MOLECULAR ECOLOGY
RESOURCES
Molecular Ecology Resources (2017) **17**, 356–361 doi: 10.1111/1755-0998.12649

COMMENT

RADseq provides unprecedented insights into molecular ecology and evolutionary genetics: comment on Breaking RAD by Lowry *et al.* (2016)

GARRETT J. MCKINNEY,* WESLEY A. LARSON,† LISA W. SEEB* and JAMES E. SEEB*
*School of Aquatic and Fishery Sciences, University of Washington, 1122 NE Boat Street, Box 355020, Seattle WA, 98195-5020, USA, †U.S. Geological Survey, Wisconsin Cooperative Fishery Research Unit, College of Natural Resources, University of Wisconsin-Stevens Point, 800 Reserve St., Stevens Point WI, 54481, USA

Abstract

In their recently corrected manuscript, "Breaking RAD: An evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation", Lowry *et al.* argue that genome scans using RADseq will miss many loci under selection due to a combination of sparse marker density and low levels of linkage disequilibrium in most species. We agree that marker density and levels of LD are important considerations when designing a RADseq study; however, we dispute that RAD-based genome scans are as prone to failure as Lowry *et al.* suggest. Their arguments ignore the flexible nature of RADseq; the availability of different restriction enzymes and capacity for combining restriction enzymes ensures that a well-designed study should be able to generate enough markers for efficient genome coverage. We further believe that simplifying assumptions about linkage disequilibrium in their simulations are invalid in many species. Finally, it is important to note that the alternative methods proposed by Lowry *et al.* have limitations equal to or greater than RADseq. The wealth of studies with positive impactful findings that have used RAD genome scans instead supports the argument that properly conducted RAD genome scans are an effective method for gaining insight into ecology and evolution, particularly for non-model organisms and those with large or complex genomes.

Keywords: RADseq, genome scan, linkage disequilibrium, adaptation

Received 22 September 2016; revision received 20 December 2016; accepted 20 December 2016

In their recently corrected manuscript, 'Breaking RAD: An evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation', Lowry *et al.* (2016) detail several concerns with the use of RADseq data for genome scans. Lowry *et al.* (2016) argue that informative RAD tags are usually distant from each other (100 Kbs to MBs), and linkage disequilibrium (LD) in populations can be restricted to small enough haplotype blocks that informative loci may not be in strong enough LD with causal SNPs to be useful in genome scans. We agree that the relatively low density of RAD loci will sometimes result in the failure to detect adaptive loci and that LD in some species is too low to efficiently conduct genome scans with RADseq. However, we dispute that RAD-based genome scans are as prone to failure as Lowry *et al.* (2016) suggest and believe that simplifying assumptions in their simulations are invalid in many species. Further, it is important to note that the alternative methods proposed by Lowry *et al.* (2016) have limitations equal to or greater than RADseq. Despite the authors' references to 'Breaking Bad', RADseq genome scans are neither broken nor a half-measure.

Correspondence: Garrett J. McKinney, Fax: 206-685-7471; E-mail: gjmckinn@uw.edu

Genomic investigations using RADseq, including genome scans, have become increasingly common since Hohenlohe *et al.* (2010) demonstrated the potential for the identification of genes and genomic regions of adaptive significance. Lowry *et al.* (2016) generate a table of citations (Table S1) as a basis for their demonstration of inadequacy of RADseq. Rather than examining the findings of these studies, Lowry *et al.* (2016) count the number of loci used in each study to estimate the genome coverage and report that these studies 'miss the vast majority of loci underlying adaptation'. What Lowry *et al.* fail to do is recognize that nearly all of these studies that conducted genome scans have positive findings. For example, Laporte *et al.* (2016) identified within-generation polygenic selection in response to anthropogenic pollutants; Cammen *et al.* (2015) identified genomic

Table 1 Example references from Lowry *et al.* (2016) Table S1 of genome scans where RADseq-based scans revealed novel insights into signatures of adaptation, defying the expectations promoted by the model of Lowry *et al.* (2016)

Organism	Major finding	Reference
Eel	Over 140 covarying loci discriminated individuals from control vs. polluted environments implicating role of sterol syntheses genes	Laporte *et al.* (2016)
Bottlenose Dolphin	Identified genomic regions surrounding polymorphic RAD loci associated with survival following harmful algal blooms	Cammen *et al.* (2015
Chinook Salmon	Genomic divergence at 566 putatively adaptive loci driven by environmental variables important to migratory cold water fish	Hecht *et al.* (2015)
Bonnethead Shark	Alleles at 30 outlier loci associated with latitude-associated selection	Portnoy *et al.* (2015)
Surgeonfish sp.	Divergent selection at 59 loci, one of which mapped to the opsin Rh2 gene, a component of the cone receptor system (light sensitivity)	Gaither *et al.* (2015)
Chinook Salmon	Map-based identification of outlier loci and genomic regions associated with local adaptation	Brieuc *et al.* (2015)
Guppy	Indications of recent, convergent directional selection in two introduced populations	Fraser *et al.* (2015)
Alcolapia Cichlid sp.	Identification of outlier SNPs under diversifying selection indicating that the cichlid radiation is an example of undergoing adaptive divergence	Ford *et al.* (2015)
Periwinkle Snail	Indications that parallel phenotypic divergence is not matched by shared genomic divergence despite a high probability of gene flow	Ravinet *et al.* (2016)
Damselfly	Shows how a strategic sampling design and integration of genomic, phenotypic and environmental data can disentangle neutral and adaptive processes	Swaegers *et al.* (2015)
Nuthatches	Found 79 loci putatively under selection; some were highly associated with climate extremes, disentangling signals of isolation by environment from IBD	Manthey & Moyle (2015)
Stickleback	Inferred demography of selective sweeps and the maintenance of adaptive divergence in the face of gene flow	Roesti *et al.* (2015)
Eel	Outlier analysis and functional annotation implicated the adaptive importance of signal transduction pathways	Pujolar *et al.* (2015)
Herring	Outlier analyses revealed 100s of directionally selected loci with variability associated with either salinity, temperature or both	Guo *et al.* (2016)

Some of the papers in S1 were not focused genome scans but rather were more general studies of population structure or genetic diversity (e.g. Jezkova *et al.* 2015; Saenz-Agudelo *et al.* 2015). Many studies listed in S1 contributed novel insights into molecular ecology and evolutionary genetics that could not have been easily obtained by approaches suggested in Lowry *et al.* (2016).

regions and candidate genes associated with ability to survive harmful algal blooms; and Brieuc *et al.* (2015) identified a set of 33 loci that together explained 79.2% of variance in run timing in Chinook salmon, an ecologically important trait. We summarize important findings of the studies examined by Lowry *et al.* (2016) (Table 1).

In more recent manuscripts, Epstein *et al.* (2016) identified genomic regions showing rapid adaptive response to transmissible cancer in Tasmanian devils using RADseq, Larson *et al.* (2016) identify genomic islands of divergence and a mutation that is likely involved in ecotypic differentiation in sockeye salmon, and Chaves *et al.* (2016) identified multiple genomic regions associated with adaptive radiation in Darwin's finches. Why do Lowry *et al.* (2016) discount the utility of RADseq in spite of empirical evidence to the contrary?

First, a major assumption of the simulation model adopted by Lowry *et al.* (2016) is that the recombination rate and therefore linkage disequilibrium are constant across a given genome. Lowry *et al.* (2016) further argue that any violation of this assumption will decrease the effectiveness of RAD genome scans. This simplifying assumption ignores the fact the power to detect loci associated with adaptive variation will increase in regions of the genome with reduced recombination rates. In zebra finch, 'recombination deserts' have been observed to extend for over 100 Mb on multiple chromosomes and 90% of the total recombination in zebra finch is concentrated in only 23% of the genome (Backstrom *et al.* 2010). Recombination deserts spanning 50 Mb have been observed in cats (Li *et al.* 2016); 34 Mb recombination deserts have been found in pigs (Ai *et al.* 2015; Li *et al.* 2016). Finally, reduced recombination has been observed across several Mb sections of multiple chromosomes in collared flycatchers (Burri *et al.* 2015) and mouse (Shifman *et al.* 2006). In cases such as these, relatively few loci can effectively interrogate large segments of chromosomes due to extended strong LD. In addition, when adaptive loci occur in regions of low recombination, linked selection can result in extended regions of divergence between populations (Burri *et al.* 2015).

LD can also increase around selected sites when selection maintains differences in allele frequencies between populations but there is still gene flow (Via & West 2008; Via 2012). This can result in genomic islands of

divergence (Nosil *et al.* 2009) that can span nearly half a chromosome (e.g. Miller *et al.* 2012; Andrew & Rieseberg 2013; Hemmer-Hansen *et al.* 2013) but more typically cover between 10 and 500 kb (Hohenlohe *et al.* 2010; Soria-Carrasco *et al.* 2014). Many studies have used RAD sequencing to discover islands of divergence including Andrew & Rieseberg (2013) who found large islands of divergence between sunflower ecotypes, Gagnaire *et al.* (2013) who discovered parallel islands of divergence in ecotypes of dwarf and normal whitefish, and Miller *et al.* (2012) who found a large genomic region linked to growth in rainbow trout. The genomic region discovered in Miller *et al.* (2012) was also found to be involved in parallel adaptation of resident and migratory forms of rainbow trout (Pearse *et al.* 2014). These studies as well as many others demonstrate that differentiation across the genome is highly heterogeneous and that RAD sequencing is a useful tool for discovering genomic regions linked to adaptively important traits.

As part of their argument that RAD genome scans are destined to fail, Lowry *et al.* (2016) present LD estimates for a variety of species in their Table 1. They cite this table as evidence that 'for most organisms, LD is far less than 100 kb'. Of course, it is unrealistic to expect that the 30 species in this table broadly represent all species and all populations of interest for scientific study, and Lowry *et al.* (2016) fail to point out that six of the species in their Table 1 (20%) have LD estimates either equal or much greater than 100 Kb, three of which even had LD estimates of 1 Mb or greater [Siberian Jays (Li & Merila 2010), bighorn sheep (Miller *et al.* 2011), chickpea (Kujur *et al.* 2015)]. Lowry *et al.* (2016) divide the genome into haplotype blocks based on the length of LD, assuming all markers within a haplotype block are linked, and plot the expected genome coverage for species with varying length of LD, number of markers, and genome size, in their Figure 3. The maximum haplotype block size plotted is 50 Kb, despite the presence of several species in their Table 1 with LD estimates equal or greater to 100 Kb. Even with 50 Kb haplotype blocks, Lowry *et al.* (2016) show that species with less than a 500-Mb genome would have 100% coverage with only a few thousand markers and species with a 1-Gb genome would have 75% coverage with 15 000 markers.

We used their approach and their R code to plot expected genome coverage for species with haplotype blocks ranging from 100 Kb to 1 Mb (Figure 1). With haplotype blocks of 100 Kb, 100% coverage was estimated for 1-Gb genomes with only 10 000 markers. This suggests that the expected coverage for RAD genome scans increases rapidly with LD. According to these same calculations, all of the species with LD of 100 Kb or

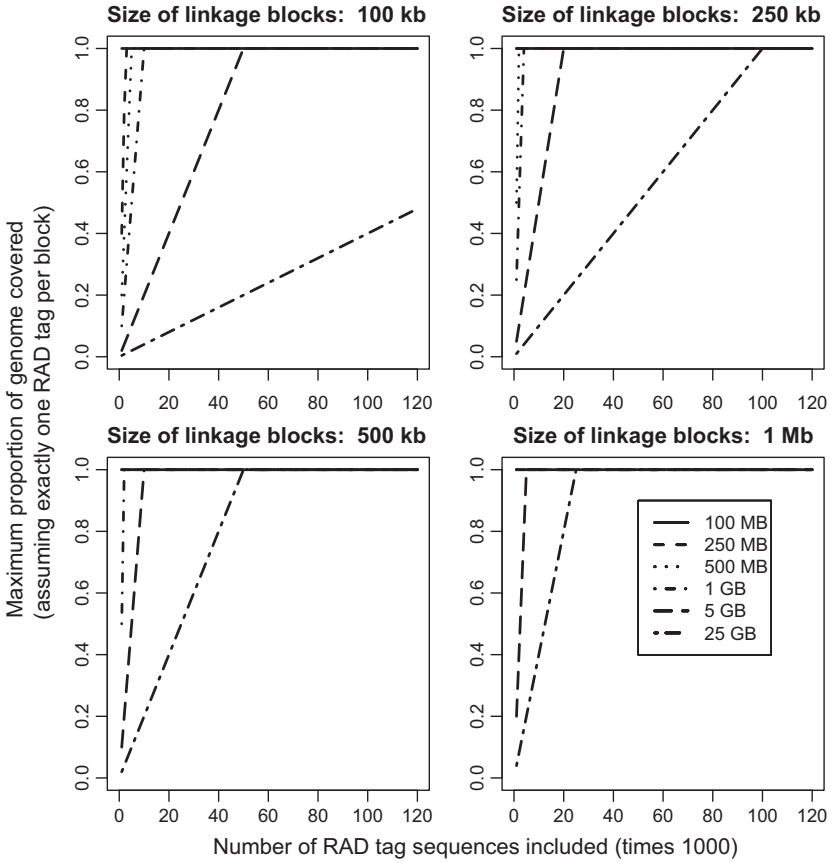

Fig. 1 The expected proportion of the genome sampled (*y*-axis) based on number of RAD markers (*x*-axis) and genome sizes (line types) for different average lengths of linkage disequilibrium. Organisms with LD on the scale of 100 Kb include wild yeast (Tsai *et al.* 2008); 250 Kb include some *Brassica* populations (Wu *et al.* 2016); 500 Kb include collared Flycatcher (400 Kb, Backstrom *et al.* 2006) and wild soya bean (Wang *et al.* 2015); 1 Mb or greater include chickpea (Kujur *et al.* 2015), Siberian jays (Li & Merila 2010) and bighorn sheep (Miller *et al.* 2011).

greater listed in Table 1 of Lowry *et al.* would have 100% genome coverage with only 5,000 RAD markers based on known or estimated genome sizes for these species (Fig. 1). With the exception of wild yeast, the genomes of these species ranged from 738 Mb to 3.5 Gb which would yield 11 000 to 53 000 RAD cut sites with an 8-base cutter or 180 000 to 854 000 cut sites with a 6-base cutter. Wild yeast has a genome size of only 12 Mb, but 47 000 RAD cut sites could still be generated using a 4-base cutter. The flexibility afforded by different restriction enzymes ensures that a well-designed study should be able to generate enough markers to efficiently cover a large portion of the genome in many species. An example of this is the recent study by Chaves *et al.* (2016) who generated 32 569 SNPs from a ~1-Gb genome using a combination of EcoRI and MseI enzymes, a 6-bp cutter and 4-bp cutter. LD was shown to decay at 20–40 Kb in this species; the 32 000 SNPs in this study would yield approximately 64% coverage assuming a haplotype block size of 20 Kb. Complete genome coverage would be achieved with a haplotype block size of ~31 Kb which is below the upper estimate of 40 Kb. Regardless of the actual genome coverage, Chaves *et al.* (2016) were able to identify several regions of the genome associated with variation in beak and body size in Darwin's finches and provide candidate genes associated with these SNPs.

Lowry *et al.* (2016) also gloss over the fact that LD can vary widely between populations within a species due to the unique evolutionary history of each population. For example, average LD in pigs has been demonstrated to vary from <10 Kb in Chinese breeds up to 400Kb in European breeds (Amaral *et al.* 2008). Species can also exhibit extensive LD in local populations even when LD is relatively low at the species level (Nordborg *et al.* 2002; Kim *et al.* 2007). As a result, species level estimates of LD may be downwardly biased relative to population level LD, as was noted by Kim *et al.* (2007) who provided the LD estimate for *Arabidopsis* that Lowry *et al.* (2016) used in their Table 1. An extreme example of this can be found in dogs where LD can be several Mb within breeds but is only tens of Kb when measured across the species (Lindblad-Toh *et al.* 2005). We recognize that the level of LD in domesticated species is not likely representative of wild species, but the variation in LD in domestic species reflects similar processes that lead to variation in LD across populations of wild species: selection, population structure, historical N_e, and bottlenecks or founder events. Unfortunately, there have been relatively few studies characterizing LD in nonmodel species (Li & Merila 2010, 2011). Increased research on LD in wild species would be beneficial for many reasons, but until actual LD values are known, we should not assume RAD genome scans are destined to fail based on estimates that may not be representative of many species or populations.

Lowry *et al.* (2016) suggest four methods as alternatives to RADseq for genome scans: transcriptome sequencing, exon capture, whole genome sequencing and pool-seq. Each of these methods can be effective for particular study designs; however, each of these methods also has equivalent or greater limitations than RADseq and most require greater initial input or cost. The many limitations of transcriptome sequencing (e.g. tissue or condition-specific expression, failure to interrogate regulatory regions and lack of data for the null model) are detailed by Lowry *et al.* (2016). Another alternative, exon capture, has many of the same limitations of transcriptome sequencing but also requires a priori knowledge of coding regions and is best implemented with a reference genome or at the very least, a transcriptome. Significant costs and developmental time associated with custom design of capture arrays are not mentioned by the authors. The third alternative, whole genome sequencing for genome scans, provides the strongest evidence for detecting local adaptation, and the existence of a whole genome sequence can be used as a resource for alignment and annotation of reduced representation (RRL) approaches. However, cost and complexity of genomes are a consideration in undertaking a whole genome sequencing project, and the number of individuals that can be sequenced may be very low for a genome scan. Finally, given a whole genome sequence, Lowry *et al.* (2016) suggest that pooled sequencing (pool-seq) is a cost-effective alternative to individual whole genome sequencing for genome scans. Pool-seq provides accurate allele frequency estimation from a large number of individuals in a pool, but small pool sizes (<40–50 individuals) will yield suboptimal results (Schlotterer *et al.* 2014). More importantly, allele frequencies in pooled samples can be skewed due to uneven representation of individuals in the pool (Anderson *et al.* 2014); this is particularly difficult to detect in the absence of individual data for comparison. Further, few chromosomes are sequenced more than once; thus, pool-seq is not well suited for linkage information or complex genomes with copy number variation. While pool-seq may be an appropriate choice for studies with large sample sizes, multiple populations, and well-established genomic resources, the efficiency and ability to obtain individual genotypes from RAD approaches suggest that RAD will continue to be a better choice for many nonmodel organisms.

In conclusion, we would like to answer one of the questions from Lowry *et al.* (2016); how bad is RAD for genomic scans of adaptation? If you examine the wealth of studies that have used RAD genome scans and consider their positive findings, the answer is not bad at all. It is nearly impossible to identify all loci subjected to selection regardless of methodology; this is not a particular failure of RADseq. We argue that RAD sequencing

has significantly advanced the field of molecular ecology by facilitating genotyping of thousands of markers without existing genomic resources. RAD sequencing, like methods before it, does not represent an endpoint. However, for most scientists, especially those working on nonmodel organisms with limited genomic resources, RADseq is the most appropriate and suitable technique available at this time for investigating population genomics and natural selection.

Acknowledgements

We thank Shawn Narum and Angus Davison for their comments during review. We also thank Marty Kardos, Gordon Luikart and Julian Catchen for their comments on draft versions of this manuscript. Any use of trade, product, or company name is for descriptive purposes only and does not imply endorsement by the U.S. Government.

References

Ai H, Fang X, Yang B *et al.* (2015) Adaptation and possible ancient interspecies introgression in pigs identified by whole-genome sequencing. *Nature Genetics*, **47**, 217–225.

Amaral AJ, Megens HJ, Crooijmans RP, Heuven HC, Groenen MA (2008) Linkage disequilibrium decay and haplotype block structure in the pig. *Genetics*, **179**, 569–579.

Anderson EC, Skaug HJ, Barshis DJ (2014) Next-generation sequencing for molecular ecology: a caveat regarding pooled samples. *Molecular Ecology*, **23**, 502–512.

Andrew RL, Rieseberg LH (2013) Divergence is focused on few genomic regions early in speciation: incipient speciation of sunflower ecotypes. *Evolution*, **67**, 2468–2482.

Backstrom N, Qvarnstrom A, Gustafsson L, Ellegren H (2006) Levels of linkage disequilibrium in a wild bird population. *Biology Letters*, **2**, 435–438.

Backstrom N, Forstmeier W, Schielzeth H *et al.* (2010) The recombination landscape of the zebra finch *Taeniopygia guttata* genome. *Genome Research*, **20**, 485–495.

Brieuc MSO, Ono K, Drinan DP, Naish KA (2015) Integration of random forest with population-based outlier analyses provides insight on the genomic basis and evolution of run timing in Chinook salmon (*Oncorhynchus tshawytscha*). *Molecular Ecology*, **24**, 2729–2746.

Burri R, Nater A, Kawakami T *et al.* (2015) Linked selection and recombination rate variation drive the evolution of the genomic landscape of differentiation across the speciation continuum of *Ficedula* flycatchers. *Genome Research*, **25**, 1656–1665.

Cammen KM, Schultz TF, Rosel PE, Wells RS, Read AJ (2015) Genome-wide investigation of adaptation to harmful algal blooms in common bottlenose dolphins (*Tursiops truncatus*). *Molecular Ecology*, **24**, 4697–4710.

Chaves JA, Cooper EA, Hendry AP *et al.* (2016) Genomic variation at the tips of the adaptive radiation of Darwin's finches. *Molecular Ecology*, **25**, 5282–5295.

Epstein B, Jones M, Hamede R *et al.* (2016) Rapid evolutionary response to a transmissible cancer in Tasmanian devils. *Nature Communications*, **7**, 12684, doi:10.1038/ncomms12684.

Ford AGP, Dasmahapatra KK, Ruber L, Gharbi K, Cezard T, Day J (2015) High levels of interspecific gene flow in an endemic cichlid fish adaptive radiation from an extreme lake environment. *Molecular Ecology*, **24**, 3421–3440.

Fraser BA, Kunstner A, Reznick DN, Dreyer C, Weigel D (2015) Population genomics of natural and experimental populations of guppies (*Poecilia reticulata*). *Molecular Ecology*, **24**, 389–408.

Gagnaire PA, Pavey SA, Normandeau E, Bernatchez L (2013) The genetic architecture of reproductive isolation during speciation-with-gene-flow in lake whitefish species pairs assessed by RAD sequencing. *Evolution*, **67**, 2483–2497.

Gaither MR, Bernal MA, Coleman RR *et al.* (2015) Genomic signatures of geographic isolation and natural selection in coral reef fishes. *Molecular Ecology*, **24**, 1543–1557.

Guo BC, Li Z, Merila J (2016) Population genomic evidence for adaptive differentiation in Baltic Sea herring. *Molecular Ecology*, **25**, 2833–3852.

Hecht BC, Matala AP, Hess JE, Narum SR (2015) Environmental adaptation in Chinook salmon (*Oncorhynchus tshawytscha*) throughout their North American range. *Molecular Ecology*, **24**, 5573–5595.

Hemmer-Hansen J, Nielsen EE, Therkildsen NO *et al.* (2013) A genomic island linked to ecotype divergence in Atlantic cod. *Molecular Ecology*, **22**, 2653–2667.

Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics*, **6**, e1000862.

Jezkova T, Riddle BR, Card DC, Schield DR, Eckstut ME, Castoe TA (2015) Genetic consequences of postglacial range expansion in two codistributed rodents (genus *Dipodomys*) depend on ecology and genetic locus. *Molecular Ecology*, **24**, 83–97.

Kim S, Plagnol V, Hu TT *et al.* (2007) Recombination and linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **39**, 1151–1155.

Kujur A, Bajaj D, Upadhyaya HD *et al.* (2015) Employing genome-wide SNP discovery and genotyping strategy to extrapolate the natural allelic diversity and domestication patterns in chickpea. *Frontiers in Plant Science*, **6**, 162, doi: 10.3389/fpls.2015.00162.

Laporte M, Pavey SA, Rougeux C *et al.* (2016) RAD sequencing reveals within-generation polygenic selection in response to anthropogenic organic and metal contamination in North Atlantic Eels. *Molecular Ecology*, **25**, 219–237.

Larson WA, Limborg MT, McKinney GJ, Schindler DE, Seeb JE, Seeb LW (2016) Genomic islands of divergence linked to ecotypic variation in sockeye salmon. *Molecular Ecology*, doi:10.1111/mec.13933.

Li MH, Merila J (2010) Extensive linkage disequilibrium in a wild bird population. *Heredity (Edinb)*, **104**, 600–610.

Li MH, Merila J (2011) Population differences in levels of linkage disequilibrium in the wild. *Molecular Ecology*, **20**, 2916–2928.

Li G, Hillier LW, Grahn RA *et al.* (2016) A high-resolution SNP array-based linkage map anchors a new domestic cat draft genome assembly and provides detailed patterns of recombination. *G3 (Bethesda)*, **6**, 1607–1616.

Lindblad-Toh K, Wade CM, Mikkelsen TS *et al.* (2005) Genome sequence, comparative analysis and haplotype structure of the domestic dog. *Nature*, **438**, 803–819.

Lowry DB, Hoban S, Kelley JL *et al.* (2016) Breaking RAD: an evaluation of the utility of restriction site associated DNA sequencing for genome scans of adaptation. *Molecular Ecology Resources*, doi: 10.1111/1755-0998.12635.

Manthey JD, Moyle RG (2015) Isolation by environment in White-breasted Nuthatches (*Sitta carolinensis*) of the Madrean Archipelago sky islands: a landscape genomics approach. *Molecular Ecology*, **24**, 3628–3638.

Miller JM, Poissant J, Kijas JW, Coltman DW, International Sheep Genomics C (2011) A genome-wide set of SNPs detects population substructure and long range linkage disequilibrium in wild sheep. *Molecular Ecology Resources*, **11**, 314–322.

Miller MR, Brunelli JP, Wheeler PA *et al.* (2012) A conserved haplotype controls parallel adaptation in geographically distant salmonid populations. *Molecular Ecology*, **21**, 237–249.

Nordborg M, Borevitz JO, Bergelson J *et al.* (2002) The extent of linkage disequilibrium in *Arabidopsis thaliana*. *Nature Genetics*, **30**, 190–193.

Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Molecular Ecology*, **18**, 375–402.

Pearse DE, Miller MR, Abadía-Cardoso A, Garza JC (2014) Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/rainbow trout. *Proceedings of the Royal Society B: Biological Sciences*, **281**, 20140021doi: 10.1098/rspb.2014.0012.

Portnoy DS, Puritz JB, Hollenbeck CM, Gelsleichter J, Chapman D, Gold JR (2015) Selection and sex-biased dispersal in a coastal shark: the influence of philopatry on adaptive variation. *Molecular Ecology*, **24**, 5877–5885.

Pujolar JM, Jacobsen MW, Bekkevold D *et al.* (2015) Signatures of natural selection between life cycle stages separated by metamorphosis in European eel. *BMC Genomics*, **16**, 600, doi: 10.1186/s12864-015-1754-3.

Ravinet M, Westram A, Johannesson K, Butlin R, Andre C, Panova R (2016) Shared and nonshared genomic divergence in parallel ecotypes of *Littorina saxatilis* at a local scale. *Molecular Ecology*, **25**, 287–305.

Roesti M, Kueng B, Moser D, Berner D (2015) The genomics of ecological vicariance in threespine stickleback fish. *Nature Communications*, **6**, 8767, doi: 10.1038/ncomms9767.

Saenz-Agudelo P, Dibattista JD, Piatek MJ *et al.* (2015) Seascape genetics along environmental gradients in the Arabian Peninsula: insights from ddRAD sequencing of anemonefishes. *Molecular Ecology*, **24**, 6241–6255.

Schlotterer C, Tobler R, Kofler R, Nolte V (2014) Sequencing pools of individuals–mining genome-wide polymorphism data without big funding. *Nature Reviews Genetics*, **15**, 749–763.

Shifman S, Bell JT, Copley RR *et al.* (2006) A high-resolution single nucleotide polymorphism genetic map of the mouse genome. *PLoS Biology*, **4**, e395, doi: 10.1371/journal.pbio.0040395.

Soria-Carrasco V, Gompert Z, Comeault AA *et al.* (2014) Stick insect genomes reveal natural selection's role in parallel speciation. *Science*, **344**, 738–742.

Swaegers J, Mergeay J, Van Geystelen A, Therry L, Larmuseau MHD, Stoks R (2015) Neutral and adaptive genomic signatures of rapid poleward range expansion. *Molecular Ecology*, **24**, 6163–6176.

Tsai IJ, Bensasson D, Burt A, Koufopanou V (2008) Population genomics of the wild yeast *Saccharomyces paradoxus*: Quantifying the life cycle. *Proceedings of the National Academy of Sciences of the United States of America*, **105**, 4957–4962.

Via S (2012) Divergence hitchhiking and the spread of genomic isolation during ecological speciation-with-gene-flow. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **367**, 451–460.

Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Molecular Ecology*, **17**, 4334–4345.

Wang Y, Shahid MQ, Huang H, Wang Y (2015) Nucleotide diversity patterns of three divergent soybean populations: evidences for population-dependent linkage disequilibrium and taxonomic status of *Glycine gracilis*. *Ecology and Evolution*, **5**, 3969–3978.

Wu Z, Wang B, Chen X *et al.* (2016) Evaluation of linkage disequilibrium pattern and association study on seed oil content in *Brassica napus* using ddRAD sequencing. *PLoS ONE*, **11**, e0146383, doi: 10.1371/journal.pone.0146383.

All authors contributed to analysis and writing the manuscript.
