View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by St Andrews Research Repository

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

Pierre de Villemereuil*, Oscar E. Gaggiotti*†, Médéric Mouterde*, & Irène Till-Bottraud*

*Université Joseph Fourier, Centre National de la Recherche Scientifique,

LECA, UMR 5553, 2233 rue de la piscine, 38400 Saint Martin d'Hères, France

†Scottish Oceans Institute, University of St Andrews,

Fife, KY16 8LB, United Kingdom

Keywords: common garden, high-throughput genotyping, genomics, adaptive traits

Corresponding author: Pierre de Villemereuil, E-mail: bonamy@horus.ens.fr

1 **Abstract**

2 The study of local adaptation is rendered difficult by many evolutionary confounding phenomenons (e.g.
3 genetic drift and demographic history). When complex traits are involved in local adaptation, phenomenons
4 such as phenotypic plasticity further hamper evolutionary biologists to study the complex relationships
5 between phenotype, genotype and environment. In this perspective paper, we suggest that the common
6 garden experiment, specifically designed to deal with phenotypic plasticity has a clear role to play in the
7 study of local adaptation, even (if not specifically) in the genomic era. After a quick review of some high-
8 throughput genotyping protocols relevant in the context of a common garden, we explore how to improve
9 common garden analyses with dense marker panel data and recent statistical methods. We then show how
10 combining approaches from population genomics and genome-wide association studies with the settings of a
11 common garden can yield to a very efficient, thorough and integrative study of local adaptation. Especially,
12 evidence from genomic (e.g. genome scan) and phenotypic origins constitute independent insights into the
13 possibility of local adaptation scenarios, and genome-wide association studies in the context of a common
14 garden experiment allow to decipher the genetic bases of adaptive traits.

15 Studying adaptation and the genetic bases of the adaptive traits is an ambitious but daunting enterprise,
16 especially for complex traits that have a polygenic basis and are strongly influenced by the environment. Indeed,
17 uncovering the evidence of genetic adaptation is almost always hampered by the pervasive effects of evolutionary
18 phenomena such as genetic drift, phenotypic plasticity, complex demographic history and complex genetic
19 architecture. In the particular case of local adaptation, evolutionary biologists have developed efficient tools to
20 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 (e.g. 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 a 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). Additionally, 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 semi-natural conditions. Common garden experiments can also be used to study genotype-by-environment (G×E) 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 G×E 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 e.g. 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 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 (i.e. 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 (e.g. environmental). These variance components can be used to estimate the heritability of the trait:

$$h^2 = \frac{V_{\text{A}}}{V_{\text{A}} + V_{\text{R}}} \tag{1}$$

It is also possible to estimate Q_{ST}, a standardized 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_{pop} over the total genetic variance (i.e. including the within-population additive variance V_{A}), and in the case of diploid species is given by:

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

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 last 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 (e.g. Eckert *et al.*, 2010; Bourret *et al.*, 2013; Fischer *et al.*, 2013). Although this method can be quite powerful, it has some limitations (e.g. 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 (e.g. 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 SNPs in a genome of size 100Mbp would represent approximately 3% of all SNPs if a SNP occurs every 300bp.

The most straightforward high-throughput genotyping method is whole genome sequencing (WGS). 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 (e.g. 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 WGS 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 WGS 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 AFLP genotyping (Vos *et al.*, 1995). Here, however, the DNA fragments (or at least some of them) are partially sequenced (approx. 100 bp) using next generation technology such as Illumina HiSeq. This kind of approach includes the genotyping-by-sequencing method (GBS, Elshire *et al.*, 2011) and the family of Restriction-site Associated DNA sequencing methods (RADseq, Miller *et al.*, 2007; Baird *et al.*, 2008).

The sequences obtained are then analysed using quality checks (i.e. selecting reads according to their sequencing quality, local coverage, availability over all or most individuals, etc.) and SNP calling pipelines, in order to identify SNP markers. Note that contrary to the AFLP approach, markers issued from RAD sequencing

118 are preferentially issued from non-polymorphic restriction sites and are co-dominant. Alternatively, when more
119 than one SNP is present on a 100bp sequence, they can be combined into a new marker with more than two
120 alleles. The rationale behind this is that very close SNPs are likely to be strongly associated due to physical
121 linkage, in which case, fewer but independent markers composed of more alleles are often preferable to strongly
122 linked SNPs. Genome representation protocols can yield up to several hundreds of thousands of SNPs, but
123 more typically tens of thousands. This can be achieved at a cost comparable or up to 10 times the cost of an
124 AFLP analysis.

125 For all of the above, it is clear that NGS makes possible the generation of a very large number of markers
126 for a moderate cost. When compared to AFLP markers, NGS marker panels are denser, the markers are
127 codominant, and less arbitrary in their interpretation (i.e. no "binning" process), hence better in every way,
128 except possibly for their cost. Micro-satellites, on the other hand, are very different: they usually provide very
129 sparse panels (up to a few dozens of markers), but highly mutable and with a large allelic diversity. Although
130 it has been argued that micro-satellites are better markers to infer relatedness (Ritland, 2000), they typically
131 yield smaller relatedness estimates than SNP or AFLP markers due to higher mutation rates (Uptmoor *et al.*,
132 2003; El Rabey *et al.*, 2013). They also yield smaller F_{ST} estimates (Edelaar and Björklund, 2011) for the same
133 reason. Finally, although in theory more accurate than SNPs for the same number of loci, they typically yield
134 one to two orders of magnitude less loci, hence they are less accurate in practice (Uptmoor *et al.*, 2003).

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

143 Common gardens 2.0: new markers and new methods

144 We are certainly not the first to encourage the evolutionary biology community to switch towards NGS technol-
145 ogy (Luikart *et al.*, 2003; Savolainen *et al.*, 2013), and it is clear that such a "revolution" is already happening
146 (reviewed in Pardo-Diaz *et al.*, 2014). However, we wish here to emphasise the interest of dense marker panels
147 in the context of a common garden experiment.

148 As stated above, a study of the genetics of complex traits such as measured in common garden experiments
149 strongly relies on the relatedness between individuals, which is often assumed, especially when individuals are
150 siblings (e.g. Hernández-Serrano *et al.*, 2014). Yet contrary to the parent-offspring relationship, the relatedness

151 between siblings varies: the commonly used value of 0.25 between half-sibs for example, is only an average, 152 expected value. Hence using realised relatedness, inferred from molecular data, can allow for better estimates 153 in the sense that *(i)* they are more robust to error in the kinship assessment (e.g. full-sibs instead of half-sibs) 154 and *(ii)* they reflect more accurately the variation in relatedness between siblings. Better relatedness estimates 155 are useful because they will improve the precision of the estimates of h^2 and Q_{ST}. Note however that many 156 markers are typically needed to obtain precise molecular estimates of relatedness (Uptmoor *et al.*, 2003). Dense 157 markers provided by high-throughput genotyping naturally fulfill this requirement.

158 A large number of markers also allows the reconstruction of the family structure. Indeed even when re159 latedness is precisely estimated, the family structure (i.e. who is the mother/father of the individuals, which 160 individuals are full- or half-sibs) is of utmost importance, in order to account for many confounding effects such 161 as dominance (Wolak and Keller, 2014), parental effects (e.g. maternal Wilson *et al.*, 2010) or selfing (Gauzere 162 *et al.*, 2013). Note that maternal effects can also be accounted for by weighting seeds (in plants, Roach and 163 Wulff, 1987) or reduced by using F2 generations (Roach and Wulff, 1987; Mousseau and Dingle, 1991). However, 164 the possibility of using one of these methods will strongly depend on the studied species. According to Jones 165 *et al.* (2010), brood size is one of the biggest limitations for parental reconstruction algorithms, due to issues 166 of unsampled alleles when too few segregating individuals are available. With many markers, even with low 167 levels of polymorphism (such as SNPs), this is no longer an issue, as it becomes possible to reconstruct a large168 enough proportion of the parental genomes to obtain high certainties of assignment, even for small brood sizes. 169 Now that efficient algorithms such as those implemented in COLONY (Jones and Wang, 2010; Wang, 2012), 170 are available, the number of markers should not be a problem. This software allows reconstructing the family 171 structure, as well as inferring parental genotypes, while accounting for selfing or genotyping errors. Indeed, 172 one crucial issue for parental inference with a large number of markers is to include possible genotyping errors, 173 which, if left unaccounted for, can severely bias the results (Wang, 2004).

174 The most innovative statistical method, especially designed to study common garden data, is probably the 175 one developed by Ovaskainen *et al.* (2011), which overcomes several problems associated with the classical F_{ST}– 176 Q_{ST} comparisons. In order to avoid clumsy comparisons between two noisy estimators, Ovaskainen *et al.* (2011) 177 conceived a model of neutral phenotypic differentiation between populations that is compared to phenotypic 178 differentiation measured in a common garden experiment (i.e. the genetic differentiation linked to the pheno179 type). When suspiciously strong phenotypic differentiation is observed compared to the neutral expectation, 180 a local adaptation hypothesis can be proposed. The neutral model of phenotypic differentiation is actually 181 a combination of a within-population "animal model" (see Kruuk, 2004, for a description of the model) and 182 an among-populations "F-model" (see Gaggiotti and Foll, 2010, for a description of the model) of phenotypic 183 evolution (Karhunen and Ovaskainen, 2012). By doing so, this model allows for a multivariate genetic analysis 184 to be performed, e.g. to infer genetic correlations and a G matrix. This is a perfect illustration of how models

185 emerging from the field of population genomics (here the F model) can be used to dramatically improve the
186 analysis of common garden datasets. This method has been implemented in the DRIFTSEL package (Karhunen
187 *et al.*, 2013). Using this method, Karhunen *et al.* (2014) demonstrated the presence of strong footprints of local
188 adaptation in several populations of nine-spine stickleback (*Pungitius pungitius*).

189 What is the use of common garden experiment in the genomic era?

190 It is well known in the domain of genome wide association studies (GWAS), which aim at uncovering the loci
191 responsible for phenotypic variation, that such analyses should be performed with extreme caution because of
192 the potential effect of hidden population structure. Especially important are the combined effects of genetic
193 drift and gene flow, and the confounding effect of phenotypic plasticity. However, both of the afore-mentioned
194 problems can be overcome. Structure between populations structure can be accounted for by using appropriate
195 models (e.g. Nicholson *et al.*, 2002; Beaumont and Balding, 2004) or methods (Frichot *et al.*, 2013) from the
196 genome scan literature. The second problem, on the other hand, is perfectly addressed by common garden
197 experiments, which were specifically designed to control for phenotypic plasticity!

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

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

205 where y_{ij} is the phenotype of individual i in population j, x_{ij} is its genotype and $\bar{x}_j$ the mean genotype
206 in population j. The parameters μ, β_B and β_W are the fixed effects of the model. Note that the within-
207 population effects can be tested independently by using a parameter β_W^j for each population j. The term u_j
208 stands for any population structure correction and e_{ij} is the residual. This equation is simply an illustration
209 of within-group centring and does not constitute a model *per se*. Accounting for population structure should
210 help in distinguishing between neutral and selective scenarios for markers associated with between-population
211 variability. As always (Korte and Farlow, 2013), the power of a GWAS to actually detect loci linked to the
212 phenotypic variability strongly depends on the extent of linkage disequilibrium and the density of markers along
213 the genome, in addition to the sample size. Hence the most useful, but most expensive, genotyping method for
214 this kind of analysis is WGS. Note also that heterogeneity in recombination/mutation rates along the genome

215 can generate false positives during such analyses (Korte and Farlow, 2013). Here the number of populations
216 is also of importance, as it will determine the power to detect significance for the parameter β_B. Note that
217 Brachi *et al.* (2013) used a different approach of multi-scale (local to worldwide variation) analysis and found
218 very different results depending on the studied scale of local adaptation. The approach that is probably the
219 most typical of the genomic era is to scan genomes for signal of selection (mostly selective sweeps and local
220 adaptation). Many methods have been developed in the last decades to detect local adaptation (Beaumont and
221 Balding, 2004; Foll and Gaggiotti, 2008; Coop *et al.*, 2010; Bonhomme *et al.*, 2010; Frichot *et al.*, 2013; Guillot
222 *et al.*, 2014; Duforet-Frebourg *et al.*, 2014). Despite considerable efforts to account for population structures,
223 these methods have been shown to display high error rates (Lotterhos and Whitlock, 2014; de Villemereuil *et al.*,
224 2014). Hence validation of the results of a genome scan must always be done using independent tests. Gene
225 ontologies and pathway analyses are the most common mean of checking these results. However, it has been
226 suggested that common garden experiments might be a very efficient complement to those analyses (Lepais and
227 Bacles, 2014; De Kort *et al.*, 2014; Rellstab *et al.*, 2015).

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

243 An important problem, when performing genome-scan analyses directly on common garden individuals, is
244 to correctly infer the source-population allele frequencies. The preferred way is simply to genotype the parents
245 of the common garden individuals. But this is not always possible (e.g. genotyping the father for plants is
246 impossible most of the time). In that case, allele frequencies inferred directly from the individuals should be
247 accurate, as long as there is no sex-dependent allelic frequency bias. But the confidence in that inference will
248 be overestimated by the fact that many related individuals were sampled. To account for this situation, a

249 conservative solution is to calculate the allele frequencies based on the individuals of the common garden, but
250 to consider that the sample size of these estimates are the number of parents that have generated the offspring.
251 With this kind of data, all population-based methods (such as Bayescan, Foll and Gaggiotti, 2008 or BayEnv,
252 Coop *et al.*, 2010) can be used. A second solution, if the confidence in parental genotypic reconstruction is high
253 enough, is to directly use the inferred genotypes of the parents both to infer allele frequencies in the population,
254 and directly as data for individual-based genome scan methods. Yet, in practice, these data will always be
255 inferred with some uncertainty, and the consequences of ignoring this uncertainty during *post-hoc* analyses is
256 unknown. Still, the interest of this approach is that individual-based methods (such as LFMM, Frichot *et al.*,
257 2013 or PCAdapt, Duforet-Frebourg *et al.*, 2014) can be used to analyse the data. A last solution is the one
258 implemented by De Kort *et al.* (2014), which consists in using only one individual *per* family. Although this
259 solution requires a sufficiently large number of families for each population, it has the compelling advantage of
260 simplicity and efficiency.

261 # Conclusion

262 Local adaptation is a play starring three actors: the environment, the phenotype and the genotype. The envi-
263 ronment selects the phenotypes, which are (partly) determined by a number of genes. The evolutionary result
264 is a change in allele frequencies of the polymorphic coding genes. Understanding the relationships between
265 the three actors requires precise but large-scale measurements, rigorous experiments and powerful statistical
266 methods. Because phenotypic plasticity is such a pervasive phenomenon and because it is nearly impossible to
267 account for its effect on *in situ* phenotypes, phenotypes should never be directly compared between different
268 populations, unless a case is made that the comparison is safe enough (low environmental contrasts or little
269 phenotypic plasticity). In contrast, common garden experiments are ideally suited to perform such kinds of
270 analyses, hence to study the phenotypic traits impacted by local adaptation. Now that dense marker panels
271 are obtainable for many individuals at a moderate cost, common garden experiments are expected to be per-
272 formed more routinely. Of course, this is unless the biological characteristics (e.g. size, behaviour, generation
273 time) prevent the applicability of this experiment. Common gardens could possibly even replace the field work
274 required to obtain tissue samples for genotyping: as we mentioned, it would still allow for population genomics
275 approaches, while guaranteeing independent validation through the study of phenotypes (Pardo-Diaz *et al.*,
276 2014; Rellstab *et al.*, 2015), hence saving the cost of another genotyping campaign. As emphasised by Lepais
277 and Bacles (2014), deciphering the genetic basis of local adaptation will only be accomplished by combining
278 all the information yielded by dense marker panels, careful experiments and *in situ* sampling and observations.
279 Replicating common garden experiments in different environments can also provide insight into complicated
280 relationships between the three actors such as genotype-by-environment interactions. High-throughput geno-
281 typing provides an abundance of genetic data. World-wide fine-scale databases (e.g. WorldClim, Hijmans *et al.*,

282 2005) and the advent of cheap *in situ* sensors also provide high quality environmental data. But collecting
283 phenotypic data is still time-consuming, tedious and sometimes expensive. It thus seems that the last challenge
284 that needs to be overcome is the development of high-throughput phenotyping allowing for a scaling-up and a
285 more widespread use of common garden experiments.

286 Aknowledgement

287 We thank the associate editor and three anonymous referees for their very thorough and relevant reviews, which
288 considerably improved the focus and quality of this manuscript. PdV was supported by a doctoral studentship
289 from the French *Ministère de la Recherche et de l'Enseignement Supérieur*. OEG was supported by the Marine
290 Alliance for Science and Technology for Scotland (MASTS).

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 non genetic 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 epistatis, 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 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, which 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 $V_{\mathrm{pop}} + 2V_{\mathrm{A}}$).

291 # References

292 Alberto F, Bouffier L, Louvet JM, Lamy JB, Delzon S, Kremer A (2011). Adaptive responses for seed and leaf
293 phenology in natural populations of sessile oak along an altitudinal gradient. *Journal of Evolutionary Biology*
294 **24**: 1442–1454.

295 Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, *et al.* (2008). Rapid SNP Discovery and
296 Genetic Mapping Using Sequenced RAD Markers. *PLoS ONE* **3**: e3376.

297 Bassar RD, Marshall MC, López-Sepulcre A, Zandonà E, Auer SK, Travis J, *et al.* (2010). Local adaptation
298 in Trinidadian guppies alters ecosystem processes. *Proceedings of the National Academy of Sciences* **107**:
299 3616–3621.

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

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

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

306 Bozinovic F, Rojas JM, Broitman BR, Vásquez RA (2009). Basal metabolism is correlated with habitat pro-
307 ductivity among populations of degus (Octodon degus). *Comparative Biochemistry and Physiology Part A:*
308 *Molecular & Integrative Physiology* **152**: 560–564.

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

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

314 Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010). Using Environmental Correlations to Identify Loci
315 Underlying Local Adaptation. *Genetics* **185**: 1411–1423.

316 Cushman SA (2014). Grand challenges in evolutionary and population genetics: the importance of integrating
317 epigenetics, genomics, modeling, and experimentation. *Evolutionary and Population Genetics* **5**: 197.

318 Davis J, Spaeth J, Huson C (1961). A Technique for Analyzing the Effects of Group Composition. *American*
319 *Sociological Review* **26**: 215–225.

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

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

325 DeFaveri J, Merilä J (2014). Local adaptation to salinity in the three-spined stickleback? *Journal of Evolutionary*
326 *Biology* **27**: 290–302.

327 Duforet-Frebourg N, Bazin E, Blum MGB (2014). Genome Scans for Detecting Footprints of Local Adaptation
328 Using a Bayesian Factor Model. *Molecular Biology and Evolution* **31**: 2483–2495.

329 Eckert AJ, Bower AD, González-Martínez SC, Wegrzyn JL, Coop G, Neale DB (2010). Back to nature: ecological
330 genomics of loblolly pine (Pinus taeda, Pinaceae). *Molecular Ecology* **19**: 3789–3805.

331 Edelaar P, Björklund M (2011). If FST does not measure neutral genetic differentiation, then comparing it with
332 QST is misleading. Or is it? *Molecular Ecology* .

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

335 El Rabey H, Salem KF, Mattar MZ (2013). The genetic diversity and relatedness of rice (Oryza sativa L.)
336 cultivars as revealed by AFLP and SSRs markers. *Life Science Journal* **10**.

337 Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, *et al.* (2011). A Robust, Simple
338 Genotyping-by-Sequencing (GBS) Approach for High Diversity Species. *PLoS ONE* **6**: e19379.

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

342 Foll M, Gaggiotti OE (2008). A Genome-Scan Method to Identify Selected Loci Appropriate for Both Dominant
343 and Codominant Markers: A Bayesian Perspective. *Genetics* **180**: 977 –993.

344 Frichot E, Schoville SD, Bouchard G, François O (2013). Testing for Associations between Loci and Environ-
345 mental Gradients Using Latent Factor Mixed Models. *Molecular Biology and Evolution* **30**: 1687–1699.

346 Gaggiotti OE, Foll M (2010). Quantifying population structure using the F-model. *Molecular Ecology Resources*
347 **10**: 821–830.

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

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

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

355 Hernández-Serrano A, Verdú M, Santos-del Blanco L, Climent J, González-Martínez SC, Pausas JG (2014). Heritability and quantitative genetic divergence of serotiny, a fire-persistence plant trait. *Annals of Botany* **114**: 571–577.

358 Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005). Very high resolution interpolated climate surfaces for global land areas. *International Journal of Climatology* **25**: 1965–1978.

360 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 Ecology & Diversity* **1**: 273–283.

362 Jones AG, Small CM, Paczolt KA, Ratterman NL (2010). A practical guide to methods of parentage analysis. *Molecular Ecology Resources* **10**: 6–30.

364 Jones OR, Wang J (2010). COLONY: a program for parentage and sibship inference from multilocus genotype data. *Molecular Ecology Resources* **10**: 551–555.

366 Karhunen M, Merilä J, Leinonen T, Cano JM, Ovaskainen O (2013). DRIFTSEL: an {R} package for detecting signals of natural selection in quantitative traits. *Molecular Ecology Resources* **13**: 746–754.

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

370 Karhunen M, Ovaskainen O, Herczeg G, Merilä 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.

372 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. *Molecular Ecology* **20**: 2318–2328.

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

377 Kruuk LEB (2004). Estimating genetic parameters in natural populations using the 'animal model'. *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* **359**: 873 –890.

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

381 Linhart YB, Grant MC (1996). Evolutionary Significance of Local Genetic Differentiation in Plants. *Annual Review of Ecology and Systematics* **27**: 237–277.

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

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

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

389 McKay JK, Bishop JG, Lin JZ, Richards JH, Sala A, Mitchell-Olds T (2001). Local adaptation across a climatic
390 gradient despite small effective population size in the rare sapphire rockcress. *Proceedings of the Royal Society*
391 *of London B: Biological Sciences* **268**: 1715–1721.

392 Miller MR, Dunham JP, Amores A, Cresko WA, Johnson EA (2007). Rapid and cost-effective polymorphism
393 identification and genotyping using restriction site associated DNA (RAD) markers. *Genome Research* **17**:
394 240–248.

395 Mousseau TA, Dingle H (1991). Maternal effects in insect life histories. *Annual review of entomology* **36**:
396 511–534.

397 Nicholson G, Smith AV, Jónsson F, Gústafsson O, Stefánsson K, Donnelly P (2002). Assessing Population
398 Differentiation and Isolation from Single-Nucleotide Polymorphism Data. *Journal of the Royal Statistical*
399 *Society Series B (Statistical Methodology)* **64**: 695–715.

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

402 Pardo-Diaz C, Salazar C, Jiggins CD (2014). Towards the identification of the loci of adaptive evolution.
403 *Methods in Ecology and Evolution* pp. n/a–n/a.

404 Rellstab C, Gugerli F, Eckert AJ, Hancock AM, Holderegger R (2015). A practical guide to environmental
405 association analysis in landscape genomics. *Molecular Ecology* pp. n/a–n/a.

406 Ritland K (2000). Marker-inferred relatedness as a tool for detecting heritability in nature. *Molecular Ecology*
407 **9**: 1195–1204.

408 Roach DA, Wulff RD (1987). Maternal Effects in Plants. *Annual Review of Ecology and Systematics* **18**:
409 209–235.

410 Savolainen O, Lascoux M, Merilä J (2013). Ecological genomics of local adaptation. *Nature Reviews Genetics*
411 **14**: 807–820.

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

415 Spitze K (1993). Population structure in Daphnia obtusa: quantitative genetic and allozymic variation. *Genetics*
416 **135**: 367–374.

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

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

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

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

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

429 Wang J (2004). Sibship Reconstruction From Genetic Data With Typing Errors. *Genetics* **166**: 1963–1979.

430 Wang J (2012). Computationally Efficient Sibship and Parentage Assignment from Multilocus Marker Data.
431 *Genetics* **191**: 183–194.

432 Wang J (2014). Marker-based estimates of relatedness and inbreeding coefficients: an assessment of current
433 methods. *Journal of Evolutionary Biology* **27**: 518–530.

434 Wilson AJ, Réale D, Clements MN, Morrissey MM, Postma E, Walling CA, *et al.* (2010). An ecologist's guide
435 to the animal model. *Journal of Animal Ecology* **79**: 13–26.

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

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