npg

REVIEW

Common garden experiments in the genomic era: new perspectives and opportunities

P de Villemereuil[1], OE Gaggiotti[1,2], M Mouterde[1] and I Till-Bottraud[1]

The study of local adaptation is rendered difficult by many evolutionary confounding phenomena (for example, genetic drift and demographic history). When complex traits are involved in local adaptation, phenomena such as phenotypic plasticity further hamper evolutionary biologists to study the complex relationships between phenotype, genotype and environment. In this perspective paper, we suggest that the common garden experiment, specifically designed to deal with phenotypic plasticity, has a clear role to play in the study of local adaptation, even (if not specifically) in the genomic era. After a quick review of some high-throughput genotyping protocols relevant in the context of a common garden, we explore how to improve common garden analyses with dense marker panel data and recent statistical methods. We then show how combining approaches from population genomics and genome-wide association studies with the settings of a common garden can yield to a very efficient, thorough and integrative study of local adaptation. Especially, evidence from genomic (for example, genome scan) and phenotypic origins constitute independent insights into the possibility of local adaptation scenarios, and genome-wide association studies in the context of a common garden experiment allow to decipher the genetic bases of adaptive traits.
Heredity (2016) **116,** 249–254; doi:10.1038/hdy.2015.93; published online 21 October 2015

INTRODUCTION

Studying adaptation and the genetic bases of the adaptive traits is an ambitious but daunting enterprise, especially for complex traits that have a polygenic basis and are strongly influenced by the environment. Indeed, uncovering the evidence of genetic adaptation is almost always hampered by the pervasive effects of evolutionary phenomena such as genetic drift, phenotypic plasticity, complex demographic history and complex genetic architecture. In the particular case of local adaptation, evolutionary biologists have developed efficient tools to overcome these challenges and the common garden experiment is one of them. The rationale behind this protocol is to control for the effects of phenotypic plasticity and, to a certain extent, genotype-by-environment interactions by growing individuals from different populations in a common environment, and by using the quantitative genetics toolbox (see Box 1) to study the genetic bases of complex traits (for example, life history, morphological and physiological traits).

Because it enables to unravel the genetic basis of complex phenotypes across various populations without the confounding effects of the corresponding environment, the common garden experiment is used to test for local adaptation signal in traits of interest such as life history traits (Kawakami *et al.*, 2011), phenology (Brachi *et al.*, 2013) and allometric relationships (Gonda *et al.*, 2011). Local adaptation might be suspected because of the existence of an environmental gradient such as latitude (Toräng *et al.*, 2015) or altitude (Alberto *et al.*, 2011), or because of the existence of several contrasting environments, such as sea and fresh water (DeFaveri and Merilä, 2014). In addition, common garden experiments are also used

to study the consequences of local adaptation for conservation (McKay *et al.*, 2001) or even for ecosystem functioning (Bassar *et al.*, 2010). Despite its name, and although it has been used extensively with plants (Linhart and Grant, 1996), this experimental approach can also be applied to a large variety of organisms including fish (Bassar *et al.*, 2010; DeFaveri and Merilä, 2014), invertebrates (Spitze, 1993; Luttikhuizen *et al.*, 2003) and small mammals (Bozinovic *et al.*, 2009). The main limitations to this experimental design are the ability to breed the species and to grow the produced offspring in laboratory or seminatural conditions. Common garden experiments can also be used to study genotype-by-environment interactions, by implementing the same design in different environments. Although replicating common garden experiments is logistically challenging, the outcomes of such experiments are highly rewarding, as genotype-by-environment effects are likely common and very important in the wild (Stinchcombe, 2014). Note finally that, although common garden experiments are closely related to reciprocal transplant experiments (which aim at testing local adaptation by showing that the average fitness of local individuals is higher than the average fitness of aliens, see for example Ågren and Schemske, 2012), there are important philosophical and practical differences between the two types of experiments. The difference is that reciprocal transplants are designed to prove local adaptation, whereas common gardens are designed to study the genetic bases of traits, regardless of whether they are adaptive or not. In practice, reciprocal transplants will typically create a differential survival, because the locals will survive better. This will be a confounding effect during the quantitative genetic analysis, because only the phenotypes of 'fit' individuals are available. Common

Box 1 Quantitative genetics glossary

Quantitative Genetics: Theoretical framework used to study the genetic basis of (mostly) quantitative polygenic traits. It uses relatedness between individuals to partition the phenotypic variance into (among others) genetic and nongenetic components.

Relatedness: Probability of shared ancestry (identity by descent (IBD)) of any two homologous alleles sampled among two individuals. Can also be defined in terms of correlation of homologous alleles between two individuals when the reference population is the sample itself. Relatedness is indeed always defined according to a reference population (Wang, 2014).

Additive genetic variance (V_A): Variance component due to the additive effects of the alleles and genes responsible for the phenotype. Under general conditions (no epistasis, no inbreeding), this is the only component transmitted to the offspring generation.

Dominance variance (V_D): Genetic variance arising from interactions between alleles within each gene responsible for the phenotype. The dominance effect is perceptible only when comparing full-sibs and in the presence of mild to strong inbreeding (Wolak and Keller, 2014).

Parental effects: Direct or indirect effects of the parental phenotype on the offspring phenotype, apart from the genetic heredity of the phenotype. This includes, in particular, maternal energetic investment in offspring.

Heritability: Proportion of the phenotypic variance genetically transmissible to the offspring generation within a population. Calculated as a ratio between V_A and the total phenotypic variance. The marker-based heritability is the proportion of phenotypic variance explained by the whole genetic marker panel that is not necessarily equal to the true heritability.

Q_{ST}: Among-population genetic differentiation index. Ratio of the among-population additive genetic variance V_{pop} to the total additive genetic variance (calculated as $V_{pop} + 2V_A$).

---

gardens, by contrast, are often designed to be 'softer' on the individuals. Nevertheless, most of the elements in this article regarding common garden experiments can also be applied to reciprocal transplants, especially if one is interested in applying them to survival or some other measure of fitness.

To perform the quantitative genetics analyses of the studied traits, individuals of controlled families (that is, group of individuals with known genealogy) are used. An average relatedness between individuals is derived from this known genealogy and allows to infer within-population additive genetic variance V_A, whereas effects due to the population of origin allows to infer the between-population additive genetic variance V_{pop}. This is so because all individuals share the same environment and, therefore, any average difference between populations must have a genetic origin. The residual variance V_R accounts for all other kinds of effects (for example, environmental). These variance components can be used to estimate the heritability of the trait:

$$h^2 = \frac{V_A}{V_A + V_R} \tag{1}$$

It is also possible to estimate Q_{ST}, a standardised measure of genetic differentiation for quantitative traits (Spitze, 1993; Edelaar et al., 2011). Q_{ST} is defined as the ratio of among-population (additive) genetic variance V_A over the total genetic variance (that is, including the within-population additive variance V_{pop}), and in the case of diploid species is given by:

$$Q_{ST} = \frac{V_{pop}}{V_{pop} + 2V_A} \tag{2}$$

This parameter is a quantitative analogue of population genetics' F_{ST} and, under a hypothesis of neutrality, both should be equal. Hence, a common approach for distinguishing between neutral drift and local adaptation scenarios is to compare Q_{ST}s and F_{ST}s. Consequently, individuals from a common garden experiment are typically genotyped to compute F_{ST}.
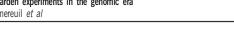
Despite the advantages of common garden experiments, the study of local adaptation in non-model species during the past decade has been strongly driven by the study of genetic markers in natural populations (Luikart et al., 2003). Typically, evolutionary biologists go to natural populations, sample tissue from the individuals and genotype them with high-throughput methods and then proceed with a genome scan analysis of selection (see, for example, Eckert et al., 2010; Bourret et al., 2013; Fischer et al., 2013). Although this method can be quite powerful, it has some limitations (for example, false positives, no information on the adaptive phenotype). Several calls have been made to independently validate the results of such analyses (see Buehler et al., 2014 for a striking example), possibly using common garden or reciprocal transplant experiments (Holderegger et al., 2008; Pardo-Diaz et al., 2014; Rellstab et al., 2015). Following these lines, this perspective paper addresses three main questions: where does the common garden experiment stand in the genomic era? In particular, what can common garden experiments bring to population genomics? Conversely, how can techniques from the genomic fields (for example, high-throughput genotyping and model-based inference of neutral evolution) extend the range and scope of common gardens?

It is important to note that population genomics aims at linking genotypes and environments through genome scans methods but often completely neglects to study the phenotypic traits under potential selection. There is much to gain by adding phenotypes into the equation (Cushman, 2014). Yet, because phenotypic plasticity is hard to distinguish from local adaptation in wild populations, it seems useless, or at least dubious, to use phenotypes directly obtained in the field. This simple fact lies at the heart of common garden experiments and we suggest here that this approach is ideally suited to jointly study genotypes, phenotypes and environments, especially when they are combined with high-throughput genotyping and powerful statistical methods. After a short introduction to the different high-throughput genotyping methods available in the context of a common garden experiment, we will discuss how those methods and powerful statistical tools can rejuvenate this classical approach. Finally, we will discuss the complementarity between population genomics and common garden experiments, and how an integrative analysis can deepen our understanding of local adaptation.

HIGH-THROUGHPUT GENOTYPING IN THE CONTEXT OF A COMMON GARDEN

High-throughput genotyping defines any genotyping method yielding a large number of markers, thus providing a dense marker panel across the genome. Given the focus on non-model species in this paper, we consider as few as 10 000 independent markers as fairly 'dense', provided that the genome of the species is not too large. For example, 10 000 single-nucleotide polymorphisms (SNPs) in a genome of size 100 Mbp would represent ~3% of all SNPs if a SNP occurs every 300 bp.

The most straightforward high-throughput genotyping method is whole-genome sequencing. This method yields the largest possible number of markers, and offers the densest genotyping. However, this technique requires high DNA quality and quantity, bioinformatics computation power and, most importantly, access to genomic resources (for example, genome assembly) within a relatively short phylogenetic range. The huge number of markers generated can also

be problematic during the analyses because of high computation/ memory requirements, high redundancy in information between linked markers and low signal-to-noise ratio. Still, whole- genome sequencing is the ultimate high-throughput genotyping method, yielding up to millions of SNP markers throughout the whole genome. With a decreasing cost and an increase in the number of species for which the whole genome has been sequenced over the years, it might soon become a recommended technique even for non-model species. A cheaper alternative to whole- genome sequencing are SNP genotyping chips, with most of the limitations above applying still.

For now, an approach likely to be best suited for non-model species is genome representation sequencing. The overall principle of this approach is to sequence only restricted, but random, parts of the genome in order to decrease the sequencing effort, and hence the overall costs and computational efforts associated with genotyping. To do so, the above approaches mainly use DNA digestion by restriction enzymes followed by a ligation of tags and primers and PCR amplification. This is akin to the principle underlying amplified fragment length polymorphism (AFLP) genotyping (Vos et al., 1995). Here, however, the DNA fragments (or at least some of them) are partially sequenced (~100 bp) using next-generation technology such as Illumina HiSeq (Illumina Inc., San Diego, CA, USA). This kind of approach includes the genotyping-by-sequencing method (Elshire et al., 2011) and the family of restriction site-associated DNA sequencing methods (Miller et al., 2007; Baird et al., 2008).

The sequences obtained are then analysed using quality checks (that is, selecting reads according to their sequencing quality, local coverage, availability over all or most individuals and so on) and SNP calling pipelines in order to identify SNP markers. Note that contrary to the AFLP approach, markers issued from restriction site-associated DNA sequencing are preferentially issued from nonpolymorphic restriction sites and are codominant. Alternatively, when more than one SNP is present on a 100-bp sequence, they can be combined into a new marker with more than two alleles. The rationale behind this is that very close SNPs are likely to be strongly associated because of physical linkage, in which case fewer but independent markers composed of more alleles are often preferable to strongly linked SNPs. Genome representation protocols can yield up to several hundreds of thousands of SNPs, but more typically tens of thousands. This can be achieved at a cost comparable or up to 10 times the cost of an AFLP analysis.

For all of the above, it is clear that next-generation sequencing makes possible the generation of a very large number of markers for a moderate cost. When compared with AFLP markers, next-generation sequencing marker panels are denser, and the markers are codominant and less arbitrary in their interpretation (that is, no 'binning' process), hence better in every way, except possibly for their cost. Microsatellites, on the other hand, are very different: they usually provide very sparse panels (up to a few dozens of markers), but highly mutable and with a large allelic diversity. Although it has been argued that microsatellites are better markers to infer relatedness (Ritland, 2000), they typically yield smaller relatedness estimates than SNP or AFLP markers because of higher mutation rates (Uptmoor et al., 2003; El Rabey et al., 2013). They also yield smaller F_{ST} estimates (Edelaar and Björklund, 2011) for the same reason. Finally, although in theory more accurate than SNPs for the same number of loci, they typically yield one to two orders of magnitude less loci, and hence they are less accurate in practice (Uptmoor et al., 2003).

A key issue is the number of individuals that need to be genotyped. Our view is that ideally all individuals from the experimental garden(s) should be genotyped, because this opens the way toward the more refined or novel analyses detailed below. However, some of the analyses suggested here (for example, genome scans) can be performed even when a subsample of individuals have been genotyped. De Kort et al. (2014), for example, have sampled one individual per family in their common garden experiment to combine it with population genomics (that is, genome scans) analyses. This cheaper subsampling procedure might be very attractive to researchers who are not interested in individual genotypes: that is, neither in the relatedness inference nor in the genome-wide association studies that are described below.

COMMON GARDENS 2.0: NEW MARKERS AND NEW METHODS
We are certainly not the first to encourage the evolutionary biology community to switch toward next-generation sequencing technology (Luikart et al., 2003; Savolainen et al., 2013), and it is clear that such a 'revolution' is already happening (reviewed in Pardo-Diaz et al., 2014). However, we wish here to emphasise the interest of dense marker panels in the context of a common garden experiment.

As stated above, a study of the genetics of complex traits such as that measured in common garden experiments strongly relies on the relatedness between individuals that is often assumed, especially when individuals are siblings (see, for example, Hernández-Serrano et al., 2014). Yet, contrary to the parent–offspring relationship, the relatedness between siblings varies: the commonly used value of 0.25 between half-sibs, for example, is only an average, expected value. Hence, using realised relatedness, inferred from molecular data, can allow for better estimates in the sense that (1) they are more robust to error in the kinship assessment (for example, full-sibs instead of half-sibs) and (2) they reflect more accurately the variation in relatedness between siblings. Better relatedness estimates are useful because they will improve the precision of the estimates of h^2 and Q_{ST}. Note however that many markers are typically needed to obtain precise molecular estimates of relatedness (Uptmoor et al., 2003). Dense markers provided by high-throughput genotyping naturally fulfill this requirement.

A large number of markers also allows the reconstruction of the family structure. Indeed, even when relatedness is precisely estimated, the family structure (that is, who is the mother/father of the individuals, which individuals are full- or half-sibs) is of utmost importance in order to account for many confounding effects such as dominance (Wolak and Keller, 2014), parental effects (for example, maternal, Wilson et al., 2010) or selfing (Gauzere et al., 2013). Note that maternal effects can also be accounted for by weighting seeds (in plants, Roach and Wulff, 1987) or reduced by using F2 generations (Roach and Wulff, 1987; Mousseau and Dingle, 1991). However, the possibility of using one of these methods will strongly depend on the studied species. According to Jones et al. (2010), brood size is one of the biggest limitations for parental reconstruction algorithms because of issues of unsampled alleles when too few segregating individuals are available. With many markers, even with low levels of polymorphism (such as SNPs), this is no longer an issue, as it becomes possible to reconstruct a large-enough proportion of the parental genomes to obtain high certainties of assignment, even for small brood sizes. Now that efficient algorithms such as those implemented in COLONY (Jones and Wang, 2010; Wang, 2012), are available, the number of markers should not be a problem. This software allows reconstructing the family structure, as well as inferring parental genotypes, while accounting for selfing or genotyping errors. Indeed, one crucial issue for parental inference with a large number of markers is to include possible genotyping errors that, if left unaccounted for, can severely bias the results (Wang, 2004).

The most innovative statistical method, especially designed to study common garden data, is probably the one developed by Ovaskainen et al. (2011) that overcomes several problems associated with the classical F_{ST}–Q_{ST} comparisons. In order to avoid clumsy comparisons between two noisy estimators, Ovaskainen et al. (2011) conceived a model of neutral phenotypic differentiation between populations that is compared with phenotypic differentiation measured in a common garden experiment (that is, the genetic differentiation linked to the phenotype). When suspiciously strong phenotypic differentiation is observed compared with the neutral expectation, a local adaptation hypothesis can be proposed. The neutral model of phenotypic differentiation is actually a combination of a within-population 'animal model' (see Kruuk, 2004 for a description of the model) and an among-populations 'F-model' (see Gaggiotti and Foll, 2010 for a description of the model) of phenotypic evolution (Karhunen and Ovaskainen, 2012). By doing so, this model allows for a multivariate genetic analysis to be performed, for example, to infer genetic correlations and a G matrix. This is a perfect illustration of how models emerging from the field of population genomics (here the F model) can be used to dramatically improve the analysis of common garden data sets. This method has been implemented in the DRIFTSEL package (Karhunen et al., 2013). Using this method, Karhunen et al. (2014) demonstrated the presence of strong footprints of local adaptation in several populations of nine-spine stickleback (Pungitius pungitius).

WHAT IS THE USE OF COMMON GARDEN EXPERIMENT IN THE GENOMIC ERA?
It is well known in the domain of genome-wide association studies, which aim at uncovering the loci responsible for phenotypic variation, that such analyses should be performed with extreme caution because of the potential effect of hidden population structure. Especially important are the combined effects of genetic drift and gene flow, and the confounding effect of phenotypic plasticity. However, both of the aforementioned problems can be overcome. Structure between population structure can be accounted for by using appropriate models (see, for example, Nicholson et al., 2002; Beaumont and Balding, 2004) or methods (Frichot et al., 2013) from the genome scan literature. The second problem, on the other hand, is perfectly addressed by common garden experiments that were specifically designed to control for phenotypic plasticity.

As a result, combining common garden experiments of non-model species with genome-wide association studies provides opportunity for multiple-population genome-wide association studies (Brachi et al., 2013; Slavov et al., 2014). For a locally adapted trait, it would even be possible to differentiate markers explaining among-population phenotypic variability (by testing for among-population effects) from markers explaining within-population variability (by testing for within-population effects). The technique of within-group centring (Davis et al., 1961; van de Pol and Wright, 2009) could be used to this end. It simply consists in distinguishing between the mean-population effect and the within-population effect of each predictor of an association model, as follows:

$$y_{ij} \sim \mu + \beta_B \overline{x}_j + \beta_W (x_{ij} - \overline{x}_j) + u_j + e_{ij}, \tag{3}$$

where y_{ij} is the phenotype of individual i in population j, x_{ij} is its genotype and $\overline{x}_j$ the mean genotype in population j. The parameters μ, β_B and β_W are the fixed effects of the model. Note that the within-population effects can be tested independently by using a parameter β_W^j for each population j. The term u_j stands for any population structure correction and e_{ij} is the residual. This equation is simply an illustration of within-group centring and does not constitute a model per se. Accounting for population structure should help in distinguishing between neutral and selective scenarios for markers associated with between-population variability. As always (Korte and Farlow, 2013), the power of a genome-wide association study to actually detect loci linked to the phenotypic variability strongly depends on the extent of linkage disequilibrium and the density of markers along the genome, in addition to the sample size. Hence, the most useful, but most expensive, genotyping method for this kind of analysis is whole-genome sequencing. Note also that heterogeneity in recombination/mutation rates along the genome can generate false positives during such analyses (Korte and Farlow, 2013). Here, the number of populations is also of importance, as it will determine the power to detect significance for the parameter β_B. Note that Brachi et al. (2013) used a different approach of multiscale (local to worldwide variation) analysis and found very different results depending on the studied scale of local adaptation. The approach that is probably the most typical of the genomic era is to scan genomes for signal of selection (mostly selective sweeps and local adaptation). Many methods have been developed in the past decades to detect local adaptation (Beaumont and Balding, 2004; Foll and Gaggiotti, 2008; Bonhomme et al., 2010; Coop et al., 2010; Frichot et al., 2013; Duforet-Frebourg et al., 2014; Guillot et al., 2014). Despite considerable efforts to account for population structures, these methods have been shown to display high error rates (de Villemereuil et al., 2014; Lotterhos and Whitlock, 2014). Hence, validation of the results of a genome scan must always be done using independent tests. Gene ontologies and pathway analyses are the most common mean of checking these results. However, it has been suggested that common garden experiments might be a very efficient complement to those analyses (De Kort et al., 2014; Lepais and Bacles, 2014; Rellstab et al., 2015).

Performing genome-scan analyses using common garden data can have many advantages. If a strong adaptive signal is detected both using both using genome scan methods (that is, using genotypic and possibly environmental data) and the phenotypic data from a common garden experiment, that will constitute two independent piece of evidence favouring the hypothesis of local adaptation (Holderegger et al., 2008). As stated above, genome scan results need to be validated anyhow (Pardo-Diaz et al., 2014; Rellstab et al., 2015), and performing a common garden experiment is an elegant way to do so. We suggested that, whenever possible, combining genome scan approaches with common garden experiments is an efficient approach to the study of local adaptation. Moreover, by comparing the loci showing strong signals of differentiation and the loci associated with among-population phenotypic differentiation, it is possible to isolate candidate loci for local adaptation with very little information regarding the functional annotation of the species' genome. Third, using the environmental information allows not only to identify the selected phenotypes (that is, strongly differentiated genetically), but also to infer the environmental variable driving the selective pressure. In particular, if a locus is strongly associated with an environmental variable and with the among-population phenotypic differentiation, one might conclude that a relationship exists between the environmental variable and the phenotype (although only correlatively: each variable is a putative proxy for the real selective/selected variable).

An important problem when performing genome-scan analyses directly on common garden individuals is to correctly infer the source-population allele frequencies. The preferred way is simply to genotype the parents of the common garden individuals. However, this is not always possible (for example, genotyping the father for plants is impossible most of the time). In that case, allele frequencies inferred

directly from the individuals should be accurate, as long as there is no sex-dependent allelic frequency bias. But the confidence in that inference will be overestimated by the fact that many related individuals were sampled. To account for this situation, a conservative solution is to calculate the allele frequencies based on the individuals of the common garden, but to consider that the sample size of these estimates are the number of parents that have generated the offspring. With these kind of data, all population-based methods (such as Bayescan, Foll and Gaggiotti, 2008, or BayEnv, Coop *et al.*, 2010) can be used. A second solution, if the confidence in parental genotypic reconstruction is high enough, is to directly use the inferred genotypes of the parents, both to infer allele frequencies in the population and directly as data for individual-based genome scan methods. Yet, in practice, these data will always be inferred with some uncertainty, and the consequences of ignoring this uncertainty during *post hoc* analyses is unknown. Still, the interest of this approach is that individual-based methods (such as Latent Factor Mixed Model, Frichot *et al.*, 2013, or PCAdapt, Duforet-Frebourg *et al.*, 2014) can be used to analyse the data. A last solution is the one implemented by De Kort *et al.* (2014) that consists in using only one individual *per* family. Although this solution requires a sufficiently large number of families for each population, it has the compelling advantage of simplicity and efficiency.

CONCLUSION

Local adaptation is a play starring three actors: the environment, the phenotype and the genotype. The environment selects the phenotypes that are (partly) determined by a number of genes. The evolutionary result is a change in allele frequencies of the polymorphic coding genes. Understanding the relationships between the three actors requires precise but large-scale measurements, rigorous experiments and powerful statistical methods. Because phenotypic plasticity is such a pervasive phenomenon and because it is nearly impossible to account for its effect on *in situ* phenotypes, phenotypes should never be directly compared between different populations, unless a case is made that the comparison is safe enough (low environmental contrasts or little phenotypic plasticity). In contrast, common garden experiments are ideally suited to perform such kinds of analyses, and hence to study the phenotypic traits affected by local adaptation. Now that dense marker panels are obtainable for many individuals at a moderate cost, common garden experiments are expected to be performed more routinely. Of course, this is unless the biological characteristics (for example size, behaviour, generation time) prevent the applicability of this experiment. Common gardens could possibly even replace the field work required to obtain tissue samples for genotyping: as we mentioned, it would still allow for population genomics approaches, while guaranteeing independent validation through the study of phenotypes (Pardo-Diaz *et al.*, 2014; Rellstab *et al.*, 2015), hence saving the cost of another genotyping campaign. As emphasised by Lepais and Bacles (2014), deciphering the genetic basis of local adaptation will only be accomplished by combining all the information yielded by dense marker panels, careful experiments and *in situ* sampling and observations. Replicating common garden experiments in different environments can also provide insight into complicated relationships between the three actors such as genotype-by-environment interactions. High-throughput genotyping provides an abundance of genetic data. World-wide fine-scale databases (for example, WorldClim, Hijmans *et al.*, 2005) and the advent of cheap *in situ* sensors also provide high-quality environmental data. However, collecting phenotypic data is still time consuming, tedious and sometimes expensive. It thus seems that the last challenge that needs to be overcome is the development of high-throughput phenotyping allowing for a scaling-up and a more widespread use of common garden experiments.

Ágren J, Schemske DW (2012). Reciprocal transplants demonstrate strong adaptive differentiation of the model organism Arabidopsis thaliana in its native range. *New Phytol* **194**: 1112–1122.

Alberto F, Bouffier L, Louvet JM, Lamy JB, Delzon S, Kremer A (2011). Adaptive responses for seed and leaf phenology in natural populations of sessile oak along an altitudinal gradient. *J Evol Biol* **24**: 1442–1454.

Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* **3**: e3376.

Bassar RD, Marshall MC, Lopez-Sepulcre A, Zandona E, Auer SK, Travis J et al. (2010). Local adaptation in Trinidadian guppies alters ecosystem processes. *Proc Natl Acad Sci USA* **107**: 3616–3621.

Beaumont MA, Balding DJ (2004). Identifying adaptive genetic divergence among populations from genome scans. *Mol Ecol* **13**: 969–980.

Bonhomme M, Chevalet C, Servin B, Boitard S, Abdallah J, Blott S et al. (2010). Detecting selection in population trees: the Lewontin and Krakauer test extended. *Genetics* **186**: 241–262.

Bourret V, Dionne M, Kent MP, Lien S, Bernatchez L (2013). Landscape genomics in Atlantic salmon (Salmo salar): searching for gene-environment interactions driving local adaptation. *Evolution* **67**: 3469–3487.

Bozinovic F, Rojas JM, Broitman BR, Vasquez RA (2009). Basal metabolism is correlated with habitat productivity among populations of degus (Octodon degus). *Comp Biochem Physiol A Mol Integr Physiol* **152**: 560–564.

Brachi B, Villoutreix R, Faure N, Hautekeete N, Piquot Y, Pauwels M et al. (2013). Investigation of the geographical scale of adaptive phenological variation and its underlying genetics in Arabidopsis thaliana. *Mol Ecol* **22**: 4222–4240.

Buehler D, Holderegger R, Brodbeck S, Schnyder E, Gugerli F (2014). Validation of outlier loci through replication in independent data sets: a test on Arabis alpina. *Ecol Evol* **4**: 4296–4306.

Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010). Using environmental correlations to identify loci underlying local adaptation. *Genetics* **185**: 1411–1423.

Cushman SA (2014). Grand challenges in evolutionary and population genetics: the importance of integrating epigenetics, genomics, modeling, and experimentation. *Front Genet* **5**: 197.

Davis J, Spaeth J, Huson C (1961). A technique for analyzing the effects of group composition. *Am Soc Rev* **26**: 215–225.

De Kort H, Vandepitte K, Bruun HH, Closset-Kopp D, Honnay O, Mergeay J (2014). Landscape genomics and a common garden trial reveal adaptive differentiation to temperature across Europe in the tree species Alnus glutinosa. *Mol Ecol* **23**: 4709–4721.

de Villemereuil P, Frichot E, Bazin E, Francois O, Gaggiotti OE (2014). Genome scan methods against more complex models: when and how much should we trust them? *Mol Ecol* **23**: 2006–2019.

DeFaveri J, Merilä J (2014). Local adaptation to salinity in the three-spined stickleback? *J Evol Biol* **27**: 290–302.

Duforet-Frebourg N, Bazin E, Blum MGB (2014). Genome scans for detecting footprints of local adaptation using a Bayesian factor model. *Mol Biol Evol* **31**: 2483–2495.

Eckert AJ, Bower AD, Gonzalez-Martinez SC, Wegrzyn JL, Coop G, Neale DB (2010). Back to nature: ecological genomics of loblolly pine (Pinus taeda, Pinaceae). *Mol Ecol* **19**: 3789–3805.

Edelaar P, Björklund M (2011). If FST does not measure neutral genetic differentiation, then comparing it with QST is misleading. Or is it? *Mol Ecol*; e-pub ahead of print 16 March 2011; doi:10.1111/j.1365-294X.2011.05051.x.

Edelaar P, Burraco P, Gomez-Mestre I (2011). Comparisons between QST and FST–how wrong have we been? *Mol Ecol* **20**: 4830–4839.

El Rabey H, Salem KF, Mattar MZ (2013). The genetic diversity and relatedness ofice (Oryza sativa L.) cultivars as revealed by AFLP and SSRs markers. *Life Sci J* **10**: 1471–1479.

Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One* **6**: e19379.

Fischer MC, Rellstab C, Tedder A, Zoller S, Gugerli F, Shimizu KK *et al*. (2013). Population genomic foot- prints of selection and associations with climate in natural populations of Arabidopsis halleri from the Alps. *Mol Ecol* **22**: 5594–5607.

Foll M, Gaggiotti OE (2008). A genome-scan method to identify selected loci appropriate for both dominant and codominant markers: a Bayesian perspective. *Genetics* **180**: 977–993.

Frichot E, Schoville SD, Bouchard G, Francois O (2013). Testing for associations between loci and environmental gradients using latent factor mixed models. *Mol Biol Evol* **30**: 1687–1699.

Gaggiotti OE, Foll M (2010). Quantifying population structure using the F-model. *Mol Ecol Resour* **10**: 821–830.

Gauzere J, Oddou-Muratorio S, Pichot C, Lefevre F, Klein E (2013). Biases in quantitative genetic analyses using open-pollinated progeny tests from natural tree populations. *Acta Bot Gallica* **160**: 227–238.

Gonda A, Herczeg G, Merila J (2011). Population variation in brain size of nine-spined sticklebacks (Pungitius pungitius) - local adaptation or environmentally induced variation? BMC. *Evol Biol* **11**: 75.

Guillot G, Vitalis R, Al Rouzic, Gautier M (2014). Detecting correlation between allele frequencies and environ-mental variables as a signature of selection. A fast computational approach for genome-wide studies. *Spatial Stat* **8**: 145–155.

Hernández-Serrano A, Verdu M, Santos-delBlancoL, ClimentJ, Gonzalez-Martinez SC, Pausas JG (2014). Heritability and quantitative genetic divergence of serotiny, a fire-persistence plant trait. *Ann Bot* **114**: 571–577.

Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. *Int J Climatol* **25**: 1965–1978.

Holderegger R, Herrmann D, Poncet B, Gugerli F, Thuiller W, Taberlet P *et al*. (2008). Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecol Divers* **1**: 273–283.

Jones AG, Small CM, Paczolt KA, Ratterman NL (2010). A practical guide to methods of parentage analysis. *Mol Ecol Resour* **10**: 6–30.

Jones OR, Wang J (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Mol Ecol Resour* **10**: 551–555.

Karhunen M, Merila J, Leinonen T, Cano JM, Ovaskainen O (2013). DRIFTSEL: an {R} package for detecting signals of natural selection in quantitative traits. *Mol Ecol Resour* **13**: 746–754.

Karhunen M, Ovaskainen O (2012). Estimating population-level coancestry coefficients by an admixture {F} model. *Genetics* **192**: 609–617.

Karhunen M, Ovaskainen O, Herczeg G, Merila J (2014). Bringing habitat information into statistical tests of local adaptation in quantitative traits: a case study of nine-spined sticklebacks. *Evolution* **68**: 559–568.

Kawakami T, Morgan TJ, Nippert JB, Ocheltree TW, Keith R, Dhakal P *et al*. (2011). Natural selection drives clinal life history patterns in the perennial sunflower species, Helianthus maximiliani. *Mol Ecol* **20**: 2318–2328.

Korte A, Farlow A (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* **9**: 29.

Kruuk LEB (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philos Trans R Soc Lond B Biol Sci* **359**: 873–890.

Lepais O, Bacles CFE (2014). Two are better than one: combining landscape genomics and common gardens for detecting local adaptation in forest trees. *Mol Ecol* **23**: 4671–4673.

Linhart YB, Grant MC (1996). Evolutionary significance of local genetic differentiation in plants. *Annu Rev Ecol Syst* **27**: 237–277.

Lotterhos KE, Whitlock MC (2014). Evaluation of demographic history and neutral parameterization on the performance of FST outlier tests. *Mol Ecol* **23**: 2178–2192.

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

Luttikhuizen PC, Drent J, Van Delden W, Piersma T (2003). Spatially structured genetic variation in a broadcast spawning bivalve: quantitative vs. molecular traits. *J Evol Biol* **16**: 260–272.

McKay JK, Bishop JG, Lin JZ, Richards JH, Sala A, Mitchell-Olds T (2001). Local adaptation across a climatic gradient despite small effective population size in the rare sapphire rockcress. *Proc R Soc Lond B Biol Sci* **268**: 1715–1721.

Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007). Rapid and cost-effective polymorphism identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Res* **17**: 240–248.

Mousseau TA, Dingle H (1991). Maternal effects in insect life histories. *Annu Rev Entomol* **36**: 511–534.

Nicholson G, Smith AV, Jaonsson F, Guastafsson O, Stefaansson K, Donnelly P (2002). Assessing population differentiation and isolation from single-nucleotide polymorphism data. *J R Stat Soc B (Stat Methodol)* **64**: 695–715.

Ovaskainen O, Karhunen M, Zheng C, Arias JMC, Merila J (2011). A new method to uncover signatures of divergent and stabilizing selection in quantitative traits. *Genetics* **189**: 621–632.

Pardo-Diaz C, Salazar C, Jiggins CD (2014). Towards the identification of the loci of adaptive evolution. *Methods Ecol Evol*; e-pub ahead of print 12 February 2015; doi:10.1111/2041-210X.12324.

Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015). A practical guide to environmental association analysis in landscape genomics. *Mol Ecol*; e-pub ahead of print 26 August 2015; doi:10.1111/mec.13322.

Ritland K (2000). Marker-inferred relatedness as a tool for detecting heritability in nature. *Mol Ecol* **9**: 1195–1204.

Roach DA, Wulff RD (1987). Maternal effects in plants. *Annu Rev Ecol Syst* **18**: 209–235.

Savolainen O, Lascoux M, Merila J (2013). Ecological genomics of local adaptation. *Nat Rev Genet* **14**: 807–820.

Slavov GT, Nipper R, Robson P, Farrar K, Allison GG, Bosch M *et al*. (2014). Genome-wide association studies and prediction of 17 traits related to phenology, biomass and cell wall composition in the energy grass Miscanthus sinensis. *New Phytol* **201**: 1227–1239.

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

Stinchcombe JR (2014). Cross-pollination of plants and animals: wild quantitative genetics and plant evolutionary genetics. In: Charmantier A, Garant D, Kruuk LE (eds). *Quantitative Genetics in the Wild*. Oxford University Press: Oxford, UK, pp 128–146.

Toräng P, Wunder J, Obeso JR, Herzog M, Coupland G, Agren J (2015). Large-scale adaptive differentiation in the alpine perennial herb Arabis alpina. *New Phytol* **206**: 459–470.

Uptmoor R, Wenzel W, Friedt W, Donaldson G, Ayisi K, Ordon F (2003). Comparative analysis on the genetic relatedness of Sorghum bicolor accessions from Southern Africa by RAPDs, AFLPs and SSRs. *Theor Appl Genet* **106**: 1316–1325.

van de Pol M, Wright J (2009). A simple method for distinguishing within- versus between-subject effects using mixed models. *Anim Behav* **77**: 753–758.

Vos P, Hogers R, Bleeker M, Reijans M, Tvd Lee, Hornes M *et al*. (1995). AFLP: a new technique for DNA fingerprinting. *Nucleic Acids Res* **23**: 4407–4414.

Wang J (2004). Sibship reconstruction from genetic data with typing errors. *Genetics* **166**: 1963–1979.

Wang J (2012). Computationally efficient sibship and parentage assignment from multi-locus marker data. *Genetics* **191**: 183–194.

Wang J (2014). Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current methods. *J Evol Biol* **27**: 518–530.

Wilson AJ, Reale D, Clements MN, Morrissey MM, Postma E, Walling CA *et al*. (2010). An ecologist's guide to the animal model. *J Anim Ecol* **79**: 13–26.

Wolak ME, Keller LF (2014). Dominance genetic variance and inbreeding in natural populations. In: Charmantier A, Garant D, Kruuk LE (eds). *Quantitative Genetics in the Wild*. Oxford University Press: Oxford, UK, pp 104–127.