REVIEW PAPER

The potential of a population genomics approach to analyse geographic mosaics of plant–insect coevolution

Kim M. C. A. Vermeer · Marcel Dicke · Peter W. de Jong

Abstract   A central issue in the evolutionary ecology of species interactions is coevolution, which involves the reciprocal selection between individuals of interacting species. Understanding the importance of coevolution in shaping species interactions requires the consideration of spatial variation in their strength. This is exactly what the, recently developed, geographic mosaic theory of coevolution addresses. Another major development in the study of population ecology is the introduction of the population genomics approach in this field of research. This approach addresses spatial processes through molecular methods. It is of particular interest that population genomics is especially applicable to natural populations of non-model species. We describe how population genomics can be used in the context of the geographic mosaic of coevolution, specifically to identify coevolutionary hot-spots, and to attribute genetic variation found at specific loci to processes of selection versus trait remixing. The proposed integration of the population genomics approach with the conceptual framework of the geographic mosaic of coevolution is illustrated with a few selected, particularly demonstrative, examples from the realm of insect–plant interactions.

Keywords   Geographic mosaic of coevolution · Natural resistance · Plant–insect interaction · Population genomics

Introduction

Organisms are under the constant threat of attack by their enemies and have evolved a range of defence strategies. How organisms evolve in the context of attack-defence relationships has been under intensive debate for several decades (e.g. Courtney 1988; Fox 1988; Rausher 1988; Thompson 1988). Since 1994 John Thompson has developed his influential theory "The Geographic Mosaic Theory of Coevolution" (Thompson 1994,

K. M. C. A. Vermeer (✉) · M. Dicke · P. W. de Jong
Laboratory of Entomology, Wageningen University, PO Box 8031, 6700 EH Wageningen,
The Netherlands
e-mail: kim.vermeer@wur.nl

1999a, 2005). The primary premise of this conceptual framework to explain coevolutionary interactions and –patterns is that populations are generally genetically and ecologically structured and that, therefore, coevolution between interacting species must be studied at different scales, including that of local populations. Coevolution *sensu stricto* is thought to occur only in a limited range of the geographical distributions of the species involved, called "coevolutionary hot spots". Beyond such local coevolution, larger-scale coevolutionary dynamics are additionally influenced by geographic selection mosaics and trait remixing (here defined as the changing of the spatial distribution of alleles by processes including gene flow across landscapes, random genetic drift within populations, extinction and recolonisation of local populations, and mutation; Thompson 2005; Gomulkiewicz et al. 2007). This view on coevolution has led to clear, testable predictions, and indeed these are taken up by an expanding range of scientists, who experimentally test these predictions in a variety of systems (e.g. Thompson 1999a, b, 2009a; Burdon and Thrall 1999; Lively 1999; Brodie et al. 2002; Benkman et al. 2003; Neuhauser et al. 2003; Zangerl and Berenbaum 2003; Toju and Sota 2006). One of the central challenges is to explain the observed geographical distribution of alleles at adaptive loci in terms of selection (stabilizing, directional, disruptive or balancing) *versus* migration and genetic drift. In the context of the geographic mosaic of coevolution, this challenge translates into the identification of coevolutionary hot spots, and to distinguish their effects on genetic variation from those of trait remixing (Thompson 2005; Gomulkiewicz et al. 2007).

At roughly the same time as the development of these novel concepts in research on coevolution, a new approach was introduced in the field of evolutionary ecology and evolutionary genetics: population genomics (Black et al. 2001; Schlotterer 2002; Luikart et al. 2003; Dicke et al. 2004). This approach was inspired by two key factors: (1) the realization that, generally, populations are genetically structured, and consist of local sub-populations that are linked by migration, and (2) the progress in the development and analysis of molecular markers, that have become widely available for non-model organisms. Essentially, population genomics is an approach that attributes the distribution of genetic variation to locus-specific effects, such as selection, *versus* genome-wide processes such as migration, by genome-wide sampling of molecular markers. A recent development is the connection of population genomics to the field of phylogeography (Brito and Edwards 2009; Hickerson et al. 2010), which has traditionally provided a neutral template for the study of the evolution of coevolving traits, and aims to describe population history and estimate demographic parameters.

The approach of population genomics shows considerable congruence with the conceptual model of a geographic mosaic of coevolution. After introducing the two key topics of this paper in more detail, we will explore how the population genomics approach can lead to more insights in the dynamics of the geographic mosaics of coevolution.

The evolution of the concept of coevolution

Coevolution, i.e. reciprocal selection between individuals of interacting species, is an important process influencing adaptations. Classic examples of coevolution are found among the interactions between insects and plants (Ehrlich and Raven 1964; Berenbaum and Zangerl 1992; Funk et al. 2002; Cornell and Hawkins 2003; Pellmyr 2003; Nuismer and Thompson 2006; Futuyma 2009). This is not surprising, because insect herbivores represent more than 25% of all multi-cellular species (Strong et al. 1984) and our green

world is abundantly covered by plants, making insect–plant interactions one of the most common interactions within ecosystems on this planet (Schoonhoven 2005).

Since the concept of coevolution was first conceived (Darwin 1859, 1862), there have been different views on how to define it. At one extreme, only reciprocal one-to-one interactions at the species level are included (Strong et al. 1984), also called pairwise coevolution by others such as Janzen (Janzen 1980). At the other extreme, coevolution is regarded as a process that is diffuse and multispecific (all species may interact with several other species; this is also known as diffuse coevolution) (Janzen 1980; Fox 1981; Futuyma and Slatkin 1983). As a result of using a very strict definition of coevolution, several authors such as Jermy (1984); Strong et al. (1984) and Schmitt et al. (1995) do not agree with the idea that co-evolution is responsible for, or has at least a major impact on, insect–plant interactions.

The two extreme views on coevolution mentioned above, i.e. pairwise coevolution and diffuse coevolution, usually ignore the spatial structure of species interactions (Thompson 1999b). As a result, much of the coevolutionary process that occurs in between these two extremes is ignored. The spatial structure of species interactions is explicitly included in the more recently developed theory of the geographic mosaic of coevolution (Thompson 1994, 2005; Thompson et al. 1997). In this theory, much of the coevolutionary process occurs at levels in between those of local populations and species. This relates to the structure that many populations exhibit. Populations are generally not homogeneous, but geographically and genetically structured (Wright 1951, 1968; Avise 2000; Thompson 2005), due to, for example, the patchy nature of their habitats and the patchy nature of biotic and abiotic environmental factors (Agrawal et al. 2001). One feature leading to structured populations is the lack of complete dispersal by individuals within a population so that mating within the population is not random. A metapopulation (Hanski 1998) often consists of subpopulations, that each have their own population dynamics and a certain degree of mating restriction between them.

Apart from the spatial structure of these metapopulations, that contributes to a geographic mosaic of coevolution, the fact that geographical ranges of interacting species usually do not completely overlap leads to spatial variation in species interactions and in the intensity of natural selection. For example, in plant–insect interactions, some patches of host plants may not be within the distributional range of a particular herbivore, while other patches might be under attack by this natural enemy, resulting in a difference in selection intensity on the host plant among patches. The distributional range of an insect species can be smaller than that of the host plant (Strong et al. 1984) or even exceed the host plant's range, for example when an insect uses multiple host plants. Both cases result in an unequal distribution of selective interactions across the distributional range. Also the geographical range of other herbivore species may have an effect on the selective interactions in a metapopulation of a species under study.

In most of north-western North America the moth *Greya politella* uses woodland stars (*Lithophragma parviflorum*) as its only host plant and acts as its pollinator when ovipositing in the flower (Thompson and Pellmyr 1992). In this case the costs of seed feeding by the moth's offspring are smaller than the benefit resulting from pollination. When other effective co-pollinators are involved, however, the pollination of *L. parviflorum* no longer depends on *G. politella*; the mutualistic effects of *G. politella* on *L. parviflorum* (Thompson and Pellmyr 1992) are swamped. Thus, other pollinators present in some subpopulations interfere with the selective interaction between *G. politella* and *L. parviflorum*.

Lodgepole pine (*Pinus contorta* ssp. *latifolia*) forms another example of a host plant experiencing differences in the interaction with different herbivores. The morphology of lodgepole pine cones in and near the Rocky Mountains has evolved differently because of different selection pressures imposed by red squirrels (*Tamiasciurus hudsonicus*) and red crossbills (*Loxia curvirostra*) (Benkman 1999; Benkman et al. 2001, 2003). Both species eat the seeds from the partially closed pine cones. When no red squirrels are present, the red crossbill imposes a selection on the cone shape (cones are relatively narrow at their base) (Benkman 1999), but when red squirrels are present, the squirrels harvest most cones before crossbills have access to them. The presence of red squirrels imposes such a strong selection on the cones (cones are short and wide at the base when squirrels are present) that red crossbills no longer influence cone morphology (Benkman et al. 2001). Therefore, selection on cone shape is imposed by red crossbills only when red squirrels are absent. So, within the lodgepole pine species range, different reciprocal selective forces act due to differences in the distributional range of the herbivores.

Environmental variation (spatial and temporal) further contributes to this mosaic of coevolution. Individuals in subpopulations can evolve traits that make them best adapted to the local environmental conditions. However, because environmental conditions vary in space and time, an evolved trait may lose its adaptive value when local conditions change. Similarly, migration between patches can cause a locally adapted genotype to spread to localities where it does not perform as well as in the original patch. In an extreme case, the adaptation might even become a maladaptation in the new patch. So, a variable environment and migration between patches may cause local mismatches of traits and maladaptation (Thompson 1999b; Schoonhoven et al. 2005).

The importance of implementing a spatial component in the theory of coevolution

Ignoring the geographical structure of populations might lead to an underestimation of the importance of coevolution. Coevolution between plants and insects is likely to be diffuse, because a herbivorous enemy of a plant species may alter the pattern of selection exerted by other natural enemies (Rausher 1996). Rausher also states that there can only be either pairwise or diffuse coevolution. Following Rausher's idea, if an interaction that evolves between a plant and its enemy is not influenced by the absence or presence of another species, this is called pairwise coevolution. All other evolutionary processes are diffuse (Iwao and Rausher 1997). By ignoring the geographical structure of populations, however, one ignores the possibility of the existence of both pairwise *and* diffuse coevolution. Where one subpopulation of the plant might be attacked by several enemies, plants in another patch can still be under selection exerted by only one enemy, thus being involved in pairwise coevolution at the local scale.

Strong et al. (1984) use geographical structure precisely as an argument why coevolution would not be common. Plant–insect interactions are variable and unpredictable in time and space and, therefore, Strong et al. (1984) argue that selection pressures differ locally, making it less likely that coevolution (which needs reciprocal and intense interactions) is actually working; in other words, they argue that the selection pressure is not stable and intense enough to drive coevolution. During some years a patch may consist of two interacting species, while in another year one of them can be absent or even more species may be present that interfere with this interaction. A reciprocal interaction would, thus, not be sustained and intense and, therefore, in the view of Strong et al. (1984) coevolution would not take place.

The geographical structure of plant–insect interactions, i.e. Strong et al's argument why coevolution is not common, is exactly what makes coevolution so flexible and such a major force in shaping ecologically relevant traits (Thompson 2009b). This is precisely what provides the raw material of the geographic mosaic of coevolution. As a certain local adaptation occurs in some places but not in others, a geographical structure arises of different selective interactions, the so-called selection mosaics. Given the overwhelming evidence that indeed populations of most, if not all, species are geographically structured even at the local scale (e.g. Bermingham and Avise 1986; Bohonak 1998; Althoff and Thompson 1999; Medrano and Herrera 2008; Nosil et al. 2008; Gomez et al. 2009a), Thompson's view of a dynamic mosaic of ever changing coevolutionary interactions, governed by (sometimes rapid) local adaptation and trait remixing through migration, in our view best describes the existing pattern of species interactions (Thompson 2005).

Geographic mosaic theory of coevolution (GMTC)

Considering the process of coevolution as a hierarchical process operating at different spatial and temporal scales, the geographic mosaic theory of coevolution consists of three components: (a) geographic selection mosaics, (b) trait remixing and (c) coevolutionary cold and hot spots (Thompson 2005). "*Geographic selection mosaics*" refers to the occurrence of geographic differences in fitness of interacting species. Selection mosaics are not just variable selective forces on interactions, but relate to spatial differences in genotype-by-genotype-by-environment interactions amongst interacting species (Thompson 2005). "*Trait remixing*" refers to the changing of the distribution of alleles by processes such as gene flow and genetic drift. Gene flow between local subpopulations and genetic drift within the subpopulations may alter the distribution of traits by changing the genetic composition of the subpopulation, which can be further modified by local extinction of subpopulations and mutation. *Coevolutionary hot spots* are local communities where reciprocal selection through mutual interactions between individuals of different species occurs. Such reciprocal selection does not occur in *cold spots*. These three components are considered to make up the raw material of the GMTC, i.e. they constitute the driving force behind coevolution.

Based on these three components of the geographic mosaic of coevolution, different patterns and dynamics of interspecific interactions are expected than when considering a population that is not geographically structured. The patterns expected under the GMTC are trait mismatching among interacting species, few species-level coevolved traits and spatial variation in traits mediating interactions among species (Thompson 2005). Trait mismatching, for example, can be caused by one of the forces that contribute to trait remixing, e.g. gene flow between hot spots and cold spots. Because local communities differ, a certain trait can be an advantage in some hot spots whereas in others it is not, or may even cause a local maladaptation. Gene flow can prevent local coevolution of traits in some local communities, thereby causing local trait mismatching and maladaptation. Few species-wide coevolved traits are expected in a landscape of geographically and genetically structured subpopulations, because only in some cases have coevolved traits become fixed in the complete metapopulation (Thompson 2005).

Verifying these ecological predictions is valuable, but not enough to test the GMTC per se (Gomulkiewicz et al. 2007; Thompson 2005, 2009a). Spatially variable traits can, for example, also occur in antagonistic interactions, even when none of the three components of the GMTC are present. Local maladaptation can also take place without geographic

selection mosaics, hot spots, cold spots and trait remixing. The same holds for the expected low number of coevolved traits found at the species level; they can also be observed when a geographic mosaic of interactions is not present. Even when gene flow is completely absent, populations can show patterns such as local maladaptation and trait mismatching (Nuismer et al. 2003). So, when the three patterns are all present, this does not necessarily mean that there is a geographical mosaic of coevolution.

It is important to realize that in addition to demonstrating the occurrence of the three *patterns*, i.e. local maladaptation, spatially variable traits and a paucity of coevolved traits at the species level, unraveling the components of the underlying *processes* that generate these patterns is vital to support the GMTC. The processes will reveal whether the patterns are indeed caused by the components of the GMTC. The components are the three components mentioned earlier: selection mosaics, cold and hot spots, and trait remixing. Since 2005, rapid progress has been made in formulating and testing of predictions from the GMTC (Thompson 2009a; Laine 2009). Approaches to test the underlying components are described by Gomulkiewicz et al. (2007).

The processes that generate the patterns predicted by the geographical mosaic theory of coevolution in the first place act on genetic variation. For a genotype to have effect on another genotype of an interacting species, i.e. the very basis of coevolution, there must be enough genetic variation in both species. There is a surprising paucity in empirical evidence of the genetics underlying natural adaptations, while the underlying genetic basis is a fundamental part of the understanding of natural adaptation (Orr 2005; Stinchcombe and Hoekstra 2008). Thompson emphasizes the need to study the *genetics* underlying traits that have evolved through coevolution and the processes influencing the spread and distribution of these genetic factors (Thompson 1999b, 2005). Hence, such knowledge is vital for understanding the evolutionary processes that shape the geographical mosaic of coevolution. To study genetic variation and how this influences the predictions of the GMTC, means that one cannot simply use a single disciplinary field, but must use a multidisciplinary approach, including ecology, genetics and molecular biology. A promising way to investigate the geographic mosaic theory of coevolution is to exploit a population genomics approach. Population genomics provides an interface between population genetics and molecular biology (Black et al. 2001), which links the molecular genetic basis of (coevolutionary) forces to their consequences at the population level.

Population genomics

Since Black et al. (2001) introduced the population genomics approach in the field of plant–insect interactions (Black et al. 2001), it has proven to be useful in a range of studies (e.g. Rogers and Bernatchez 2005; Egan et al. 2008; Herrera and Bazaga 2008; Minder and Widmer 2008; Nosil et al. 2008; Schneider 2008; Butlin 2010; Ikeda and Setoguchi 2010). Briefly, the population genomics approach as developed by Black et al. (2001), involves the following. By sampling numerous markers throughout the genome in individuals from one or more populations, parts of the genome can be identified that differ in variation across populations from that of neutral markers, due to locus-specific effects such as selection (Black et al. 2001). Genome-wide effects, that affect all markers in the genome, include genetic drift (founder effects and population bottlenecks), migration and inbreeding (Black et al. 2001). What is done in the population genomics approach is that genome-wide effects and locus-specific effects are differentiated to gain more insight in
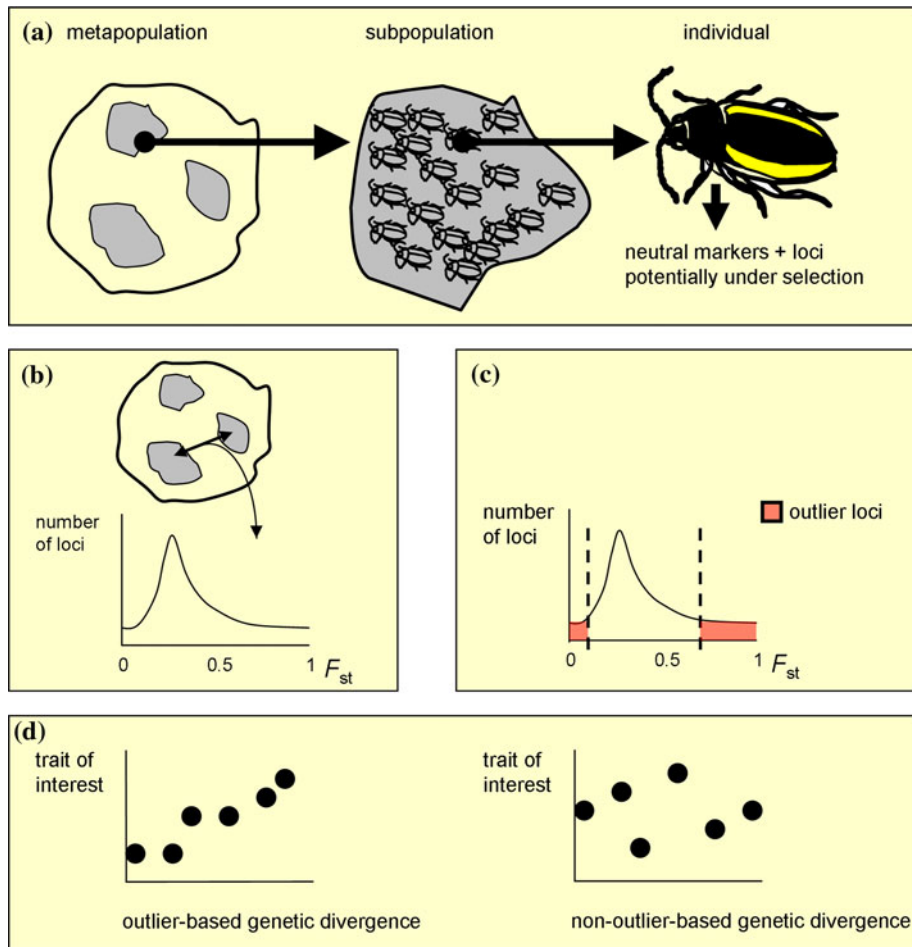
Fig. 1 The population genomics approach consists of four steps. **a** Sampling numerous (*neutral*) molecular markers—such as microsatellites—as well as testing loci that are suspected to be under selection in individuals of different tentative subpopulations. **b** A frequency distribution is drawn of a measure of the variation between different samples, in this example the Fst value for each locus. **c** Screen the frequency distribution for statistical outliers. **d** Validation step to verify whether the outlier is caused by selection (or by hitch-hiking with/on a genome region that is under selection). In this example, non-outlier- and outlier-based genetic distances are correlated with differences in the ecological trait of interest

evolutionary processes that influence variation within and across populations (Luikart et al. 2003; Stinchcombe and Hoekstra 2008).

In more detail, the procedure of a population genomics approach to detect selection is as follows (see also Fig. 1): The first step is to formulate clear biological hypotheses with respect to contrasts between specific samples, and candidate loci to be included. A second step is the sampling of numerous loci (most of them thought to be neutral) in numerous individuals of the different groups of interest, e.g. putative subpopulations, by genotyping individuals for molecular markers such as microsatellites. Then a frequency distribution is produced of a measure of variation between the samples of interest, based on the range of loci that has been sampled. For neutral loci (e.g. most microsatellites), this is expected to

be a continuous distribution with a certain maximum with stochastic variation around it. Subsequently, a screen for outlier loci that exhibit values of the measure of variation that deviate from values for the rest of the genome is carried out. Obviously, the correct identification of outlier loci is crucial to the success of this approach. Methods and software to do this are summarized in the excellent reviews of Luikart et al. (2003) and Butlin (2010). Assuming that the variation at the majority of the sampled loci is the result of genome-wide effects, the outlier loci are thought to mark adaptive variation, since they are the result of locus-specific effects such as selection. Examples of how different forms of selection (stabilizing, directional, disruptive and balancing) influence the measure of variation between samples of interest are given in Black et al. (2001). The selection of candidate-loci in step 1 of the population genomics approach (see above), and the validation step that follows below, increase the rigor with which outliers can be adequately attributed to selection, rather than other locus-specific effects (e.g. mutation, assortative mating, and recombination). Other ways to confirm the presence of genuine outliers, and the cause of outlier behaviour, are reviewed in Luikart et al. (2003). Among others, repeatability of the detection of outlier loci in independent samples strengthens support for the correct identification of outliers, and, if the independent samples are from different localities with similar putative selection gradients, supports the interpretation that the outliers mark adaptive loci. Further validation of selection being indeed the cause of the deviation of the outlier locus should be obtained in a final step of the analysis. For the hawk moth-pollinated violet *Viola cazorlensis* this was done by comparing the phenotypic divergence of the floral trait of *V. cazorlensis* with the genetic divergence of the outlier loci (Herrera and Bazaga 2008). A strong relationship was found between the genetic divergence of the outlier loci and three floral traits, i.e. the length of the peduncle, spur and upper petal of the flower, while there was no relationship between the genetic divergence of the neutral loci and any of the investigated floral traits. In this validation, it is also important to consider alternative factors that may explain the presence of outlier loci. For example, Manel et al. (2009) correlated outliers in a phytophagous insect (the large pine weevil, *Hylobius abietis* (L.) to host plant use versus abiotic factors.

A specific advantage of the population genomics approach is that it enables the scanning of the genome for ecologically relevant (Thompson 2009b) genetic variation without having to know the phenotypes (Stinchcombe and Hoekstra 2008). Because the population genomics approach implies that the genome is scanned for outlier loci by using anonymous markers, no prior knowledge is needed regarding the phenotype of a trait of interest. Other advantages of the population genomics approach are that a possible breeding history is not needed and that this approach can be applied to naturally occurring non-model organisms. Population genomics becomes particularly useful if segregating sites can be related to traits of interest, by identifying functional genetic polymorphisms (Storz 2005; Vasemagi and Primmer 2005). For the lake whitefish (*Coregonus clupeaformis*), for example, a population genomics approach was combined with adaptive QTL (quantitative trait loci) mapping to examine growth rate differences between dwarf and normal ecotypes (Rogers and Bernatchez 2005). Significantly higher levels of divergence were found for several growth-associated QTLs.

Integration of population genomics approach and GMTC

In the context of the geographic mosaic of coevolution, the population genomics approach is an excellent tool to investigate the two basic processes leading to the three predictions of this theory (Fig. 2):

(a)   *Detection of variable selection leading to the identification of hot spots and cold spots that cause spatial variation in traits mediating interactions among species.*

Cold spots are easy to identify if one considers a patch where only one of the two interacting species occurs (Gomulkiewicz et al. 2007). Population genomics is, therefore, not needed to identify cold spots. Hot spots, however, can only be identified by showing that reciprocal selection occurs. If a certain trait is expected to be under selection, the population genomics approach can be used to establish whether the locus of such a candidate trait is indeed under selection, like in the previous example of the growth rate differences between two ecotypes of the whitefish (Rogers and Bernatchez 2005). Traits that are expected to be under selection should be checked in both interacting species to identify a genuine hot spot, since there, by definition, selection should be reciprocal.

(b)   *Detecting trait remixing and more specifically, gene flow*

In a geographical mosaic of coevolution, trait remixing is thought to be the underlying cause of trait mismatching in interacting species. Gene flow, next to local extinction and random genetic drift, is one of the processes that influences trait remixing and can, therefore, influence geographic selection mosaics. If a population displays considerable genetic structure at certain loci, the influence of gene flow on homogenizing the variation at such loci in the metapopulation is limited. The population structure, however, needs not be the same at all loci (Thompson 2005); if selection at a certain locus is stronger than migration, then gene flow at that locus may be limited between subpopulations. So, only measuring average gene flow and showing population structure by using molecular markers, is not enough to prove the existence of a geographical mosaic of coevolution. Population genomics is useful to separate the processes of selection and migration/dispersal, which both influence gene flow. When a frequency distribution is produced of the variation of the sampled loci, statistical parameters—such as the Fst statistic (Wright 1951), or any other appropriate measure of differentiation (Fst is not an appropriate measure under all circumstances, (see Gregorius et al. 2007; Gillet and Gregorius 2008); other measures, each with their own (dis) advantages (see e.g. Meirmans and Hedrick 2010; Sefc et al. 2007) include: Qst (Spitze 1993), Rst (Slatkin 1995), Gst (Nei 1987), Θ (Weir and Cockerham 1984), Φ (Excoffier et al. 1992), C (Xu et al. 2009), G'st (Hedrick 2005), D (Jost 2008), δ (Gregorius et al. 2007), Dm (Nei 1973)—can be estimated for the loci that are considered to only undergo genome-wide effects. After removing outliers, the mean Fst of these neutral loci, with a stochastic variation around it, should indicate to what extent subpopulations are differentiated with respect to genome-wide effects (Wright 1951), i.e. to what extent hot and cold spots are linked.

For *Viola cazorlensis* it was demonstrated that phenotypic traits such as the length of the flower petals are subject to selection by the pollinating hawk moth, because gene flow is high between the patches of *V. cazorlensis* and yet the phenotypic floral traits differed much between the patches (Herrera and Bazaga 2008). A strong average gene flow between the patches was found by sampling many neutral loci. Further details of the way population genomics can be employed to study the processes leading to a geographic mosaic of coevolutionary interactions are given in the examples below.

Examples of integrating the GMTC and population genomics

An interesting example of ecological adaptation is the interaction between the oligophagous flea beetle, *Phyllotreta nemorum* L., and one of its host plants, the crucifer *Barbarea*
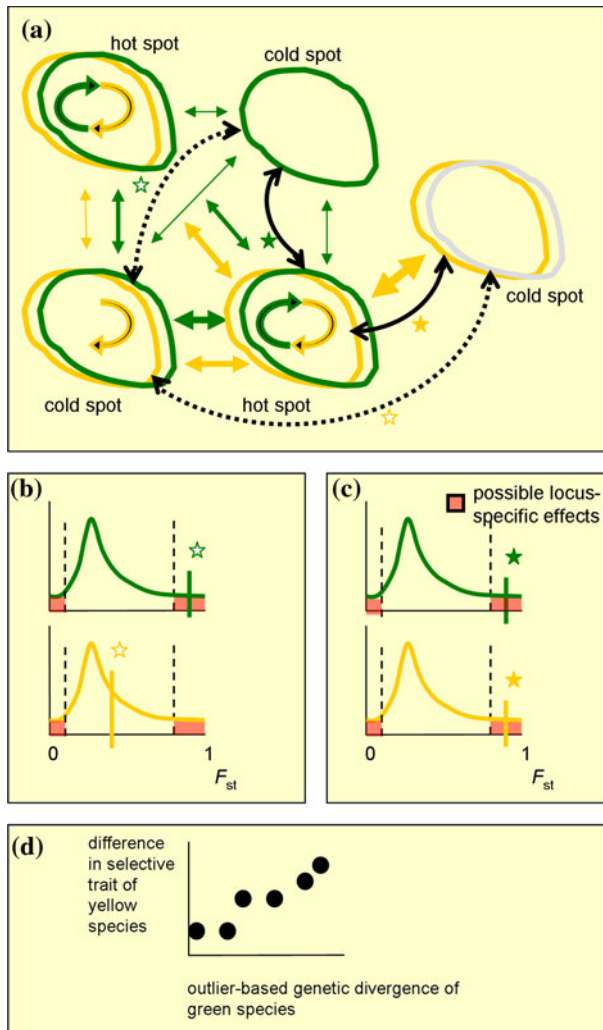
vulgaris ssp. arcuata (Opiz.) Simkovics. *Barbarea vulgaris* ssp. *arcuata* consists of two varieties (Nielsen 1997): the P-type which can be used as food by all individuals of *P. nemorum* and the G-type which is unsuitable for most flea beetles during summer. These types can hybridize in the field, yielding hybrids with intermediate chemical defence. Flea beetle individuals can be susceptible or resistant to the defence of the *B. vulgaris* G-type.

The host plants, as well as the flea beetles, have a patchy geographical distribution (de Jong et al. 2001, 2009). Resistance traits in both plant- and beetle varieties vary geographically and temporally (Nielsen and de Jong 2005; Toneatto et al. 2010). The defence of *B. vulgaris* varies in that some patches consist of the P-type and some of the G-type, thereby varying in suitability as well. In some patches nearly all *P. nemorum* are resistant to *B. vulgaris'* defence (G-type patches), whereas in other patches only some or none of the flea beetles are resistant (P-type patches). Most *B. vulgaris* patches (>80%) are free of *P. nemorum*, whereas *P. nemorum* is also found on other host plants than *B. vulgaris*, most

◀ **Fig. 2** Implementation of a population genomics approach to analyse the GMC: using population genomics to detect selection and to identify hot spots and cold spots. **a** A hypothetical example has been drawn for a set of local populations of two interacting species (respectively *green* and *yellow* patches). *Green* and *yellow arrows* between patches represent dispersal of the respective species, with *thicker arrows* indicating more dispersal between the local populations. *Curved arrows* within patches show selective interactions within local populations; a *green arrow* shows selection from the *green species* on the *yellow species*, and a *yellow arrow* vice versa. The occurrence of both a *yellow* and a *green arrow* within one patch indicates reciprocal selection, so that such a patch becomes a coevolutionary hot spot. In some patches, selection only acts from one species upon the other, and in other patches, no selection occurs between the two species of interest, e.g. when one of them is not present in that local patch. This is, for example, the case in the right-hand patch, where the yellow species is interacting with a third species (*grey*), and the green species is absent. **b**, **c** Examples of detection of selection within two local populations (the ones at the bottom of part **a** of the figure) by applying and Fst-based population genomics approach, comparing each of the two populations with local patches where only one of the two species is present (and hence no selection is exerted between them). These comparisons are indicated with curved, black (respectively *dashed* and *solid*) double-headed arrows in part **a**, accompanied by *green* and *yellow open*, and *closed asterisks*, respectively, for the two different pairwise comparisons. The *vertical bars* in **b** and **c** show the Fst value for candidate loci, suspected to be under selection. In **b**, the Fst for the candidate locus for the green species is an outlier, but for the yellow species it is not. This shows that the yellow species exerts selection on the green one, but not vice versa, and the local community is thus a coevolutionary cold spot. In **c**, the Fst-values for the candidate loci for both interacting species are outliers, showing that reciprocal selection may be occurring, and thus indicating a coevolutionary hot spot. **d** A validation involves the correlation of outlier Fst values of one species with the hypothesized selective pressure imposed by the other species (the latter needs to be determined empirically)

notably *Sinapis arvensis* L. On the non-*B. vulgaris* patches, frequencies of *P. nemorum* that are resistant to *B. vulgaris* G-type are relatively low (de Jong and Nielsen 1999). This leads to a geographic mosaic of interactions between *B. vulgaris* and *P. nemorum* (Nielsen and de Jong 2005).

Crosses revealed that the flea beetle's resistance is caused by a dominant major resistance gene (R-gene). The exact genetic basis at the DNA level of the resistance trait of *P. nemorum* is not yet known, although a candidate gene has been identified (K.M.C.A. Vermeer et al., unpublished results). For *B. vulgaris*, the biochemical basis of the defence against phytophagous insects (including *P. nemorum*) has been unraveled (Kuzina et al. 2009; Nielsen et al. 2010), and efforts are being made to elucidate the genetic basis of the production of this chemical defence (Kuzina et al. 2011). The presence of the different types of *B. vulgaris* on the one hand, and the different resistance genotypes of the flea beetles on the other hand, as well as their geographical distribution has prompted the question to what extent processes with genome-wide effect, such as dispersal/migration, and locus-specific effects, like selection, are responsible for these observed patterns. A population genomics approach is the obvious tool to examine this question. As prerequisites for application of this approach, microsatellite markers have been developed for both the flea beetles (Verbaarschot et al. 2007 and unpublished) and *B. vulgaris* (Toneatto et al. 2010). Furthermore, loci/markers are being identified for putative adaptive traits in both interacting species. One of the specific biological questions that can be addressed with population genomics is: can the difference in frequency of resistant flea beetles between other plants and *B. vulgaris* G-type be attributed to a limited dispersal of the beetles, or is selection on resistance in the flea beetles involved? By sampling flea beetles on *B. vulgaris* G-type and on other host plants and applying a population genomics approach using the microsatellite markers and the candidate gene for resistance, any involvement of selection can be detected. Samples of flea beetles on other host plants than *B. vulgaris* G-type, for example having different frequencies of resistant beetles (e.g. at different distances from a

B. vulgaris G-type patch) can also be contrasted in a population genomics approach. In this way, hypothesized selection against resistance on these other plants by disruption of coadapted gene complexes (de Jong and Nielsen 2002) can be detected and distinguished from effects of dispersal. Analogously, the population genomics approach can be used to detect selection on chemical defence in *B. vulgaris*. For example, patches with and without presence of *P. nemorum* can be compared and the distribution of alleles involved in the level of chemical defence can be attributed to selection versus dispersal. By combining the outcomes for the beetles and *B. vulgaris*, tentative coevolutionary hot spots, where selection for both interacting species is detected, can be identified (Fig. 2). This way of applying population genomics would be a slightly different approach than that from Black's original perspective; whereas they *identified* candidate loci by using outlier loci, this investigation would study whether the a priori candidate gene is an outlier, in order to draw conclusions with respect to the involvement of selection in the observed geographical distribution of the different alleles of the trait of interest.

Another example of a population genomics approach in plant–insect interactions is the study of the leaf beetle *Neochlamisus bebbianae* and its host plants (Funk 1998; Egan et al. 2008; Funk and Nosil 2008). The populations of the leaf beetle are associated with either maple trees or willow trees. These leaf beetle populations are partially differentiated in host preference and performance traits and they exhibit premating reproductive barriers (Funk 1998; Egan and Funk 2006). By using population genomics, Egan and co-workers (Egan et al. 2008) aim to investigate the contribution of host-plant related divergent selection to genetic differentiation during ecological speciation.

An ecologically comparative genome scan of AFLPs for pairwise population comparisons of the beetle *N. bebbianae* was made. Then different-host population comparisons were contrasted with same-host population comparisons. Outlier AFLPs in the first set of comparisons, that were not outliers in the second set of comparisons, should represent regions experiencing host-specific selection. Certain regions were found that matched these criteria and those candidate gene regions are now further investigated to find specific genes (and their function) that contribute to ecological speciation (Nosil et al. 2008). Not only will these genes provide insight to the process of speciation, they are also nice examples of genes encoding for traits that vary across the populations, differing in cold and hot spots.

Local adaptation and maladaptation can be found even in a generalist plant-pollinator system when combining the strength of population genomics with the geographic mosaic theory of coevolution. Gomez and co-workers found that *Erysimum mediohispanicum* forms a selection mosaic of varying selective regimes mediated by different pollinators (Gomez et al. 2009a, b). Although *E. mediohispanicum* and its interactors represent a generalist system, meaning that more than one-to-one interactions are present, it is composed of geographic mosaics of selection. This illustrates once again how important it is to also focus on the spatial scale of interactions, when investigating (co)evolution. Also generalist systems, involving many interacting organisms, can consist of different selective regimes caused by specific components of the community.

One of the predictions of the GMTC is the presence of a mosaic of local adaptation and maladaptation across the population. For *E. mediohispanicum* and its pollinators this was tested by comparing the plant's attractiveness to pollinators (Gomez et al. 2009a). Plants originating from hot spots (i.e. spots where selection on the plant by pollinator assemblages is large) were more attractive than plants from cold spots. Randomly amplified polymorphic DNA (RAPD) was used to evaluate genetic differences between populations. Since most RAPDs are thought to be neutral, they can be used to calculate the genetic distance between populations. The genetic difference between populations was compared

to differences in adaptation (being the attractiveness of the plants to pollinator assemblages, which can vary across different localities), to find out whether this trait was associated with the genetic differences between populations. Gene flow between the investigated populations was low, considering the large genetic differences found between populations with molecular markers. Knowing that plant fitness is highly influenced by pollinator visitation rate—plants in hot spots being more attractive than plants in cold spots—this indicates that *E. mediohispanicum* in hot spots is locally adapted, but maladapted in cold spots.

Conclusion and future perspectives

The concept of coevolution has been presented more than a century ago. So far, especially temporal aspects have been included in the study of coevolution, although temporal studies of coevolution in the true sense by monitoring reciprocal genetic change over time are relatively rare (J.N. Thompson, pers. comm.). To fully understand the evolutionary ecology of insect–plant interactions, the inclusion of spatial variation in selection pressures and reciprocal interactions between individuals is essential. The geographical mosaic theory of coevolution has proven to be a valuable approach in this respect. Furthermore, by applying a population genomic approach when examining the geographical mosaic of coevolution, one can understand more of the genetic variation influencing adaptive traits under natural conditions and thereby more of coevolution. Given the important evolutionary questions that are open in the context of the effects of climate change on biodiversity, community structure, and species interactions, combining population genomics with a coevolutionary approach that includes geographical aspects is likely to yield valuable information to understand the past and predict future scenarios.

References

Agrawal AF, Brodie ED, Wade MJ (2001) On indirect genetic effects in structured populations. Am Nat 158:308–323

Althoff DM, Thompson JN (1999) Comparative geographic structures of two parasitoid-host interactions. Evolution 53:818–825

Avise JC (2000) Stability, equilibrium and molecular aspects of conservation in marine species. Hydrobiologia 420:Xi–Xii

Benkman CW (1999) The selection mosaic and diversifying coevolution between crossbills and lodgepole pine. Am Nat 153:S75–S91

Benkman CW, Holimon WC, Smith JW (2001) The influence of a competitor on the geographic mosaic of coevolution between crossbills and lodgepole pine. Evolution 55:282–294

Benkman CW, Parchman TL, Favis A et al (2003) Reciprocal selection causes a coevolutionary arms race between crossbills and lodgepole pine. Am Nat 162:182–194

Berenbaum MR, Zangerl AR (1992) Genetics of physiological and behavioral resistance to host furanocoumarins in the parsnip webworm. Evolution 46:1373–1384

Bermingham E, Avise JC (1986) Molecular zoogeography of fresh-water fishes in the southeastern United-States. Genetics 113:939–965

Black WC IV, Baer CF, Antolin MF et al (2001) Population genomics: genome-wide sampling of insect populations. Annu Rev Entomol 46:441–469

Bohonak AJ (1998) Genetic population structure of the fairy shrimp *Branchinecta coloradensis* (Anostraca) in the Rocky Mountains of Colorado. Can J Zool 76:2049–2057

Brito PH, Edwards SV (2009) Multilocus phylogeography and phylogenetics using sequence-based markers. Genetica 135:439–455

Brodie ED Jr, Ridenhour BJ, Brodie ED III (2002) The evolutionary response of predators to dangerous prey: hotspots and coldspots in the geographic mosaic of coevolution between garter snakes and newts. Evolution 56:2067–2082

Burdon JJ, Thrall PH (1999) Spatial and temporal patterns in coevolving plant and pathogen associations. Am Nat 153:S15–S33

Butlin RK (2010) Population genomics and speciation. Genetica 138:409–418

Cornell HV, Hawkins BA (2003) Herbivore responses to plant secondary compounds: a test of phytochemical coevolution theory. Am Nat 161:507–522

Courtney S (1988) If it's not coevolution, it must be predation. Ecology 69:910–911

Darwin CR (1859) On the origin of species by means of natural selection, or the preservation of favoured races in the struggle for life. John Murray, London

Darwin CR (1862) On the various contrivances by which British and foreign orchids are fertilised by insects, and on the good effects of intercrossing. John Murray, London

de Jong PW, Nielsen JK (1999) Polymorphism in a flea beetle for the ability to use an atypical host plant. Proc R Soc Lond B 266:103–111

de Jong PW, Nielsen JK (2002) Host plant use of *Phyllotreta nemorum*: do coadapted gene complexes play a role? Ent Exp Appl 104:207–215

de Jong PW, de Vos H, Nielsen JK (2001) Demic structure and its relation with the distribution of an adaptive trait in danish flea beetles. Mol Ecol 10:1323–1332

de Jong PW, Breuker CJ, de Vos H et al (2009) Genetic differentiation between resistance phenotypes in the phytophagous flea beetle, *Phyllotreta nemorum*. J Ins Sci 9:69

Dicke M, van Loon JJA, de Jong PW (2004) Ecogenomics benefits community ecology. Science 305:618–619

Egan SP, Funk DJ (2006) Individual advantages to ecological specialization: insights on cognitive constraints from three conspecific taxa. Proc R Soc Lond B 273:843–848

Egan SP, Nosil P, Funk DJ (2008) Selection and genomic differentiation during ecological speciation: isolating the contributions of host association via a comparative genome scan of *Neochlamisus bebbianae* leaf beetles. Evolution 62:1162–1181

Ehrlich PR, Raven PH (1964) Butterflies and plants—a study in coevolution. Evolution 18:586–608

Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to the human mitochondrial DNA restriction data. Genetics 131:479–491

Fox LR (1981) Defense and dynamics in plant-herbivore systems. Am Zool 21:853–864

Fox LR (1988) Diffuse coevolution within complex communities. Ecology 69:906–907

Funk DJ (1998) Isolating a role for natural selection in speciation: host adaptation and sexual isolation in *Neochlamisus bebbianae* leaf beetles. Evolution 52:1744–1759

Funk DJ, Nosil P (2008) Comparative analyses and ecological speciation in herbivorous insects. In: Tilmon KJ (ed) Specialization, speciation, and radiation: the evolutionary biology of herbivorous insects. University of California Press, Berkeley, pp 117–135

Funk DJ, Filchak KE, Feder JL (2002) Herbivorous insects: model systems for the comparative study of speciation ecology. Genetica 116:251–267

Futuyma DJ (2009) Evolution. Sinauer Associates Inc., Sunderland

Futuyma DJ, Slatkin M (1983) Coevolution. Sinauer Associates Inc., Sunderland

Gillet EM, Gregorius H-R (2008) Measuring differentiation among populations at different levels of genetic integration. BMC Genet 9:60

Gomez JM, Abdelaziz M, Camacho JPM et al (2009a) Local adaptation and maladaptation to pollinators in a generalist geographic mosaic. Ecol Lett 12:672–682

Gomez JM, Perfectti F, Bosch J et al (2009b) A geographic selection mosaic in a generalized plant-pollinator-herbivore system. Ecol Monogr 79:245–263

Gomulkiewicz R, Drown DM, Dybdahl MF et al (2007) Dos and don'ts of testing the geographic mosaic theory of coevolution. Heredity 98:249–258

Gregorius H-R, Degen B, König A (2007) Problems in the analysis of genetic differentiation among populations—a case study in quercus robur. Silvae Genet 56:190–199

Hanski I (1998) Metapopulation dynamics. Nature 396:41–49

Hedrick PW (2005) A standardized genetic differentiation measure. Evolution 59:1633–1638

Herrera CM, Bazaga P (2008) Population-genomic approach reveals adaptive floral divergence in discrete populations of a hawk moth-pollinated violet. Mol Ecol 17:5378–5390

Hickerson MJ, Carstens BC, Cavender-Bares J et al (2010) Phylogeography's past, present and future: 10 years after Avise, 2000. Mol Phylogen Evol 54:291–301

Ikeda H, Setoguchi H (2010) Natural selection on PHYE by latitude in the Japanese archipelago: insight from locus specific phylogeographic structure in *Arcteria nana* (Ericaceae). Mol Ecol 19:2779–2791

Iwao K, Rausher MD (1997) Evolution of plant resistance to multiple herbivores: quantifying diffuse coevolution. Am Nat 149:316–335

Janzen DH (1980) When is it coevolution? Evolution 34:611–612

Jermy T (1984) Evolution of insect host plant relationships. Am Nat 124:609–630

Jost L (2008) Gst and its relatives do not measure differentiation. Mol Ecol 17:4015–4026

Kuzina V, Ekstrøm LT, Andersen SB et al (2009) Identification of defence compounds in *Barbarea vulgaris* against the herbivore *Phyllotreta nemorum* by an eco-metabolomic approach. Plant Physiol 151:1977–1990

Kuzina V, Nielsen JK, Augustin JM, Torp AM, Bak S, Andersen SB (2011) *Barbarea vulgaris* linkage map and quantitative trait loci for saponins, glucosinolates, hairiness and resistance to the herbivore *Phyllotreta nemorum*. Phytochemistry (in press)

Laine AL (2009) Role of coevolution in generating biological diversity: spatially divergent selection trajectories. J Exp Bot 60:2957–2970

Lively CM (1999) Migration, virulence, and the geographic mosaic of adaptation by parasites. Am Nat 153:S34–S47

Luikart G, England PR, Tallmon D et al (2003) The power and promise of population genomics: from genotyping to genome typing. Nat Rev Genet 4:981–994

Manel S, Conord C, Després L (2009) Genome-scan to assess the respective role of host-plant and environmental constraints on the adaptation of a widespread insect. BMC Evol Biol 9:288

Medrano M, Herrera CM (2008) Geographical structuring of genetic diversity across the whole distribution range of *Narcissus longispathus*, a habitat-specialist, Mediterranean narrow endemic. Ann Bot 102:183–194

Meirmans PG, Hedrick PW (2010) Assessing population structure: Fst and related measures. Mol Ecol Res (early view)

Minder AM, Widmer A (2008) A population genomic analysis of species boundaries: neutral processes, adaptive divergence and introgression between two hybridizing plant species. Mol Ecol 17:1552–1563

Nei M (1973) Analysis of gene diversity in subdivided populations. Proc Natl Acad Sci USA 70:3321–3323

Nei M (1987) Molecular evolutionary genetics. Columbia University Press, Columbia

Neuhauser C, Andow DA, Heimpel GE et al (2003) Community genetics: expanding the synthesis of ecology and genetics. Ecology 84:545–558

Nielsen JK (1997) Variation in defences of the plant *Barbarea vulgaris* and in counteradaptations by the flea beetle *Phyllotreta nemorum*. Entomol Exp Appl 82:25–35

Nielsen JK, de Jong PW (2005) Temporal and host-related variation in frequencies of genes that enable *Phyllotreta nemorum* to utilize a novel host plant, *Barbarea vulgaris*. Entomol Exp Appl 115:265–270

Nielsen JK, Nagao T, Okabe H et al (2010) Resistance in the plant, *Barbarea vulgaris*, and counter-adaptations in flea beetles mediated by saponins. J Chem Ecol 36:277–285

Nosil P, Egan SP, Funk DJ (2008) Heterogeneous genomic differentiation between walking-stick ecotypes: "Isolation by adaptation" and multiple roles for divergent selection. Evolution 62:316–336

Nuismer SL, Thompson JN (2006) Coevolutionary alternation in antagonistic interactions. Evolution 60:2207–2217

Nuismer SL, Gomulkiewicz R, Morgan MT (2003) Coevolution in temporally variable environments. Am Nat 162:195–204

Orr HA (2005) The genetic theory of adaptation: a brief history. Nat Rev Genet 6:119–127

Pellmyr O (2003) Yuccas, yucca moths, and coevolution: a review. Ann Missi Botl Gard 90:35–55

Rausher MD (1988) Is coevolution dead? Ecology 69:898–901

Rausher MD (1996) Genetic analysis of coevolution between plants and their natural enemies. Tr Genet 12:212–217

Rogers SM, Bernatchez L (2005) Integrating QTL mapping and genome scans towards the characterization of candidate loci under parallel selection in the lake whitefish (*Coregonus clupeaformis*). Mol Ecol 14:351–361

Schlotterer C (2002) Towards a molecular characterization of adaptation in local populations. Curr Op Genet Dev 12:683–687

Schmitt TM, Hay ME, Lindquist N (1995) Constraints on chemically mediated coevolution—multiple functions for seaweed secondary metabolites. Ecology 76:107–123

Schneider CJ (2008) Exploiting genomic resources in studies of speciation and adaptive radiation of lizards in the genus *Anolis*. Integr Comp Biol 48:520–526

Schoonhoven LM (2005) Insect–plant relationships: the whole is more than the sum of its parts. Entomol Exp Appl 115:5–6

Schoonhoven LM, van Loon JJA, Dicke ME (2005) Insect–plant biology. Oxford University Press, Oxford

Sefc KM, Payne RB, Sorensen MD (2007) Genetic differentiation after founder events: an evaluation of Fst estimators with empirical and simulated data. Evol Ecol Res 9:21–39

Slatkin M (1995) A measure of population subdivision based on microsatellite allele frequencies. Genetics 139:457–462

Spitze K (1993) Population structure in *Daphnia obtusa*: quantitative genetic and allozymic variation. Genetics 135:367–374

Stinchcombe JR, Hoekstra HE (2008) Combining population genomics and quantitative genetics: finding the genes underlying ecologically important traits. Heredity 100:158–170

Storz JF (2005) Using genome scans of DNA polymorphism to infer adaptive population divergence. Mol Ecol 14:671–688

Strong DR, Lawton JH, Southwood R (1984) Insects on plants: community patterns and mechanisms. Blackwell, Oxford

Thompson JN (1988) Coevolution and alternative hypotheses on insect plant interactions. Ecology 69:893–895

Thompson JN (1994) The coevolutionary process. Univ of Chicago Press, Chicago

Thompson JN (1999a) The evolution of species interactions. Science 284:2116–2118

Thompson JN (1999b) Specific hypotheses on the geographic mosaic of coevolution. Am Nat 153:S1–S14

Thompson JN (2005) The geographic mosaic of coevolution. University of Chicago Press, Chicago

Thompson JN (2009a) The coevolving web of life. Am Nat 173:125–140

Thompson JN (2009b) Which ecologically important traits are most likely to evolve rapidly? Oikos 118:1281–1283

Thompson JN, Pellmyr O (1992) Mutualism with pollinating seed parasites amid co-pollinators—constraints on specialization. Ecology 73:1780–1791

Thompson JN, Cunningham BM, Seagraves KA et al (1997) Plant polyploidy and insect/plant interactions. Am Nat 150:730–743

Toju H, Sota T (2006) Imbalance of predator and prey armament: geographic clines in phenotypic interface and natural selection. Am Nat 167:105–117

Toneatto F, Nielsen JK, Ørgaard M et al (2010) Genetic and sexual separation between insect resistant and susceptible *Barbarea vulgaris* plants in Denmark. Mol Ecol 19:3456–3465

Vasemagi A, Primmer CR (2005) Challenges for identifying functionally important genetic variation: the promise of combining complementary research strategies. Mol Ecol 14:3623–3642

Verbaarschot P, Calvo D, Esselink GD et al (2007) Isolation of polymorphic microsatellite loci from the flea beetle *Phyllotreta nemorum* L. (Coleoptera: Chrysomelidae). Mol Ecol Notes 7:60–62

Weir BS, Cockerham CC (1984) Estimating F-statistics for the analysis of population substructure. Evolution 38:1358–1370

Wright S (1951) The genetical structure of populations. Ann Eugen 15:323–354

Wright S (1968) Evolution and the genetics of populations: a treatise in four volumes. University of Chicago Press, Chicago

Xu H, Sarkar B, George V (2009) A new measure of population structure using multiple single nucleotide polymorphisms and its relationship with Fst. BMC Res Notes 2:21

Zangerl AR, Berenbaum MR (2003) Phenotype matching in wild parsnip and parsnip webworms: causes and consequences. Evolution 57:806–815