinatorial mechanisms of molecular diversification or regulation are evolved mechanisms of overcoming pleiotropic constraints: they allow a small number of genes to influence a large number of developmental outcomes (Carroll 2005). These mechanisms release the spectrum of available phenotypes from restrictions imposed by the discrete

Box 3 Evolutionary conflicts of interest from the perspective of a selfish allele

In his early writings, Richard Dawkins popularized an idea that originated during the modern synthesis and was given shape by biologists studying the evolution of social behaviour (Fisher 1930; Hamilton 1964). 'The Selfish Gene' (or perhaps more precisely, the selfish allele) alludes to conflicts that arise as different allelic variants of a gene 'compete' with one another over evolutionary time to maximize their own rate of replication (Dawkins 1976). Thus, viewed from the perspective of a selfish allele, evolutionary conflicts of interest come in four basic varieties: delineated by the elements whose phenotypic effects place them into conflict (alleles) and by the arena in which these conflicts play out (individuals). We describe each of these four types of conflict and the genetic mechanisms that evolve under their influence (Fig. 4).

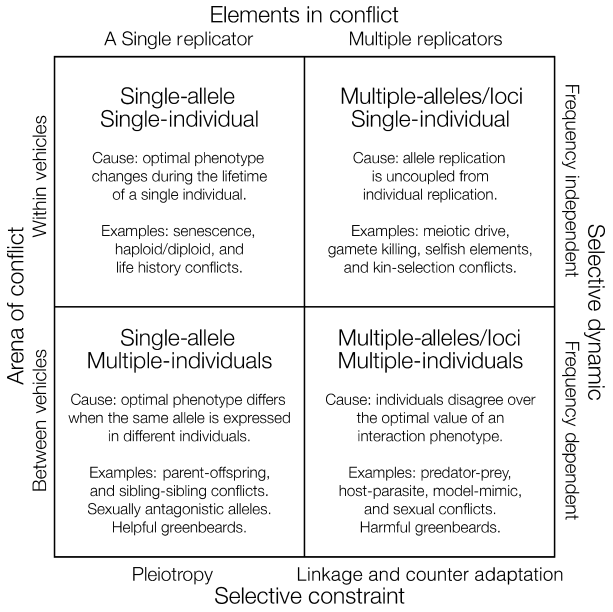

Fig. 4 Evolutionary conflicts of interest from the perspective of a selfish allele.

Single-allele, single-individual conflict

In terms of genetic composition, the simplest evolutionary conflicts occur when an allele produces a phenotype that is not optimal for every circumstance in an individual's life. The result is evolution to a compromise that maximizes the individual's lifetime reproductive success at the cost of a reproductive output that is lower than it would be if each circumstance was independently optimized. A classic example is senescence, which may be caused by an allele whose phenotypic effects on reproductive output early and late in life oppose one another (Williams 1957). Single-allele conflicts within individuals can also occur when an allele's phenotype is not perfectly suited to every environment that an individual will encounter during a lifetime or when phenotypic effects in germ and soma, or during the haploid and diploid phases of a life cycle, are in opposition (Joseph & Kirkpatrick 2004; Choi *et al.* 2008). It may not be possible to eliminate these conflicts entirely, but selection is expected to favour genetic mechanisms that dissociate conflicting phenotypic effects such as regulatory modifications of expression or alternative gene splicing. Aggregated across many genes, single-allele, single-individual conflicts are also a likely source of selection for developmental mechanisms that create spatial and temporal compartmentalization (organelles, tissues or metamorphic life cycles for example) and circumstance-appropriate developmental responses such as phenotypic plasticity.

Single-allele, multiple-individual conflict

Conflicts over optimal phenotype also extend among individuals that share an identical allele but use it in different contexts (parents and offspring, or males and females) or in local competition for a resource (siblings). Here, selec-

Box 3 *(Continued)*

tion favours alleles with phenotypes that maximize inclusive fitness (Hamilton 1964; Grafen 1984). Once again, selection should favour regulatory mechanisms that reduce the scope for conflict as much as is possible, though conflict may persist at evolutionary equilibrium. Single-allele, multiple-individual conflicts are expected to favour conditional expression mechanisms: upregulation of expression in contexts where an allele increases the fitness of the individual that controls the allele's expression and vice versa. For instance, many genes have patterns of regulation that are specific to context: gender, physiological state or social rank. Of particular interest in this class of conflicts are the epigenetic regulatory mechanisms that evolve in response. In many organisms, male and female parents bestow their offspring with epigenetic tags that create patterns of gene expression in the offspring that invoke each parent's optimal resource acquisition strategy (Haig 1996; Wilkins & Haig 2003). Single-allele, multiple-individual contests can also lead to frequency-dependent selection, where the fitness pay-off to an allele with a given phenotype depends on its relative abundance in the population.

All single-allele conflicts, whether they occur in single individuals or among multiple individuals, are ultimately a result of the pleiotropic phenotypic effects of an allele expressed in multiple contexts. The genetic architectures and regulatory mechanisms that evolve to mediate single-allele conflicts should bear evidence of these compromises.

Multiple-alleles/loci, single-individual conflict

Conflicts of interest can also occur when alleles encode phenotypes that uncouple their replication from that of their individual. These are the so-called ultra-selfish genes, whose similarity with selfish individuals in groups led to a rejection of naive group selection thinking in favour of the concept of allele-level selection (Dawkins 1976), and more recently to a refined theory of multilevel selection (Wilson & Wilson 2007). Conflict caused by over-replication of selfish alleles occurs when they create additional copies of themselves within an individual genome (transposons and selfish genetic elements) or over-represent themselves across a set of gametes or zygotes produced by a sexual individual (germ-soma cheating, meiotic drive, gamete killing and gestational drive).

Conflicts within an individual can also occur when alleles alter the allocation of reproductive effort to favour their own replication (sex ratio conflict, sexual conflict in simultaneous hermaphrodites) or recognize and direct benefits towards other individuals that carry copies of an identical allele at the expense of their own individual's replication (altruism). In all of these cases, the selfish allele produces a phenotype which enhances its own rate of replication at a cost to other alleles of the same gene, other genes in the genome or to the individual or group as a whole. This process can favour the evolution of suppressors that counter the selfish allele's effect and to bouts of antagonistic co-evolution between selfish alleles and their suppressor. These conflicts are particularly acute for loci that are physically linked to selfish alleles, and the presence of a selfish allele is often associated with allelic effects that free the selfish allele from linkage (Slotkin & Martienssen 2007) and with genomic architectures that modify the local rate of recombination to bind it to a suppressor (Harada *et al.* 2008).

Multiple-alleles/loci, multiple-individual conflict

The final and most expansive class of evolutionary conflicts is those between different alleles that exert their phenotypic effects independently in different individuals. In these cases, individuals disagree over the optimal value of an interaction phenotype, and alleles evolve to maximize each individual's fitness. These evolutionary conflicts can occur between individuals of the same species (males and females), or different species (hosts and parasites, or predators and prey). Intraspecific conflicts are subject to two additional constraints that do not affect interspecific conflicts: physical linkage between loci encoding the interaction phenotype, and antagonistic pleiotropy. Pleiotropic effects take on an additional importance in a multiple-allele context if they limit the extent to which co-evolution can exaggerate the divergence in one of the conflicting traits and thus shape the direction of evolutionary change.

Many of the conflicts described above can also be manifest as multiple-allele, multiple-individual conflicts. For example: parent–offspring conflict can involve phenotypes encoded by single alleles or multiple alleles/loci. Intracellular parasites can benefit themselves by biasing a host's reproductive allocation just as the host's own selfish alleles sometimes do. The evolutionary dynamics that result from multiple-allele, multiple-individual conflicts can be very complex, involving frequency-dependent selection, modifications of linkage and antagonistic co-evolution.

Box 3 *(Continued)*

These influences can lead the genetic constituents to evolve towards stable equilibria, limit cycles, diversification of one or both partners, or chronic co-evolution driven by successive counter-adaptations (Dawkins & Krebs 1979; Hayashi *et al.* 2007). In instances where phenotypes are encoded by a small number of loci, dominance and epistasis can exert profound effects on the direction of evolutionary change (Hayashi *et al.* 2007).

genetic causes of phenotypic variation. Their existence may thus help explain why the phenotypic gambit is so often successful in describing evolution, despite not explicitly considering its genetic causes.

For example, variation in morphological phenotypes is often caused by cis-regulatory changes, perhaps because these mechanisms can exert their effect on developmental processes independently (Carroll 2005; Stern & Orgogozo 2008). For the same reason, we might predict that alternative splicing of a protein-coding region or gene duplication would be favoured when the same protein is expressed in multiple contexts, or required to function in multiple tasks within an organism (Gilad *et al.* 2003; Hattori *et al.* 2009). In a multiple-individual context, the pleiotropic constraints imposed by single-allele conflicts (such as alleles selected for function in both parents and offspring, or in both males and females) often result in context-dependent or epigenetic regulatory mechanisms, because these alter the expression of the allele in different individuals and thus separate its conflicting pleiotropic effects (Wilkins & Haig 2003; Innocenti & Morrow 2010).

Pleiotropic constraints can also influence evolutionary conflicts between alleles or loci, but multiple-allele conflicts are often more overtly associated with modifications of physical linkage and rapid evolution because of bouts of counter-adaptation. Alleles that replicate selfishly within individuals are in conflict with elements in physically linked regions of the genome (Hurst *et al.* 1996). This linkage can favour genetic features that modify the local genomic recombination rate (Haig & Grafen 1991) or gene architectures that physically prevent recombination such as the gene-within-a-gene structure of the ciona sperm–egg self-incompatibility system (Harada *et al.* 2008). Selfishly replicating elements often display the counter-strategy; for example, they have genetic mechanisms of copying themselves to other locations in the genome, perhaps to avoid being physically bound to genes that suppress their replication (Kazazian 2004).

Chronic rapid evolutionary change is another recognized outcome of evolutionary conflict between alleles or loci. Evolutionary contests between proteins (for example, those involved in sperm–egg or host–parasite conflict) can drive rates of molecular evolution well in excess of the neutral expectation (Hughes 2002; Swanson & Vacquier 2002). When selection continually favours diversification, gene architectures that generate remarkable allelic diversity can evolve. The oyster sperm gene bindin generates thousands of alleles within a single species by a combination of alternative splicing, recombination and positive selection occurring on a gene whose primary structure is a series of repeats of a single functional motif (Moy *et al.* 2008). Similar genetic mechanisms create polymorphism in recognition systems that underlie reproductive self-incompatibility (Wang *et al.* 2001), kin-discrimination (Gibbs *et al.* 2008) and host-immune response (Litman *et al.* 2007).

Evolved genetic architectures can cause violations of the phenotypic gambit by limiting or expanding the potential for diversification, creating new equilibria that may not be evolutionarily stable in unconstrained systems. The reverse is also true, and in that evolved genetic architectures and regulatory mechanisms can release phenotypic variation from pleiotropic constraints imposed by simple genetic systems (Box 3). These kinds of adaptive genetic responses have become *prima facie* evidence of their respective evolutionary conflicts and provide a set of predictions for assigning fitness pay-offs to genetic variants, and for recognizing potential genetic responses to evolutionary conflict when presented with genetic mechanisms of unusual form and unknown function.

Conclusion

Evolutionary studies of behaviour often focus on neurally based decision-making, but conflict and collaboration in biological systems existed long before brains evolved, and many behaviours must therefore be mediated by direct physical interactions with other individuals or with the environment. These traits negotiate interesting and important biological interactions and offer a wealth of opportunity for evolutionary biology. Their genotype–phenotype relationship is mechanistically simple and their genetic basis easier to uncover than that of most quantitative traits; moreover, the current and historical effects of selection can be rigorously determined using appropriate models of molecular evolution, and the spectrum of phenotypic effects created by amino acid and regulatory mutation of candidate genes can be evaluated experimentally. Most importantly, the find-

ings of these manipulations can be compared to bouts of natural adaptation, or evolution to equilibria in nature.

The tractability of molecular behaviours offers the opportunity to study a difficult, but fundamentally important evolutionary question: how do the genetic and developmental mechanisms that underlie phenotypic variation influence adaptation and evolution? Early evidence suggests that in traits with a simple genetic basis, genetic mechanisms can exert a profound influence on phenotypic response to selection and genetic architectures that influence the scope of evolutionary conflict can and do evolve. Evolution in these circumstances may be impossible to describe with the phenotypic gambit, and such traits are therefore worth the effort required to study them from a genetic perspective.

Acknowledgements

This work was supported by: National Institute of Health (HD42563) and National Science Foundation (DEB-0716761) grants to WJS, and Natural Sciences and Engineering Research Council of Canada (NSERC) grants to BJC. SAS is a recipient of an NSERC PGSD International Graduate Fellowship. Thanks to the SFU FAB*Lab, the UW Swanson Lab, the UW Evol-Syst group, JB Joy, CE Parent, P Nosil, JG Tyerman, R Vos, G Findlay, J Aagaard and P Gagneux for many engaging discussions. Special thanks to HD Bradshaw, F Breden, DD Heath, AO Mooers, CL Peichel, BJ Swalla and VD Vacquier for their immeasurably valuable contributions to these ideas as they developed.

References

Akey JM (2009) Constructing genomic maps of positive selection in humans: where do we go from here? *Genome Research*, **19**, 711–722.

Andolfatto P (2005) Adaptive evolution of non-coding DNA in *Drosophila*. *Nature*, **437**, 1149–1152.

Arnold SJ (1992) Constraints on phenotypic evolution. *The American Naturalist*, **140**(Suppl. 1), S85–S107.

Bauer DuMont VL, Flores HA, Wright MH, Aquadro CF (2007) Recurrent positive selection at bgcn, a key determinant of germ line differentiation, does not appear to be driven by simple coevolution with its partner protein bam. *Molecular Biology and Evolution*, **24**, 182–191.

Benabentos R, Hirose S, Sucgang R *et al.* (2009) Polymorphic members of the lag gene family mediate kin discrimination in *Dictyostelium*. *Current Biology*, **19**, 567–572.

Bishop JG (2005) Directed mutagenesis confirms the functional importance of positively selected sites in polygalacturonase inhibitor protein. *Molecular Biology and Evolution*, **22**, 1531–1534.

Bradshaw HD, Schemske DW (2003) Allele substitution at a flower colour locus produces a pollinator shift in monkeyflowers. *Nature*, **426**, 176–178.

Bridgham JT, Ortlund EA, Thornton JW (2009) An epistatic ratchet constrains the direction of glucocorticoid receptor evolution. *Nature*, **461**, 515–519.

Buck MJ, Lieb JD (2004) ChIP-chip: considerations for the design, analysis, and application of genome-wide chromatin immunoprecipitation experiments. *Genomics*, **83**, 349–360.

Carroll SB (2005) Evolution at two levels: on genes and form. *PLoS Biology*, **3**, e245.

Choi S-K, Yoon S-R, Calabrese P, Arnheim N (2008) A germline-selective advantage rather than an increased mutation rate can explain some unexpectedly common human disease mutations. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 10143–10148.

Chouard T (2010) Revenge of the hopeful monster. *Nature*, **463**, 864–867.

Clark NL, Gasper J, Sekino M *et al.* (2009) Coevolution of interacting fertilization proteins. *PLoS Genetics*, **5**, e1000570.

Colosimo PF, Hosemann KE, Balabhadra S *et al.* (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science*, **307**, 1928–1933.

Crespi B (2001) The evolution of social behavior in microorganisms. *Trends in Ecology and Evolution*, **16**, 178–183.

Dawkins R (1976) *The Selfish Gene*. Oxford University Press, Oxford.

Dawkins R, Krebs JR (1979) Arms races between and within species. *Proceedings of the Royal Society of London, Series B: Biological Science*, **205**, 489–511.

Eckmann CR, Crittenden SL, Suh N, Kimble J (2004) GLD-3 and control of the mitosis/meiosis decision in the germline of *Caenorhabditis elegans*. *Genetics*, **168**, 147–160.

Elde NC, Child SJ, Geballe AP, Malik HS (2009) Protein kinase R reveals an evolutionary model for defeating viral mimicry. *Nature*, **457**, 485–489.

Felsenstein J (1985) Phylogenies and the comparative method. *The American Naturalist*, **125**, 1–15.

Fiegna F, Velicer GJ (2005) Exploitative and hierarchical antagonism in a cooperative bacterium. *PLoS Biology*, **3**, e370.

Field SF, Matz MV (2010) Retracing evolution of red fluorescence in GFP-Like proteins from Faviina corals. *Molecular Biology and Evolution*, **27**, 225–233.

Findlay GD, Yi X, Maccoss MJ, Swanson WJ (2008) Proteomics reveals novel *Drosophila* seminal fluid proteins transferred at mating. *PLoS Biology*, **6**, e178.

Fisher RA (1930) *The Genetical theory of Natural Selection*, Clarendon Press, Oxford.

Gardner A, West SA (2010) Greenbeards. *Evolution*, **64**, 25–38.

Geyer LB, Palumbi SR (2003) Reproductive character displacement and the genetics of gamete recognition in tropical sea urchins. *Evolution*, **57**, 1049–1060.

Gibbs KA, Urbanowski ML, Greenberg EP (2008) Genetic determinants of self identity and social recognition in bacteria. *Science*, **321**, 256–259.

Gilad Y, Bustamante CD, Lancet D, Pääbo S (2003) Natural selection on the olfactory receptor gene family in humans and chimpanzees. *American Journal of Human Genetics*, **73**, 489–501.

Glazier AM, Nadeau JH, Aitman TJ (2002) Finding genes that underlie complex traits. *Science*, **298**, 2345–2349.

Grafen A (1984) Natural selection, kin selection and group selection. In: *Behavioural Ecology: An Evolutionary Approach* (eds Krebs JR, Davies NB). pp. 62–84, Blackwell Scientific Publications, Oxford.

Haig D (1996) Gestational drive and the green-bearded placenta. *Proceedings of the National Academy of Sciences of the United States of America*, **93**, 6547–6551.

Haig D, Grafen A (1991) Genetic scrambling as a defence against meiotic drive. *Journal of Theoretical Biology*, **153**, 531–558.

Haldane J (1933) The part played by recurrent mutation in evolution. *The American Naturalist*, **67**, 5–19.

Hamilton WD (1964) The genetical evolution of social behaviour. I. *Journal of Theoretical Biology*, **7**, 1–16.

Hamilton WD, Axelrod R, Tanese R (1990) Sexual reproduction as an adaptation to resist parasites (a review). *Proceedings of the National Academy of Sciences of the United States of America*, **87**, 3566–3573.

Harada Y, Takagaki Y, Sunagawa M *et al.* (2008) Mechanism of self-sterility in a hermaphroditic chordate. *Science*, **320**, 548–550.

Hattori D, Chen Y, Matthews BJ *et al.* (2009) Robust discrimination between self and non-self neurites requires thousands of Dscam1 isoforms. *Nature*, **461**, 644–648.

Hayashi TI, Vose M, Gavrilets S (2007) Genetic differentiation by sexual conflict. *Evolution*, **61**, 516–529.

Hoekstra HE, Hirschmann RJ, Bundey RA, Insel PA, Crossland JP (2006) A single amino acid mutation contributes to adaptive beach mouse color pattern. *Science*, **313**, 101–104.

Hughes AL (2002) Natural selection and the diversification of vertebrate immune effectors. *Immunological Reviews*, **190**, 161–168.

Hurst LD, Atlan A, Bengtsson BO (1996) Genetic conflicts. *Quarterly Review of Biology*, **71**, 317–364.

Innocenti P, Morrow EH (2010) The sexually antagonistic genes of *Drosophila melanogaster*. *PLoS Biology*, **8**, e1000335.

Jiggins FM, Hurst GDD, Yang Z (2002) Host–symbiont conflicts: positive selection on an outer membrane protein of parasitic but not mutualistic Rickettsiaceae. *Molecular Biology and Evolution*, **19**, 1341–1349.

Joseph S, Kirkpatrick M (2004) Haploid selection in animals. *Trends in Ecology and Evolution*, **19**, 592–597.

Kazazian HH (2004) Mobile elements: drivers of genome evolution. *Science*, **303**, 1626–1632.

Koufopanou V, Bell G (1991) Developmental mutants of Volvox: does mutation recreate the patterns of phylogenetic diversity? *Evolution*, **45**, 1806–1822.

Levitan DR, Ferrell DL (2006) Selection on gamete recognition proteins depends on sex, density, and genotype frequency. *Science*, **312**, 267–269.

Lewontin RC (1978) Adaptation. *Scientific American*, **239**, 156–169.

Litman GW, Dishaw LJ, Cannon JP, Haire RN, Rast JP (2007) Alternative mechanisms of immune receptor diversity. *Current Opinion in Immunology*, **19**, 526–534.

Marsh M, Helenius A (2006) Virus entry: open sesame. *Cell*, **124**, 729–740.

Matute DR (2010) Reinforcement of gametic isolation in *Drosophila*. *PLoS Biology*, **8**, e1000341.

Maynard Smith J (1978) Optimization theory in evolution. *Annual Review of Ecology and Systematics*, **9**, 31–56.

McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.

Mercer J, Helenius A (2008) Vaccinia virus uses macropinocytosis and apoptotic mimicry to enter host cells. *Science*, **320**, 531–535.

Moore T, Haig D (1991) Genomic imprinting in mammalian development: a parental tug-of-war. *Trends in Genetics*, **7**, 45–49.

Moore T, Moore H (2002) Marsupial sperm pairing: a case of 'sticky' green beards? *Trends in Ecology and Evolution*, **17**, 112–113.

Moy GW, Springer SA, Adams SL, Swanson WJ, Vacquier VD (2008) Extraordinary intraspecific diversity in oyster sperm bindin. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 1993–1998.

Nielsen R (2009) Adaptionism-30 years after Gould and Lewontin. *Evolution*, **63**, 2487–2490.

Nielsen R, Bustamante C, Clark AG *et al.* (2005) A scan for positively selected genes in the genomes of humans and chimpanzees. *PLoS Biology*, **3**, e170.

Nielsen R, Hellmann I, Hubisz M, Bustamante C, Clark AG (2007) Recent and ongoing selection in the human genome. *Nature Reviews Genetics*, **8**, 857–868.

Olson MV (1999) When less is more: gene loss as an engine of evolutionary change. *American Journal of Human Genetics*, **64**, 18–23.

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

Palumbi SR (1999) All males are not created equal: fertility differences depend on gamete recognition polymorphisms in sea urchins. *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 12632–12637.

Parker GA, Maynard Smith J (1990) Optimality theory in evolutionary biology. *Nature*, **348**, 27–33.

Paterson S, Vogwill T, Buckling A *et al.* (2010) Antagonistic coevolution accelerates molecular evolution. *Nature*, **464**, 275–278.

Presgraves DC (2010) The molecular evolutionary basis of species formation. *Nature Reviews Genetics*, **11**, 175–180, doi:10.1038/nrg2718.

Presgraves DC, Balagopalan L, Abmayr SM, Orr HA (2003) Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature*, **423**, 715–719.

Schwikowski B, Uetz P, Fields S (2000) A network of protein–protein interactions in yeast. *Nature Biotechnology*, **18**, 1257–1261.

Shapiro MD, Marks ME, Peichel CL *et al.* (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature*, **428**, 717–723.

Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nature Reviews Genetics*, **8**, 272–285.

Smukalla S, Caldara M, Pochet N *et al.* (2008) FLO1 is a variable green beard gene that drives biofilm-like cooperation in budding yeast. *Cell*, **135**, 726–737.

Springer SA, Crespi BJ (2007) Adaptive gamete-recognition divergence in a hybridizing *Mytilus* population. *Evolution*, **61**, 772–783.

Stern DL, Orgogozo V (2008) The loci of evolution: how predictable is genetic evolution? *Evolution*, **62**, 2155–2177.

Swanson WJ, Vacquier VD (1997) The abalone egg vitelline envelope receptor for sperm lysin is a giant multivalent molecule. *Proceedings of the National Academy of Sciences of the United States of America*, **94**, 6724–6729.

Swanson WJ, Vacquier VD (2002) The rapid evolution of reproductive proteins. *Nature Reviews Genetics*, **3**, 137–144.

Tajima F (1989) Statistical method for testing the neutral mutation hypothesis by DNA polymorphism. *Genetics*, **123**, 585–595.

Tang K, Thornton KR, Stoneking M (2007) A new approach for using genome scans to detect recent positive selection in the human genome. *PLoS Biology*, **5**, e171.

Teeter KC, Payseur BA, Harris LW *et al.* (2008) Genome-wide patterns of gene flow across a house mouse hybrid zone. *Genome Research*, **18**, 67–76.

Wang X, Hughes AL, Tsukamoto T, Ando T, Kao T (2001) Evidence that intragenic recombination contributes to allelic diversity of the S-RNase gene at the self-incompatibility (S) locus in *Petunia inflata*. *Plant Physiology*, **125**, 1012–1022.

Watson FL, Püttmann-Holgado R, Thomas F *et al.* (2005) Extensive diversity of Ig-superfamily proteins in the immune system of insects. *Science*, **309**, 1874–1878.

Weinreich DM, Delaney NF, Depristo MA, Hartl DL (2006) Darwinian evolution can follow only very few mutational paths to fitter proteins. *Science*, **312**, 111–114.

Wilkins JF, Haig D (2003) What good is genomic imprinting: the function of parent-specific gene expression. *Nature Reviews Genetics*, **4**, 359–368.

Williams G (1957) Pleiotropy, natural selection, and the evolution of senescence. *Evolution*, **11**, 398–411.

Wilson DS, Wilson EO (2007) Rethinking the theoretical foundation of sociobiology. *Quarterly Review of Biology*, **82**, 327–348.

Yang Z, Bielawski JP (2000) Statistical methods for detecting molecular adaptation. *Trends in Ecology and Evolution*, **15**, 496–503.

Yang Z, Swanson WJ, Vacquier VD (2000) Maximum-likelihood analysis of molecular adaptation in abalone sperm lysin reveals variable selective pressures among lineages and sites. *Molecular Biology and Evolution*, **17**, 1446–1455.

Yokoyama S, Zhang H, Radlwimmer FB, Blow NS (1999) Adaptive evolution of color vision of the Comoran coelacanth (*Latimeria chalumnae*). *Proceedings of the National Academy of Sciences of the United States of America*, **96**, 6279–6284.

Yokoyama S, Tada T, Zhang H, Britt L (2008) Elucidation of phenotypic adaptations: molecular analyses of dim-light vision proteins in vertebrates. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 13480–13485.

Zhang J (2006) Parallel adaptive origins of digestive RNases in Asian and African leaf monkeys. *Nature Genetics*, **38**, 819–823.

Zuckerkandl E, Pauling L (1965) Molecules as documents of evolutionary history. *Journal of Theoretical Biology*, **8**, 357–366.

S.A.S. studies the evolution of simple behavioural systems to explore what the phenotypic effects of molecular divergence can tell us about the selective forces that cause adaptive evolution. This work is a result of Masters research with B.J.C. and the FAB*Lab at Simon Fraser University, and PhD research with W.J.S. and the Department of Biology at the University of Washington. Stevan's thesis projects examined the effect of adaptive gamete recognition protein divergence on sperm-egg compatibility and the evolution of reproductive isolation between naturally hybridizing blue mussel species.
