

## NEWS AND VIEWS

## MEETING REVIEW

**Uncovering the genetic basis of adaptive change: on the intersection of landscape genomics and theoretical population genetics**

STÉPHANE JOOST,\* SÉVERINE VUILLEUMIER,† JEFFREY D. JENSEN,‡§ SEAN SCHOVILLE,¶ KEVIN LEEMPOEL,\* SYLVIE STUCKI,\* IVO WIDMER,\* CHRISTELLE MELODELIMA,\*\* JONATHAN ROLLAND†† and STÉPHANIE MANEL‡‡§§  
 \*Laboratory of Geographic Information Systems (LASIG), School of Civil and Environmental Engineering (ENAC), École Polytechnique Fédérale de Lausanne (EPFL), Bâtiment GC, Station 18, 1015 Lausanne, Switzerland; †Department of Ecology and Evolution, University of Lausanne, Biophore Building, 1015 Lausanne, Switzerland; ‡Institute of Bioengineering, School of Life Sciences, École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland; §Swiss Institute of Bioinformatics, 1015, Lausanne, Switzerland; ¶CNRS, TIMC-IMAG UMR 5525, Université Joseph Fourier, 38041 Grenoble, France; \*\*Laboratoire d'Ecologie Alpine, UMR-CNRS 5553, Université Joseph Fourier, 38041 Grenoble, France; ††Centre de mathématiques appliquées, Ecole Polytechnique, 91128 Palaiseau Cedex, France; ‡‡Laboratoire Population Environnement Développement, UMR 151 UP/IRD, Université Aix Marseille, 3 place Victor Hugo, 13331 Marseille Cedex 03, France; §§UMR Botanique et BioinformaTique de l'Architecture des Plantes (AMAP), TA A51/PS2, 34398 Montpellier Cedex 5, France

A workshop recently held at the École Polytechnique Fédérale de Lausanne (EPFL, Switzerland) was dedicated to understanding the genetic basis of adaptive change, taking stock of the different approaches developed in theoretical population genetics and landscape genomics and bringing together knowledge accumulated in both research fields. Indeed, an important challenge in theoretical population genetics is to incorporate effects of demographic history and population structure. But important design problems (e.g. focus on populations as units, focus on hard selective sweeps, no hypothesis-based framework in the design of the statistical tests) reduce their capability of detecting adaptive genetic variation. In parallel, landscape genomics offers a solution to several

of these problems and provides a number of advantages (e.g. fast computation, landscape heterogeneity integration). But the approach makes several implicit assumptions that should be carefully considered (e.g. selection has had enough time to create a functional relationship between the allele distribution and the environmental variable, or this functional relationship is assumed to be constant). To address the respective strengths and weaknesses mentioned above, the workshop brought together a panel of experts from both disciplines to present their work and discuss the relevance of combining these approaches, possibly resulting in a joint software solution in the future.

**Keywords:** Adaptation, Genome scans, Landscape genomics, Selection, Theoretical population genetics

Received 05 March 2013; revised 15 April 2013; accepted 15 April 2013

There is a strong interest in identifying adaptive patterns of genetic diversity in both model and nonmodel species (e.g. Schoville *et al.* 2012). To achieve this task, theoretical population genetics attempts to characterize the relative roles of adaptive and neutral processes in shaping genomic variation (e.g. Crisci *et al.* 2012) by means of increasingly sophisticated methods and ever larger genomic data sets. In parallel, landscape genetics (Manel *et al.* 2003) aims to provide information about the interaction between landscape features and micro-evolutionary processes, and more specifically, the subfield of landscape genomics (Joost *et al.* 2007) uses correlative approaches between genetic data and environmental variables to identify regions of the genome possibly under selection. Both approaches have strengths and weaknesses and could benefit from pooling their respective assets.

Within this research framework, two 'MESSAGE' ('Méthodologies et Statistiques Spatiales Appliquées à la Génétique Environnementale') workshops took place in the context of a project funded by the Germaine de Staël Program, whose goal is to enhance collaborations between research teams from France and Switzerland (Swiss Academy of Engineering Sciences and Campus France). The aim of 'MESSAGE' was to initiate a multidisciplinary collaborative framework constituted of regular theoretical and training meetings dealing with statistical methodologies used in landscape genetics. It resulted in a stimulating environment for PhD students and postdoctoral researchers from the Ecole Polytechnique Fédérale de Lausanne (EPFL) in Switzerland, the Joseph Fourier University in Grenoble (France) and the University Aix-Marseille (France). The discussions enabled PhD students and postdoctoral researchers to present their work in progress, to expand their theoretical knowledge and allowed for a general discussion

of current technical skills. The workshops themselves focused upon inviting renowned scientists based in France and Switzerland to discuss important topics within their respective fields.

A first workshop took place in December 2011 and was entitled 'Detection of genomic regions under natural selection in landscape genetics'. It gave the opportunity to Matthieu Foll, Felix Gugerli, Stéphane de Mita and Christian Parisod to share their latest work with 'MESSAGE' participants as well as with an enlarged audience. The second workshop we report here was organized on 13 December 2012 at EPFL. It was dedicated to the idea of bringing together knowledge accumulated in both theoretical population genetics and landscape genomics, to discuss respective strengths and weaknesses and to consider the relevance of combining these approaches in order to identify genomic regions under selection, possibly resulting in a joint software solution in the future.

First, Jeffrey Jensen reminded the audience that one of the principle aims of population genetics is to develop an understanding of the relative roles of adaptive and nonadaptive processes in shaping patterns of genomic variation and in describing the evolutionary trajectory of mutations in natural populations. Particularly with the advent of next-generation sequencing, researchers are becoming increasingly focused on two issues: (i) quantifying the distribution of selection coefficients of newly arising, segregating and fixed mutations; and (ii) describing the neutral demographic history of the population under consideration.

With this, the last decade has seen a tremendous proliferation of computational approaches for the estimation of these population genetic parameters. These methodologies take both likelihood- and approximate Bayesian computation (ABC, an approach that allows for the consideration of more complex models as it avoids the need for a likelihood function, see review of Beaumont (2010))-based forms and rely on expected patterns in the site frequency spectrum (i.e. the frequencies of segregating mutations), linkage disequilibrium (i.e. the association between segregating mutations) and divergence data (i.e. the fixed differences between population/species) – individually or in combination. While intimately related, these estimators have largely been developed for three separate purposes: demographic estimation (e.g. Thornton & Andolfatto 2006), genomic scans for adaptive fixations (e.g. Pavlidis *et al.* 2010) and recurrent hitchhiking estimation (e.g. Jensen *et al.* 2008).

And yet, comparing between existing estimators (see Table 1), a number of notable discrepancies arise. Of the recently proposed recurrent hitchhiking estimators for example (e.g. Li & Stephan 2006; Andolfatto 2007; Macpherson *et al.* 2007; Jensen *et al.* 2008), estimates of the strength and rate of selection differ by orders of magnitude when applied to similar data sets. However, divergence-based methods (such as those based on the MK (McDonald & Kreitman 1991) or HKA (Hudson *et al.* 1987) frameworks) are counting the effective fixations of many weakly selected mutations over longer evolution time – describing

long-term rates of fixation at putatively selected sites (e.g. nonsynonymous) relative to putatively neutral sites (e.g. synonymous), while polymorphism-based methods are most impacted by the recent fixation of strongly beneficial mutations. Thus, it is indeed possible that these estimates are compatible – simply estimating different tails of the true underlying distribution of selection coefficients.

One commonality between approaches, however, is the inability to adequately account for the demographic history of the population in question. Nonequilibrium models are well known to mimic patterns of positive selection in polymorphism data (e.g. see reviews of Nielsen 2005; Thornton *et al.* 2007), and there is accumulating evidence that they may similarly impact divergence-based approaches (e.g. Andolfatto 2008). While attempts at joint estimation have been made (e.g. Williamson *et al.* 2005; Li & Stephan 2006; Eyre-Walker & Keightley 2009), they are accomplished in a stepwise manner. Thus, the demographic model is likely to over fit the data, accounting for much of the selection signature in the genome. Similarly, in the absence of demographic estimation, selection models are likely to be biased towards higher rates and strengths of adaptation in an attempt to fit the diversity-reducing and frequency spectrum-skewing effects produced by the underlying population history.

Thus, the challenge to the field is clear – it is essential to develop an estimator capable of jointly inferring the action of both non-neutral and nonequilibrium models simultaneously. This will require at least two components: (i) the ability to identify patterns that distinguish selective from demographic effects. The most promising avenue in this regard seems to be patterns in linkage disequilibrium (LD) that appear to be largely robust to demographic perturbation, as the specific pattern generated (i.e. strong LD flanking the fixation owing to hitchhiking effects, with reduced LD spanning the fixation owing to independent recombination events on either side of the fixation) appears difficult to reproduce under most neutral demographic models (Stephan *et al.* 2006; Jensen *et al.* 2007; Pavlidis *et al.* 2010). Additionally, the combination of polymorphism- and divergence-based inference may be effectively used to estimate different tails of the true underlying distribution – as discussed above; and (ii) a computational framework capable of handling whole genomes worth of data, a large number of summary statistics and the accurate inference of multiple parameters of interest. A good deal of recent work (e.g. Wegmann *et al.* 2009; Bazin *et al.* 2010) seems to suggest ABC-based approaches to be the most likely way forward.

Sean Schoville provided several criticisms of the theoretical population genetics outlier-detection methods that concern constraints in the design of these statistical tests to detect adaptive genetic variation. In particular, the focus on populations as units may be problematic: indeed, populations must be well defined and this is not straightforward in species that exhibit weak genetic structure, isolation by distance across continuous landscapes or when the study has to rely on few individuals. Another issue is that

**Table 1** Nonexhaustive list of current approaches for detecting adaptive change in theoretical population genetics and landscape genomics. Column five highlights the inability of theoretical population genetics estimators to *adequately* account for the demographic history of the population in question, the most promising avenue in this regard being patterns in linkage disequilibrium. The question marks in the same column indicate that the approach has not been tested yet.

| Estimator  | Type of question/<br>Situations  | Comments  | Robust to            |                     | References  |
|--|--|---|----------------------|---------------------|---|
|  |  |   | Population structure | Demographic history |   |
| <b>Theoretical population genetics</b>   |  |   |                      |                     |   |
| Site frequency spectrum  | Recent fixation of strongly beneficial mutations (impacting large genomic regions)           | Nonequilibrium models mimic patterns of positive selection  | Yes                  | Not adequate        | Nielsen (2005); Thornton <i>et al.</i> (2007)                                     |
| Linkage disequilibrium   | Demographic perturbation, positive selection (genetic hitchhiking)                           | False positives may arise in the presence of gene conversion  | Yes                  | Yes                 | Stephan <i>et al.</i> (2006); Jensen <i>et al.</i> (2007); Jones & Wakeley (2008) |
| Inter- and intraspecific divergence  | Fixation of selected mutations over longer evolutionary time                                 | Nonequilibrium models mimic patterns of positive selection  | Yes                  | Not adequate        | Beaumont & Balding (2004); Yang & Nielsen (2002); Andolfatto (2008)               |
| <b>Landscape genomics</b>  |  |   |                      |                     |   |
| Logistic regression model (correlative approach)   | Individual-based approach (spatial distribution of alleles), environmental variables         | Independent of any theoretical population genetic model, fast computation, generate more false positives under standard conditions (De Mita <i>et al.</i> 2013) | No                   | No                  | Joost <i>et al.</i> (2007)  |
| Generalized Estimating Equations (GEE)   | Consider spatial autocorrelation between individuals collected at the same sampling location | Detect false positive signatures of selection due to spatial autocorrelation.   | No                   | No                  | Poncet <i>et al.</i> (2010);  |
| Moran's eigenvector maps (MEM)   | Incorporate the effect of unaccounted environmental variables                                | Large-scale analysis  | No                   | No                  | Manel <i>et al.</i> (2010);   |
| Regression with null distribution used in statistical testing from a separate 'control data set' (covariance matrix) | Weak genetic structure, isolation by distance across continuous landscapes                   | Difficult to designate an appropriate control data set <i>a priori</i> (if neutral loci are not known)  | Yes                  | ?                   | Hancock <i>et al.</i> (2008); Coop <i>et al.</i> (2010); Günther & Coop (2012)    |
|  |  | Integrated nested Laplace approximations (INLA), fast computation   | Yes                  | ?                   | Guillot (2012)  |
| Mixed modelling (latent factor mixed model, LFMM)  | Few individuals (not enough data to create control data set)                                 | No need for a control data set, statistical learning techniques, fast computation   | Yes                  | ?                   | Frichot <i>et al.</i> (2013)  |

outlier-detection methods are most powerful in identifying loci that exhibit strong shifts in allele frequency occurring after the appearance of a novel mutation. These 'hard selective sweeps' may be infrequent in species with moderate to small population sizes (Hernandez *et al.* 2011) and only play a more prominent role in large population species such as *Drosophila* (Jensen *et al.* 2008). A last drawback mentioned is the lack of an ecological hypothesis supporting the statistical test. Outlier-detection methods identify loci with extreme levels of genetic differentiation, which can be examined a posteriori for relationships to some environmental variable. For example, Pavlidis *et al.* (2012) showed that false positives from data sets simulated under neutrality can lead to significant enrichment of gene categories with intriguing (and positively misleading) biological functions. But it can be argued that a strength of population genetic approaches is indeed the 'genotype-first' aspect – that is, the ability to identify putatively adaptive regions of the genome in an unbiased manner without a need for a priori phenotypic information, which may subsequently be connected to a phenotype/selective pressure.

Continuing, Séverine Vuilleumier presented processes driving the fate of a novel allele in structured populations. She emphasized that understanding gain or loss of allele is a longstanding research topic, dating back to the beginning of population genetics (Fisher 1922; Wright 1931). This early interest reflects its importance for numerous related critical and practical issues that aim (i) to control the invasion of new mutants (e.g. pathogen control and drug/vaccine or pesticide resistance); (ii) to maintain genetic diversity (e.g. conservation genetics) or (iii) to understand species evolution and species adaptation to novel environments (Parmesan 2006).

She demonstrated that considering environmental spatial heterogeneity strongly impacts fixation probability of a new locally adapted allele, that is, the probability that a selected mutant increases to a significant frequency and invades populations. She presented the three steps that commonly describe the fate of a new mutant allele (Wright 1931; Fisher 1922): (i) a new mutant allele appears in a population; (ii) the mutant allele segregates and undergoes stochastic processes; and (iii) the mutant allele is ultimately fixed or lost. In a structured and heterogeneous population, different strengths of selection, genetic drift and migration will determine whether the mutant is lost in the stochastic phase or spread throughout the entire population. It has been shown that when migration is strong, spatial heterogeneity in selection has a small effect on the fate of a beneficial allele, and the related fixation probability is close to what is predicted by Kimura's panmictic formula. However, when migration is weak, a new selected allele is more likely to be fixed in heterogeneous environments (Gavrilets & Gibson 2002; Vuilleumier *et al.* 2008). Those predictions hold in systems where population sizes are relatively homogeneous and migration follows an island model.

Then, Vuilleumier showed that accounting for heterogeneity in population size and selective patterns – a common

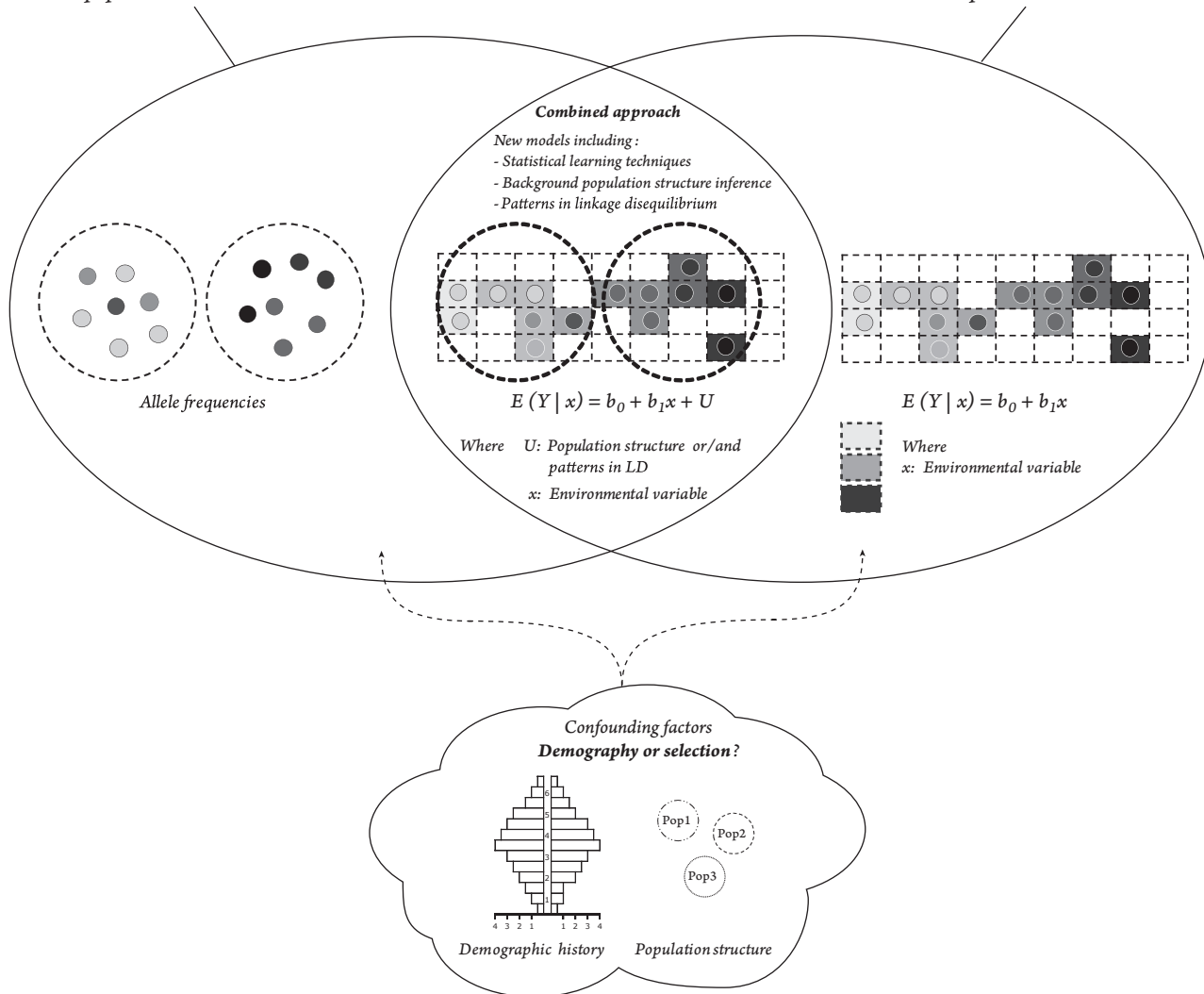
empirical situation – considerably complicates this theoretical description. For example, accounting for spatial variation in carrying capacity makes more difficult the description of the effective migration pattern among populations of different sizes. She presented three illustrative cases from Vuilleumier *et al.* (2010). First, it can be assumed that in a structured population, each individual has the same probability to emigrate from and immigrate to a population (source-sink migration). Second, settlement can be limited by competition in such a way that each individual would have the same probability to emigrate from population, but immigration into population is constrained by habitat availability. Third, it can be considered following the ideal free distributions (Fretwell & Lucas 1970; McPeck & Holt 1992) that among populations, emigration and immigration are balanced.

Then, on the basis of these migration models, Séverine Vuilleumier investigated how assumptions about isolation by distance, carrying capacity, productivity, extinction and recolonization dynamics translate into effective gene flow and into the fixation probability of selected alleles. Accounting for heterogeneity reveals that migration and structure can severely affect the probability that a locally adapted allele will succeed in settling locally and eventually invade a metapopulation. The main lesson learned from the migration models is that selected alleles can spread to populations where selection is not acting and become fixed, and that signatures of selection can be unrelated to the spatial distribution of the selective pressure (Vuilleumier *et al.* 2010).

The importance of spatial population structure has also been raised by Schoville in describing the functioning of environmental correlation methods. Based on the individual sampling, these approaches identify loci that have strong correlations with environmental variables (Joost *et al.* 2007). The main hypothesis is that deviations from a null distribution can be attributed to selection, rather than other factors that influence background patterns of genetic variation. Recent simulation studies (De Mita *et al.* 2013) have shown that these methods have increased power to detect selection along environmental gradients compared with population genetics outlier-detection approaches. A number of case studies, including Eckert *et al.*'s (2010) analysis of the loblolly pine, show that environmental correlation methods often lead to the identification of ecologically relevant loci that are not detected by outlier-detection. But as is the case with outlier-detection methods, the existence of population structure will also lead to allele frequency changes that are unrelated to selection (Holderegger *et al.* 2008). Several new models have been proposed to account for population structure, illustrating the intersection of theoretical population genetics and landscape genomics. One approach is simply to estimate the null distribution used in statistical testing from a separate control data set (Hancock *et al.* 2008) to account for both the demographic history and population structure, so that the probability ( $p$  value) at which the null hypothesis is rejected is properly adjusted. Alternatively, several methods have pro-

**Theoretical Population Genetics** seeks to explain why allele frequencies differ between populations

**Landscape Genomics** attempts to explain the pattern of spatial genetic variation in relation to local adaptation to the environment



**Fig. 1** Detecting adaptive change requires further integration of theoretical population genetics (on the left of the Venn diagram) and landscape genomics (on the right) to propose an approach that considers both genomic architecture – not restricted to allele frequencies – and environmental variation. The two disciplines lack the ability to identify patterns that distinguish selective from demographic effects. But patterns in linkage disequilibrium (LD) are promising in this regard and may be combined with correlative approaches to improve models including background population structure inference. In the landscape genomics equation,  $E(Y|x)$  is read ‘the expected value of  $Y$ , given the value  $x$ ’,  $b_0$  is the constant at the origin,  $b_1$  is the regression coefficient multiplied by some value of the environmental variable ( $x$ ) and  $U$  is an additional parameter able to take into account population structure or patterns in linkage disequilibrium.

posed incorporating an estimate of the covariance matrix among individuals in the regression model, using geographical coordinates only to consider spatial autocorrelation between individuals collected at the same sampling location (Poncet *et al.* 2010), Moran’s eigenvector maps to incorporate the effect of unaccounted environmental variables (Manel *et al.* 2010) or genetic structure (Hancock *et al.* 2008; Coop *et al.* 2010; see Table 1). In the latter approaches, a separate control data set is necessary to esti-

mate genetic relatedness. However, in practice, the same data are often used to estimate the covariance matrix and genetic adaptation, and as a result, these tests suffer from circularity. This has the potential to reduce the statistical power in these tests, particularly if some part of the estimated population structure is adaptive differentiation. To address this problem, Schoville presented a latent factor mixed model (LFMM) approach that simultaneously fits a model of population structure and environmental effects

(Frichot *et al.* 2013). Population structure is modelled as  $K$  latent factors (independent, linear combinations of the genetic data estimated from its joint distribution), while the environmental covariates are modelled as fixed effects. The authors compared the performance of this approach with the ones mentioned above and found that the application of the LFMM approach resulted in both a low false positive rate and low false negative rate, suggesting that it had better power to detect environmental correlations in the presence of population structure (see also De Mita *et al.* 2013 on the latter point).

Schoville concluded with a list of future directions necessary to improve the environmental correlation approach, among which the improvement in measurements of environmental variation (i.e. high resolution climatic data), the use of novel variables (e.g. species interactions, pollutants), the use of statistical methods to extract (e.g. principal component analysis) or select (e.g. decision trees) environmental variables and the implementation of more rigorous experimental designs. To complete the list, Kevin Leempoel interestingly proposed a modelling avenue worth investigating. He presented geographically weighted regressions (GWR) that allow regression coefficients to vary over space to study local behaviours (Brunsdon *et al.* 1996; Nakaya *et al.* 2005). A model is built for each geo-referenced individual sampled, where the interactions between samples are inversely proportional to their pairwise distance. The weighting scheme consists of a decreasing monotonic function of the distance and bandwidth over the typical range of interaction strengths. The significance of each local coefficient is assessed by a  $t$ -test, and the spatial variability of a coefficient is evaluated by fitting an alternate model where the coefficient of interest is fixed, while the others may vary. The main focus of GWR is to locally identify where associations between loci and environmental variables are the most significant, whereas global correlative approaches are only able to process general models over the whole study areas. This statistical tool may be invaluable when allele dispersal is limited across geographical space (isolation by distance) to identify local regimes where associations are significant, while it would not be useful to process regressions for long-distance dispersal organisms.

Even if landscape genomics offers a promising alternative to outlier-detection methods (e.g. multiple scales investigations, fast processing capacities, integration of spatial heterogeneity), the models make implicit assumptions that should be carefully considered. Indeed, the functional relationship between the geographical distribution of alleles and the environmental variable is assumed to be constant, and this might not be the case if the environment has changed over time or if there are genetic background effects (i.e. epistatic effects). In addition, the correlative approaches assume that selection has had enough time to create a functional relationship between the allele distributions and the environmental variable. Recent environmental changes, such as habitat fragmentation, degradation or climate change, may lead to genetic adaptation in a few generations, but only if selection is relatively strong (Kawecki & Ebert 2004). In situations

where selection is strong, both population genetics and landscape genomics offer powerful statistical tests, up to the point where selection leads to complete fixation of selected alleles. Currently, only population genetics approaches are designed to detect complete selective sweeps.

This overview of theoretical population genetics' and landscape genomics' strengths and drawbacks highlights the need for further integration of these disciplines (see Fig. 1). Indeed, decisive advantages would result in a combined approach able to capitalize mainly on the robust theoretical framework of theoretical population genetics, on the incorporation of the effect of landscape spatial heterogeneity at multiple scales and on fast computation of large genomic data sets. Interestingly, analytical frameworks recently implemented and simultaneously based on mutational frequencies, ecological modelling and statistical learning techniques appear effective (Frichot *et al.* 2013; Guillot 2012). The correlative framework is thus probably flexible enough to move a step forward and to integrate recent statistical developments from population genetics, thus taking advantage of progress towards differentiating selection from demography.

## References

- Andolfatto P (2007) Hitchhiking effects of recurrent beneficial amino acid substitutions in the *Drosophila melanogaster* genome. *Genome Research*, **17**, 1755–1762.
- Andolfatto P (2008) Controlling type-I error of the McDonald–Kreitman test in genomewide scans for selection on noncoding DNA. *Genetics*, **180**, 1767–1771.
- Bazin E, Dawson KJ, Beaumont MA (2010) Likelihood-free inference of population structure and local adaptation in a Bayesian hierarchical model. *Genetics*, **185**, 587–602.
- Beaumont MA (2010) Approximate Bayesian computation in evolution and ecology. *Annual Review of Ecology, Evolution, and Systematics*, **41**, 379–406.
- Beaumont MA, Balding DJ (2004) Identifying adaptive genetic divergence among populations from genome scans. *Molecular Ecology*, **13**, 969–980.
- Brunsdon C, Fotheringham AS, Charlton ME (1996) Geographically weighted regression: a method for exploring spatial nonstationarity. *Geographical Analysis*, **28**, 281–298.
- Coop G, Witonsky D, Di Rienzo A, Pritchard JK (2010) Using environmental correlations to identify loci underlying local adaptation. *Genetics*, **185**, 1411–1423.
- Crisci JL, Poh Y-P, Bean A, Simkin A, Jensen JD (2012) Recent progress in polymorphism-based population genetic inference. *Journal of Heredity*, **103**, 287–296.
- De Mita S, Thuillet A-C, Gay L, *et al.* (2013) Detecting selection along environmental gradients: analysis of eight methods and their effectiveness for outbreeding and selfing populations. *Molecular Ecology*, **22**, 1383–1399.
- Eckert AJ, van Heerwaarden J, Wegrzyn JL, *et al.* (2010) Patterns of population structure and environmental associations to aridity across the range of loblolly pine (*Pinus taeda* L., Pinaceae). *Genetics*, **185**, 969–982.
- Eyre-Walker A, Keightley PD (2009) Estimating the rate of adaptive molecular evolution in the presence of slightly deleterious mutations and population size change. *Molecular Biology and Evolution*, **26**, 2097–2108.

- Fisher RA (1922) On the dominance ratio. *Proceedings of the Royal Society of Edinburgh*, **42**, 321–341.
- Fretwell S, Lucas H (1970) On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical Development. *Acta Biotheoretica*, **19**, 16–36.
- Frichot E, Schoville S, Bouchard G, François O (2013) Testing for associations between loci and environmental gradients using latent factor mixed models. *Molecular Biology and Evolution* doi:10.1093/molbev/mst063.
- Gavrilets S, Gibson N (2002) Fixation probabilities in a spatially heterogeneous environment. *Population Ecology*, **44**, 51–58.
- Guillot G (2012) Detection of correlation between genotypes and environmental variables. A fast computational approach for genome-wide studies. In: *eprint arXiv:1206.0889*.
- Günther T, Coop G (2012) Robust identification of local adaptation from allele frequencies. In: *eprint arXiv: 1209.3029*.
- Hancock AM, Witonsky DB, Gordon AS, *et al.* (2008) Adaptations to climate in candidate genes for common metabolic disorders. *PLoS Genetics*, **4**, e32.
- Hernandez RD, Kelley JL, Elyashiv E, *et al.* (2011) Classic selective sweeps were rare in recent human evolution. *Science*, **331**, 920–924.
- Holderegger R, Herrmann D, Poncet B, *et al.* (2008) Land ahead: using genome scans to identify molecular markers of adaptive relevance. *Plant Ecology & Diversity*, **1**, 273–283.
- Hudson RR, Kreitman M, Aguadé M (1987) A test of neutral molecular evolution based on nucleotide data. *Genetics*, **116**, 153–159.
- Jensen JD, Thornton KR, Bustamante CD, Aquadro CF (2007) On the utility of linkage disequilibrium as a statistic for identifying targets of positive selection in nonequilibrium populations. *Genetics*, **176**, 2371–2379.
- Jensen JD, Thornton KR, Andolfatto P (2008) An approximate Bayesian estimator suggests strong, recurrent selective sweeps in *Drosophila*. *PLoS Genetics*, **4**, e1000198.
- Jones DA, Wakeley J (2008) The Influence of gene conversion on linkage disequilibrium around a selective sweep. *Genetics*, **180**, 1251–1259.
- Joost S, Bonin A, Bruford MW, *et al.* (2007) A spatial analysis method (SAM) to detect candidate loci for selection: towards a landscape genomics approach to adaptation. *Molecular Ecology*, **16**, 3955–3969.
- Kawecki TJ, Ebert D (2004) Conceptual issues in local adaptation. *Ecology Letters*, **7**, 1225–1241.
- Li HP, Stephan W (2006) Inferring the demographic history and rate of adaptive substitution in *Drosophila*. *Plos Genetics*, **2**, 1580–1589.
- Macpherson JM, Sella G, Davis JC, Petrov DA (2007) Genome-wide spatial correspondence between nonsynonymous divergence and neutral polymorphism reveals extensive adaptation in *Drosophila*. *Genetics*, **177**, 2083–2099.
- Manel S, Schwartz M, Luikart G, Taberlet P (2003) Landscape genetics: combining landscape ecology and population genetics. *Trends in Ecology & Evolution*, **18**, 189–197.
- Manel S, Poncet BN, Legendre P, Gugerli F, Holderegger R (2010) Common factors drive adaptive genetic variation at different spatial scales in *Arabis alpina*. *Molecular Ecology*, **19**, 3824–3835.
- McDonald JH, Kreitman M (1991) Adaptive protein evolution at the Adh locus in *Drosophila*. *Nature*, **351**, 652–654.
- McPeck MA, Holt RD (1992) The evolution of dispersal in spatially and temporally varying environments. *American Naturalist*, **140**, 1010–1027.
- Nakaya T, Fotheringham AS, Brunson C, Charlton M (2005) Geographically weighted Poisson regression for disease association mapping. *Statistics in Medicine*, **24**, 2695–2717.
- Nielsen R (2005) Molecular signatures of natural selection. *Annual Review of Genetics*, **39**, 197–218.
- Parmesan C (2006) Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics*, **37**, 637–669.
- Pavlidis P, Jensen JD, Stephan W (2010) Searching for footprints of positive selection in whole-genome SNP data from non-equilibrium populations. *Genetics*, **185**, 907–922.
- Pavlidis P, Jensen JD, Stephan W, Stamatakis A (2012) A critical assessment of storytelling: gene ontology categories and the importance of validating genomic scans. *Molecular Biology and Evolution*, **29**, 3237–3248.
- Poncet BN, Herrmann D, Gugerli F, *et al.* (2010) Tracking genes of ecological relevance using a genome scan in two independent regional population samples of *Arabis alpina*. *Molecular Ecology*, **19**, 2896–2907.
- Schoville SD, Bonin A, François O, *et al.* (2012) Adaptive genetic variation on the landscape: methods and cases. *Annual Review of Ecology, Evolution, and Systematics*, **43**, 23–43.
- Stephan W, Song YS, Langley CH (2006) The hitchhiking effect on linkage disequilibrium between linked neutral loci. *Genetics*, **172**, 2647–2663.
- Thornton K, Andolfatto P (2006) Approximate Bayesian inference reveals evidence for a recent, severe bottleneck in a Netherlands population of *Drosophila melanogaster*. *Genetics*, **172**, 1607–1619.
- Thornton KR, Jensen JD, Becquet C, Andolfatto P (2007) Progress and prospects in mapping recent selection in the genome. *Heredity*, **98**, 340–348.
- Vuilleumier S, Yearsley JM, Perrin N (2008) The fixation of locally beneficial alleles in a metapopulation. *Genetics*, **178**, 467–475.
- Vuilleumier S, Goudet J, Perrin N (2010) Evolution in heterogeneous populations From migration models to fixation probabilities. *Theoretical Population Biology*, **78**, 250–258.
- Wegmann D, Leuenberger C, Excoffier L (2009) Efficient approximate Bayesian computation coupled with Markov chain Monte Carlo without likelihood. *Genetics*, **182**, 1207–1218.
- Williamson SH, Hernandez R, Fedel-Alon A, *et al.* (2005) Simultaneous inference of selection and population growth from patterns of variation in the human genome. *Proceedings of the National Academy of Sciences, USA*, **102**, 7882–7887.
- Wright S (1931) Evolution in Mendelian populations. *Genetics*, **16**, 97–159.
- Yang ZH, Nielsen R (2002) Codon-substitution models for detecting molecular adaptation at individual sites along specific lineages. *Molecular Biology and Evolution*, **19**, 908–917.

---

S.J. and S.M. conceived and designed the research project. S.J., K.L., S.S.T., I.W., C.M., and J.R. organized the workshop. S.J., S.V., J.D.J. and S.S.C. wrote the paper. I.W. and S.J. elaborated the figure and the table. All co-authors participated in the conception of the paper, revised it critically and approved the final version to be published.

---

doi: 10.1111/mec.12352