# Trait mapping in diverse arthropods by bulked segregant analysis

Andre H. Kurlovs[1,*], Simon Snoeck[1,*], Olivia Kosterlitz[2,†],

Thomas Van Leeuwen[1,3] , and Richard M. Clark[2,4,**]

[1] Department of Plants and Crops, Faculty of Bioscience Engineering, Ghent University, Coupure links 653, 9000, Ghent, Belgium.

[2] School of Biological Sciences, University of Utah, 257 South 1400 East, Salt Lake City, Utah, 84112, USA.

[3] Institute for Biodiversity and Ecosystem Dynamics (IBED), University of Amsterdam (UvA), Science Park 904, 1908 XH, Amsterdam, The Netherlands.

[4] Center for Cell and Genome Science, University of Utah, 257 South 1400 East, Salt Lake City, Utah, 84112, USA.

*these authors contributed equally

† Present address: Department of Biology, University of Washington, Life Sciences Building 5 East, Seattle, WA, 98195, USA.

**Corresponding author: Clark, Richard M. (richard.m.clark@utah.edu). Postal address: 257 S 1400 E, Rm 201, Salt Lake City, Utah, USA 84112. Telephone number: 801-585-9722

e-mail addresses of all authors: andre.kurlovs@ugent.be, simonp.snoeck@ugent.be, oliviakosterlit@gmail.com, thomas.vanleeuwen@ugent.be, richard.m.clark@utah.edu

## Highlights

- Bulked segregant analysis (BSA) is a cross-based method for rapid trait mapping
- BSA does not require genotyping of individuals from large mapping populations
- BSA methods allow genetic mapping in non-model species with few genetic resources
- Advances in genome sequencing are facilitating adoption of BSA methods more broadly

**Abstract**

Bulked segregant analysis (BSA) is a cross-based method for genetic mapping in sexually reproducing organisms. The method's use of bulked (pooled) samples markedly reduces the genotyping effort associated with traditional linkage mapping studies. Further, it can be applied to species with life histories or physical attributes (as for micro-insects) that render genetic mapping with other methods impractical. Recent studies in both insects and mites have revealed that advanced BSA experimental designs can resolve causal loci to narrow genomic intervals, facilitating follow-up investigations. As high-quality genomes become more widely available, BSA methods are poised to become an increasingly important tool for the rapid mapping of both monogenic and polygenic traits in diverse arthropod species.

**Introduction**

Insect and mite taxa are species-rich, and exhibit staggeringly diverse life histories and habitat preferences. They are also of high importance for human welfare as vectors of major human diseases, or as herbivores that cause ~US$470 billion in crop yield loss worldwide [1]. Chemical control with insecticides and acaricides has been crucial to keep pest populations low, and has fueled the evolution of resistance [2]. Nevertheless, elucidating the genetic architecture of resistance, as well as the immense variation in other phenotypes, has been challenging. For instance, it was only in 2002 that a variant at a cytochrome P450 (CYP) gene was implicated in *Drosophila melanogaster* as conferring resistance to the infamous pesticide dichlorodiphenyltrichloroethane (DDT) [3], which had been in use since the early 1940s. This and more recent accomplishments using advanced genetic designs, including genome-wide association mapping (GWAS), have been made possible in this species by the extensive genetic and genomic resources developed over decades by the Drosophila community [4,5]. Apart from *D. melanogaster* or its congeners, however, such genetic resources are virtually non-existent in arthropods, and uncovering the molecular-genetic underpinnings of phenotypic variation can be daunting, even for monogenic traits. While traditional linkage mapping studies have identified loci for trait variation in a number of insect species (e.g., [6–8]), they require intensive genotyping of single individuals in segregating populations. Notwithstanding technical advances, genotyping hundreds of individuals remains time-consuming, tedious, and costly. In practice, this limits sample size, and hence power (the ability to detect a genotype-to-phenotype association) and mapping resolution (the size of a genomic region to which a causal variant is localized). Moreover, phenotyping single individuals may not always be readily feasible [9]. For instance, the minute size of many arthropods, including micro-insects or mite herbivores or parasites that can be only a few hundred μm in length [10,11], can make phenotyping individuals, as well as genotyping them, challenging or not possible in practice. To circumvent these obstacles, a growing number of studies have employed bulked segregant analysis (BSA) genetic

methods. In this review, we present a primer on BSA methods, highlight recent successes in both insects and mites, and discuss how barriers to more widespread adoption are being rapidly overcome.

**BSA genetic mapping: experimental design and detection of causal loci**

BSA concepts were first elaborated by plant geneticists in 1991 [12,13], with Michelmore *et al*. [13] introducing the term "BSA." Since then, BSA approaches (Figure 1) have been used in many studies in plants to identify loci underlying both monogenic and quantitative traits [14–16]. As the methods are generally applicable to sexually reproducing organisms, they have also been adopted (in various forms) in studies in the yeast *Saccharomyces cerevisiae* [17–19], the roundworm *Caenorhabditis elegans* [20], vertebrates [21,22], and major arthropod taxa [9,23–25]. In its simplest iteration, two parental strains with a contrasting phenotype are crossed to generate an F2 population (Figure 1A,B), and separate bulk DNA samples are prepared from pools of individuals that exhibit phenotypic extremes for the trait of interest. In the straightforward case of monogenic inheritance, one pool will be fixed at the causal locus and the region surrounding it (an effect of linkage), while at unlinked genomic locations, alleles from both parents will be represented (Figure 1E,F). For quantitative traits, fixation is not expected, but differences in allele frequencies between offspring pools can be used to locate quantitative trait loci (QTL) (Figure 2) [24,26–28]. Although crosses with parental strains that are inbred simplifies genotyping and downstream analyses, it is not always a requirement [29,30]. Despite the conceptual simplicity and wide-spread success of BSA studies, a caveat is that detection of epistatic relationships is confounded as information about the co-occurrence of variants in individuals is lost upon pooling [31].

As single DNA samples are prepared and genotyped from each of contrasting bulks, the laborious step of individually processing hundreds of samples is eliminated. Importantly, pooling of individuals can allow for relatively large bulk sizes, which can increase mapping resolution, as more recombination events are captured (Figure 1B). To increase mapping resolution further, segregating populations can also be propagated beyond the F2 generation [9], allowing additional recombination events to accumulate (Figure 1C). The classic BSA design can also be modified in other ways. For instance, selection can be applied over multiple generations (instead of focusing on a single time point; Figure 1D) [27,28]; this design is attractive for instances where the genetic architecture is polygenic, and for which multiple rounds of selection may be required to reveal detectable changes in allele frequencies at loci of minor effect size (see the Figure 1 legend for additional experimental considerations). The multigenerational selection component of this approach bears similarity to evolve and resequence (E&R) experimental designs that have also attracted attention, especially from Drosophila geneticists [32–34]. However, E&R studies typically use large, genetically diverse founding populations, as does the related pool-GWAS method [5,35], and not crosses starting with defined strains selected to vary markedly in phenotypes of interest, as typify BSA study designs. Nevertheless, E&R and pool-GWAS have been

shown to be powerful where requisite population resources exist, and their advantages and disadvantages have been investigated or reviewed in several recent and comprehensive works [36–39].

Regardless of the design, the ultimate goal of BSA mapping is to identify loci responsible for variation in traits of interest. In accomplishing this goal, many genotyping methods have been employed, with high-throughput, short-read sequencing (e.g., Illumina sequencing [40]) emerging as a dominant approach. By sequencing genomic DNA from pools [9,24,25,27–30], or in some cases cDNA (RNA-seq) [23], marker discovery and genotyping can be performed simultaneously at comparatively low cost where reference genomes are available (genetic variation is inferred from read alignments to a known genome sequence, Figure 1E). Exploiting this data type, Pool [41] developed a particularly elegant approach for QTL detection with BSA data, including the identification of linked causal loci, although the method was tailored for *Drosophila* species. More recently, Mansfeld and Grumet [42] released the QTLseqr package, which implements or elaborates several proposed methods to detect QTL in BSA data [43,44]. Further, Wybouw *et al.* [27] recently developed a permutation-based approach for QTL detection applicable to BSA designs, although this method requires moderate-to-high replication of segregating populations. An advantage of the latter two approaches is that only a single input file in the Variant Call Format (VCF) is required for QTL detection. Pipelines of well supported programs, as well as best practice guidelines for their use in generating VCF files, are now robustly established and lower the informatic barrier to entry for BSA studies (see [27,28] for example pipelines with the respective software and citations).

**Beyond theory and simulations: lessons from experimental studies in insects and mites**

Apart from the genetic architecture (monogenic versus polygenic) and species-specific characteristics that can impact the practical implementation of BSA experimental designs, factors that affect power and resolution include the sizes of bulks, generation number, and marker density (Figure 1) [16,31,41,45]. For these parameters, theory and lessons from simulated data suggest that more is usually better than less (e.g., to increase recombination [41]), and recent experimental work in arthropods reflects this. To date, most BSA studies in insects have focused on traits that are easy to score, like pigmentation, and have started with parental strains with large phenotypic differences (e.g., sensitive and highly pesticide-resistant strains). Despite some notable exceptions, small- to modestly-sized F2 or backcross populations of several hundreds of individuals have often been used. Mirroring initial studies in plants, BSA studies in insects have often used sparse genetic data; examples of marker types used to date include simple sequence repeat (microsatellite) markers [46–48], amplified fragment length polymorphism markers [49–51], restriction fragment length polymorphism markers [52], restriction-site associated DNA markers [53], single feature polymorphisms ascertained from microarrays [54–57], random amplified DNA fingerprinting [58], a combination of random amplified polymorphic DNA and microsatellite markers [59], and in a few cases dense marker data obtained from resequencing (e.g.,

[23,25]). The number of genetic markers used in these studies varied greatly, with sparsely (and unevenly) distributed markers likely contributing to low QTL mapping resolution in some cases. Nevertheless, in the beet armyworm *Spodoptera exigua,* BSA mapping that used RNA-seq data for genotyping, aided by knowledge of potential candidate genes and their characterization with RNAi, led to the identification of ATP-binding cassette (ABC) transporters as underlying insensitivity to *Bacillus thuringiensis* (Bt) toxin in a resistant strain [23]. In several other studies, genes in BSA peak regions have been proposed as candidates, such as *doublesex* in the control of sex-related mimicry in the butterfly *Papilio polytes* [53]. In many instances, however, only broad regions of chromosomes harboring many genes were associated with phenotypes. Notwithstanding the introduction of methods like genome editing to a growing number of insect species [60], along with other methodological advances that facilitate functional assessment (validation) of candidates [61], characterizing long lists of genes in broad QTL intervals remains a daunting task.

A small set of studies have now illustrated how more advanced BSA designs, in concert with the effectively saturating genotypic data afforded by high-throughput sequencing, can resolve causal loci to much narrower chromosome regions. A number of these studies have been performed with the two-spotted spider mite, *Tetranychus urticae*, an agricultural pest known for its rapid evolution of pesticide resistance and host plant use [27,62,63]. These mites are small, ~600 μm in length for females, with males substantially smaller. Although single *T. urticae* individuals can be genotyped at a moderate number of loci by PCR [29,64,65], marker-based genotyping at a genome-wide scale is challenging, hindering QTL identification by traditional linkage mapping approaches that have been successful in larger-bodied insects. However, *T. urticae* strains can be crossed, and segregating populations can be expanded to thousands of individuals on detached leaves or on whole plants [9,27–30]. In two studies that started with crosses of *T. urticae* strains sensitive or resistant to mite growth inhibitor compounds, narrow BSA peaks for recessive, monogenic resistance were resolved to a single tiny chromosome region [9,30]. Both studies used large segregating populations that were expanded over ~6 generations; in the larger and more powered of these studies, peaks of haplotype fixation were either within or less than 20 kb from *chitin synthase 1* (*CHS1*) [30], which was subsequently demonstrated by a genome editing approach in *D. melanogaster* to encode the target-site for the growth inhibitor compounds used for selection [66]. With a conceptually similar design, Bryon *et al.* [29] also used BSA mapping, in concert with follow-up studies, to identify mutations in a horizontally transferred gene, *phytoene desaturase*, as causal for the absence of carotenoid-based pigmentation in albino *T. urticae* strains. A reanalysis of Bryon *et al.*'s data with the methods presented in the Figure 2 legend revealed that the BSA peak, as assessed with replicate populations used in that study, was ~95 kb from the causal gene.

Similar designs have also elucidated the genetic basis of polygenic traits. For example, in *D. melanogaster,* Bastide *et al.* [24] used multigeneration, long-term segregant populations and sequencing

of bulks to identify 19 distinct QTL regions for high-altitude melanism in multiple crosses. The BSA peaks for QTL were localized to small genomic intervals that included candidate pigmentation genes, suggesting that many are real. For further corroboration, Bastide *et al.* [24] extended their analyses of genes at BSA peaks to populations of light and dark flies. In some cases, strongly differentiated genetic variants were identified in and nearby candidates, a signal that loci identified in BSA scans explain pigmentation differences in population samples. Two studies in 2019 in *T. urticae* have also highlighted the promise of continuous selection over many generations, followed by the isolation and sequencing of bulks (Figure 1D), to unravel the genetic architecture of complex modes of inheritance [27,28]. In each study, replicated segregant populations derived from pesticide-sensitive and -resistant strains were selected with discriminating pesticide doses for ~25-50 generations. In these cases, resistance phenotypes were known to be polygenic, and multiple resistance QTL were identified, as shown by reproducible shifts toward alleles contributed by resistant parents that were observed in selected as compared to unselected populations (Figure 2). Although new candidate loci identified in this work have yet to be investigated in follow-up studies, the identity of genes near BSA peaks suggests that the QTL are likely real. For instance, several BSA peaks fell within a mere tens of kb from genes encoding known target sites of the pesticides used for selection (Figure 3). Further, major QTL were also resolved to tiny chromosome regions harboring genes with suspected or known roles in xenobiotic detoxification, a major route to resistance [2]. These genes encode a potential xenobiotic receptor, copy variable CYPs, and also cytochrome P450 reductase, which is required for CYP activity, and that was identified as a candidate for resistance to multiple compounds in independent crosses [27,28].

**Concluding remarks**

For most arthropods, the genetic tools that have enabled model-organism geneticists to link genes and alleles to phenotypic variation are poorly developed, if not even entirely absent. For many species, however, strains with contrasting phenotypes – including for pesticide resistance, pigmentation or other visual features, or variation in life history traits – are available or can be readily isolated. Where strains can be crossed, BSA genetic mapping holds great promise for the identification of causal loci. Nevertheless, given the species richness of insects and their relatives, and high levels of intraspecific phenotypic and genotypic variation, the method has been relatively little used to date. We believe this is poised to change rapidly. The emergence of cost-effective, high-throughput short-read sequencing for simultaneous marker discovery and dense genotyping has recently revolutionized BSA studies. However, assembled genomes are nonetheless needed to fully exploit short-read sequence data (Figure 1E,F), and currently limit broader adoption of the method (or for that matter, other genetic approaches as well) [27]. The extent of this limitation was starkly revealed in BSA studies with *T. urticae*, as assembly incompleteness hampered QTL detection. Even though the *T. urticae* draft genome had a scaffold N50 of 2.99 Mb [67], it was only after allele frequency data from BSA populations was used to order scaffolds into a chromosome-level assembly that comprehensive QTL detection was possible

[27]. While the *T. urticae* genome sequence was produced with the high-quality Sanger method, *de novo* assemblies with short-read data are typically far more fragmented. However, newer single-molecule, long-read sequencing technologies, like PacBio and Oxford Nanopore [68,69], are now poised to overcome the genome assembly challenge, as they enable assemblies with scaffold sizes in the many Mb range, and potentially ones of chromosome lengths. These disruptive technologies are already in use for technically challenging genome assembly projects in insects [70]. Given the success of recent BSA studies in insects and mites, coupled with the evolving potential to rapidly generate high-quality genome assemblies, BSA approaches should be considered as a prominent tool for the rapid elucidation of the genetic architecture of trait variation. This is especially true for species for which the biological resources or attributes that have facilitated traditional linkage-based genetic mapping approaches, or alternatively ones like E&R and pool-GWAS, are lacking. However, as recent work in *D. melanogaster* has shown [24], BSA approaches can be an important tool for model-organism geneticists as well.

Declarations of interest: none.

**Acknowledgements**

We are grateful to W. Dermauw, N. Wybouw, and R. Greenhalgh for helpful comments. Articles in our review were selective as they were chosen to highlight the promise and implementation of BSA experimental designs for trait mapping in non-model arthropods; we apologize in advance to our many colleagues for the articles we did not have space to feature. RMC acknowledges funding from the USA National Science Foundation (award number 1457346). TVL acknowledges funding from Research Foundation Flanders (FWO, Belgium) (grant numbers G009312N, G053815N), and the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant numbers 772026-POLYADAPT, 773902-SUPERPEST).

**References**


1. Sharma S, Kooner R, Arora R: **Insect pests and crop losses**. In *Breeding Insect Resistant Crops for Sustainable Agriculture*. Edited by Arora R, Sandhu S. Springer Singapore; 2017:45–66.

2. Feyereisen R, Dermauw W, Van Leeuwen T: **Genotype to phenotype, the molecular and physiological dimensions of resistance in arthropods**. *Pestic Biochem Physiol* 2015, **121**:61–77.

3. Daborn PJ, Yen JL, Bogwitz MR, Le Goff G, Feil E, Jeffers S, Tijet N, Perry T, Heckel D, Batterham P, et al.: **A single P450 allele associated with insecticide resistance in *Drosophila***. *Science* 2002, **297**:2253–2256.

4. Hales KG, Korey CA, Larracuente AM, Roberts DM: **Genetics on the Fly: A Primer on the Drosophila Model System**. *Genetics* 2015, **201**:815–842.

5. Groen SC, Whiteman NK: **Using Drosophila to study the evolution of herbivory and diet specialization**. *Curr Opin Insect Sci* 2016, **14**:66–72.

6. Zhan S, Huang J, Guo Q, Zhao Y, Li W, Miao X, Goldsmith MR, Li M, Huang Y: **An integrated genetic linkage map for silkworms with three parental combinations and its application to the mapping of single genes and QTL**. *BMC Genomics* 2009, **10**:389.

7. Linnen CR, O'Quin CT, Shackleford T, Sears CR, Lindstedt C: **Genetic Basis of Body Color and Spotting Pattern in Redheaded Pine Sawfly Larvae (Neodiprion lecontei)**. *Genetics* 2018, **209**:291–305.

8. Smith HA, White BJ, Kundert P, Cheng C, Romero-Severson J, Andolfatto P, Besansky NJ: **Genome-wide QTL mapping of saltwater tolerance in sibling species of Anopheles (malaria vector) mosquitoes**. *Heredity* 2015, **115**:471–479.

9. Van Leeuwen T, Demaeght P, Osborne EJ, Dermauw W, Gohlke S, Nauen R, Grbic M, Tirry L, Merzendorfer H, Clark RM: **Population bulk segregant mapping uncovers resistance mutations and the mode of action of a chitin synthesis inhibitor in arthropods**. *Proc Natl Acad Sci* 2012, **109**:4407–4412.

10. Bailey SF, Keifer H.: **The Tomato Russet Mite, *Phyllocoptes destructor* Keifer: Its Present Status**. *J Econ Entomol* 1943, **36**:706–712.

11. Polilov AA: **Small Is Beautiful: Features of the Smallest Insects and Limits to Miniaturization**. *Annu Rev Entomol* 2015, **60**:103–121.

12. Giovannoni JJ, Wing RA, Ganal MW, Tanksley SD: **Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations**. *Nucleic Acids Res* 1991, **19**:6553–6568.

13. Michelmore RW, Paran I, Kesseli RV: **Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations.** *Proc Natl Acad Sci* 1991, **88**:9828–9832.

14. Shen X, Zhou M, Lu W, Ohm H: **Detection of *Fusarium* head blight resistance QTL in a wheat population using bulked segregant analysis**. *Theor Appl Genet* 2003, **106**:1041–1047.

15. Schneeberger K, Ossowski S, Lanz C, Juul T, Petersen AH, Nielsen KL, Jørgensen J-E, Weigel D, Andersen SU: **SHOREmap: simultaneous mapping and mutation identification by deep sequencing**. *Nat Methods* 2009, **6**:550–551.

16. Zou C, Wang P, Xu Y: **Bulked sample analysis in genetics, genomics and crop improvement**. *Plant Biotechnol J* 2016, **14**:1941–1955.

17. Ehrenreich IM, Torabi N, Jia Y, Kent J, Martis S, Shapiro JA, Gresham D, Caudy AA, Kruglyak L: **Dissection of genetically complex traits with extremely large pools of yeast segregants**. *Nature* 2010, **464**:1039–1042.

18. Swinnen S, Schaerlaekens K, Pais T, Claesen J, Hubmann G, Yang Y, Demeke M, Foulquié-Moreno MR, Goovaerts A, Souvereyns K, et al.: **Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast using pooled-segregant whole-genome sequence analysis**. *Genome Res* 2012, **22**:975–984.

19. Segrè AV, Murray AW, Leu J-Y: **High-resolution mutation mapping reveals parallel experimental evolution in yeast**. *PLOS Biol* 2006, **4**:e256.

20. Doitsidou M, Poole RJ, Sarin S, Bigelow H, Hobert O: **C. elegans mutant identification with a one-step whole-genome-sequencing and SNP mapping strategy**. *PLOS ONE* 2010, **5**:e15435.

21. Arnold CN, Xia Y, Lin P, Ross C, Schwander M, Smart NG, Müller U, Beutler B: **Rapid identification of a disease allele in mouse through whole genome sequencing and bulk segregation analysis**. *Genetics* 2011, **187**:633–641.

22. Bowen ME, Henke K, Siegfried KR, Warman ML, Harris MP: **Efficient mapping and cloning of mutations in zebrafish by low-coverage whole-genome sequencing**. *Genetics* 2012, **190**:1017–1024.

23. Park Y, González-Martínez RM, Navarro-Cerrillo G, Chakroun M, Kim Y, Ziarsolo P, Blanca J, Cañizares J, Ferré J, Herrero S: **ABCC transporters mediate insect resistance to multiple Bt toxins revealed by bulk segregant analysis**. *BMC Biol* 2014, **12**:46.

24. Bastide H, Lange JD, Lack JB, Yassin A, Pool JE: **A variable genetic architecture of melanic evolution in *Drosophila melanogaster***. *Genetics* 2016, **204**:1307–1319.

●● In this study, which used an advanced, multigenerational BSA genetic design, the authors were able to identify numerous QTL for putatively adaptive pigmentation differences in *D. melanogaster* populations. The findings highlight the complexity of adaptive evolution, even for a comparatively simple trait.

25. Jagadeesan R, Fotheringham A, Ebert PR, Schlipalius DI: **Rapid genome wide mapping of phosphine resistance loci by a simple regional averaging analysis in the red flour beetle, *Tribolium castaneum***. *BMC Genomics* 2013, **14**:650.

26. Bryon A, Kurlovs AH, Van Leeuwen T, Clark RM: **A molecular-genetic understanding of diapause in spider mites: current knowledge and future directions: Molecular genetics of mite diapause**. *Physiol Entomol* 2017, **42**:211–224.

27. Wybouw N, Kosterlitz O, Kurlovs AH, Bajda S, Greenhalgh R, Snoeck S, Bui H, Bryon A, Dermauw W, Van Leeuwen T, et al.: **Long-term population studies uncover the genome structure and genetic basis of xenobiotic and host plant adaptation in the herbivore *Tetranychus urticae***. *Genetics* 2019, **211**:1409–1427.

●● This 2019 study with the spider mite *Tetranychus urticae* started with a classic BSA experimental set up, but included an evolve-and-resequence component. QTL for both pesticide resistance and host adaptation (to tomato) were identified, and the former were resolved to small chromosome regions that suggested specific candidate genes and alleles. The authors also exploited the continuous nature of allele frequency changes in long-term populations to resolve the *T. urticae* draft assembly to the expected three chromosomes.

28. Snoeck S, Kurlovs AH, Bajda S, Feyereisen R, Greenhalgh R, Villacis-Perez E, Kosterlitz O, Dermauw W, Clark RM, Van Leeuwen T: **High-resolution QTL mapping in *Tetranychus urticae* reveals acaricide-specific responses and common target-site resistance after selection by different METI-I acaricides**. *Insect Biochem Mol Biol* 2019, **110**:19–33.

●● In this 2019 study the authors began with a traditional BSA design, but included elements of evolve-and-resequence, to identify multiple QTL for resistance to three acaricides with the same mode of action. The authors concluded that resistance was explained by a combination of target-site and metabolic resistance, with variation among compounds. For a case in which the causal variant was already suspected, BSA peaks localized to within about 30 kb, confirming the utility of BSA studies to resolve QTL to small genomic intervals.

29. Bryon A, Kurlovs AH, Dermauw W, Greenhalgh R, Riga M, Grbić M, Tirry L, Osakabe M, Vontas J, Clark RM, et al.: **Disruption of a horizontally transferred phytoene desaturase abolishes carotenoid accumulation and diapause in *Tetranychus urticae***. *Proc Natl Acad Sci* 2017, **114**:E5871–E5880.

●● In this work with *Tetranychus urticae*, a monogenic trait (spontaneous lack of carotenoid pigmentation in a laboratory strain) was mapped to a small genomic interval with a BSA approach. While the experimental design used populations expanded over many generations before selection, only several parallel replicate selections were used. The study revealed that easy to implement BSA experimental designs can revolve large-effect variants to small chromosome regions; the authors were able to identify the causal gene in follow-up studies.

30. Demaeght P, Osborne EJ, Odman-Naresh J, Grbić M, Nauen R, Merzendorfer H, Clark RM, Van Leeuwen T: **High resolution genetic mapping uncovers chitin synthase-1 as the target-site of the structurally diverse mite growth inhibitors clofentezine, hexythiazox and etoxazole in *Tetranychus urticae***. *Insect Biochem Mol Biol* 2014, **51**:52–61.

31. Schneeberger K: **Using next-generation sequencing to isolate mutant genes from forward genetic screens**. *Nat Rev Genet* 2014, **15**:662–676.

32. Burke MK, Dunham JP, Shahrestani P, Thornton KR, Rose MR, Long AD: **Genome-wide analysis of a long-term evolution experiment with *Drosophila***. *Nature* 2010, **467**:587–590.

33. Turner TL, Stewart AD, Fields AT, Rice WR, Tarone AM: **Population-Based Resequencing of Experimentally Evolved Populations Reveals the Genetic Basis of Body Size Variation in Drosophila melanogaster**. *PLOS Genet* 2011, **7**:e1001336.

34. Turner TL, Miller PM: **Investigating Natural Variation in Drosophila Courtship Song by the Evolve and Resequence Approach**. *Genetics* 2012, **191**:633–642.

35. Bastide H, Betancourt A, Nolte V, Tobler R, Stöbe P, Futschik A, Schlötterer C: **A genome-wide, fine-scale map of natural pigmentation variation in Drosophila melanogaster**. *PLoS Genet* 2013, **9**:e1003534.

36. Schlötterer C, Tobler R, Kofler R, Nolte V: **Sequencing pools of individuals — mining genome-wide polymorphism data without big funding**. *Nat Rev Genet* 2014, **15**:749–763.

37. Schlötterer C, Kofler R, Versace E, Tobler R, Franssen SU: **Combining experimental evolution with next-generation sequencing: a powerful tool to study adaptation from standing genetic variation**. *Heredity* 2015, **114**:431–440.

38. Kofler R, Schlötterer C: **A guide for the design of evolve and resequencing studies**. *Mol Biol Evol* 2014, **31**:474–483.

39. Baldwin-Brown JG, Long AD, Thornton KR: **The power to detect quantitative trait loci using resequenced, experimentally evolved populations of diploid, sexual organisms**. *Mol Biol Evol* 2014, **31**:1040–1055.

40. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, Hall KP, Evers DJ, Barnes CL, Bignell HR, et al.: **Accurate whole human genome sequencing using reversible terminator chemistry**. *Nature* 2008, **456**:53–59.

41. Pool JE: **Genetic mapping by bulk segregant analysis in *Drosophila* : experimental design and simulation-based inference**. *Genetics* 2016, **204**:1295–1306.

●● In this study, the author developed an analysis method to identify QTL from BSA data, including estimates of their effect sizes and confidence intervals. A novel aspect of the approach is that the method can evaluate overlapping peaks in BSA scans to assess (resolve) if multiple causal loci are likely to explain observed shifts in allele frequencies.

42. Mansfeld BN, Grumet R: **QTLseqr: An R package for bulk segregant analysis with next-generation sequencing**. *Plant Genome* 2018, **11**:180006.

●● QTLseqr is a computationally fast, user-friendly package in the widely-used R programming language that takes data from a VCF file as input, that is, genetic data on pools of offspring with contrasting phenotypes in a BSA design, and implements several methods for the identification of

statistically significance QTL intervals. In addition to QTL detection, the package generates plots for their visualization.


43. Magwene PM, Willis JH, Kelly JK: **The statistics of bulk segregant analysis using next generation sequencing**. *PLoS Comput Biol* 2011, **7**:e1002255.

44. Takagi H, Abe A, Yoshida K, Kosugi S, Natsume S, Mitsuoka C, Uemura A, Utsushi H, Tamiru M, Takuno S, et al.: **QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations**. *Plant J Cell Mol Biol* 2013, **74**:174–183.

45. Bazakos C, Hanemian M, Trontin C, Jiménez-Gómez JM, Loudet O: **New strategies and tools in quantitative genetics: how to go from the phenotype to the genotype**. *Annu Rev Plant Biol* 2017, **68**:435–455.

46. Chandrakanth N, Moorthy SM, Ponnuvel KM, Sivaprasad V: **Identification of microsatellite markers linked to thermotolerance in silkworm by bulk segregant analysis and in silico mapping**. *Genetika* 2015, **47**:1063–1078.

47. Behrens D, Huang Q, Geßner C, Rosenkranz P, Frey E, Locke B, Moritz RFA, Kraus FB: **Three QTL in the honey bee *Apis mellifera L.* suppress reproduction of the parasitic mite *Varroa destructor*: QTL-Mapping of *Varroa* resistance in honeybees**. *Ecol Evol* 2011, **1**:451–458.

48. Jairin J, Leelagud P, Pongmee A, Srivilai K: **Chromosomal location of a recessive red-eye mutant gene in the brown planthopper *Nilaparvata lugens* (Stål) (Insecta: Hemiptera)**. *Adv Entomol* 2017, **5**:33–39.

49. Baxter SW, Papa R, Chamberlain N, Humphray SJ, Joron M, Morrison C, ffrench-Constant RH, McMillan WO, Jiggins CD: **Convergent evolution in the genetic basis of müllerian mimicry in *Heliconius* butterflies**. *Genetics* 2008, **180**:1567–1577.

50. Heckel DG, Gahan LJ, Liu YB, Tabashnik BE: **Genetic mapping of resistance to *Bacillus thuringiensis* toxins in diamondback moth using biphasic linkage analysis**. *Proc Natl Acad Sci U S A* 1999, **96**:8373–8377.

51. Henniges-Janssen K, Reineke A, Heckel DG, Groot AT: **Complex inheritance of larval adaptation in *Plutella xylostella* to a novel host plant**. *Heredity* 2011, **107**:421–432.

52. Morlais I, Severson DW: **Identification of a polymorphic mucin-like gene expressed in the midgut of the mosquito, *Aedes aegypti*, using an integrated bulked segregant and differential display analysis**. *Genetics* 2001, **158**:1125–1136.

53. Kunte K, Zhang W, Tenger-Trolander A, Palmer DH, Martin A, Reed RD, Mullen SP, Kronforst MR: **Doublesex is a mimicry supergene**. *Nature* 2014, **507**:229–232.

54. Desjardins CA, Gadau J, Lopez JA, Niehuis O, Avery AR, Loehlin DW, Richards S, Colbourne JK, Werren JH: **Fine-scale mapping of the *Nasonia* genome to chromosomes using a high-density genotyping microarray**. *G3 GenesGenomesGenetics* 2013, **3**:205–215.

55. Lai C-Q, Leips J, Zou W, Roberts JF, Wollenberg KR, Parnell LD, Zeng Z-B, Ordovas JM, Mackay TFC: **Speed-mapping quantitative trait loci using microarrays**. *Nat Methods* 2007, **4**:839–841.

56. Ocorr KA, Crawley T, Gibson G, Bodmer R: **Genetic variation for cardiac dysfunction in *Drosophila***. *PLoS ONE* 2007, **2**:e601.

57. Zhang Z, Hsieh B, Poe A, Anderson J, Ocorr K, Gibson G, Bodmer R: **Complex genetic architecture of cardiac disease in a wild type inbred strain of *Drosophila melanogaster***. *PLoS ONE* 2013, **8**:e62909.

58. Schlipalius DI, Cheng Q, Reilly PEB, Collins PJ, Ebert PR: **Genetic linkage analysis of the lesser grain borer *Rhyzopertha dominica* identifies two loci that confer high-level resistance to the fumigant phosphine**. *Genetics* 2002, **161**:773–782.

59. Gadau J, Gerloff CU, Krüger N, Chan H, Schmid-Hempel P, Wille A, Page RE: **A linkage analysis of sex determination in *Bombus terrestris* (L.) (Hymenoptera: Apidae)**. *Heredity* 2001, **87**:234–242.

60. Trible W, Olivos-Cisneros L, McKenzie SK, Saragosti J, Chang N-C, Matthews BJ, Oxley PR, Kronauer DJC: **orco Mutagenesis Causes Loss of Antennal Lobe Glomeruli and Impaired Social Behavior in Ants**. *Cell* 2017, **170**:727-735.e10.

• This study is one of a growing number in which genome-editing methods have been implemented in species other than *D. melanogaster*. This is an important development as even though BSA genetic mapping (or mapping by other methods) can resolve causal loci to small genomic intervals, establishing causality of genes and variants is still laborious, and is especially so in the absence of tools that allow functional characterization.

61. Turner TL: **Fine-mapping natural alleles: quantitative complementation to the rescue**. *Mol Ecol* 2014, **23**:2377–2382.

62. Van Leeuwen T, Vontas J, Tsagkarakou A, Dermauw W, Tirry L: **Acaricide resistance mechanisms in the two-spotted spider mite *Tetranychus urticae* and other important Acari: A review**. *Insect Biochem Mol Biol* 2010, **40**:563–572.

63. Gould F: **Rapid host range evolution in a population of the phytophagous mite *Tetranychus urticae* Koch**. *Evolution* 1979, **33**:791–802.

64. Bajda S, Dermauw W, Panteleri R, Sugimoto N, Douris V, Tirry L, Osakabe M, Vontas J, Van Leeuwen T: **A mutation in the PSST homologue of complex I (NADH:ubiquinone oxidoreductase) from Tetranychus urticae is associated with resistance to METI acaricides**. *Insect Biochem Mol Biol* 2017, **80**:79–90.

65. Riga M, Bajda S, Themistokleous C, Papadaki S, Palzewicz M, Dermauw W, Vontas J, Leeuwen TV: **The relative contribution of target-site mutations in complex acaricide resistant phenotypes as assessed by marker assisted backcrossing in Tetranychus urticae**. *Sci Rep* 2017, **7**:9202.

66. Douris V, Steinbach D, Panteleri R, Livadaras I, Pickett JA, Van Leeuwen T, Nauen R, Vontas J: **Resistance mutation conserved between insects and mites unravels the benzoylurea insecticide mode of action on chitin biosynthesis**. *Proc Natl Acad Sci U S A* 2016, **113**:14692–14697.

• The authors used genome editing with CRISPR/Cas9 in *D. melanogaster* to introduce and validate variants in genes encoding chitin synthase identified in other arthropod species as causal for resistance

to a collection of chitin synthesis disrupting compounds. The study highlights the utility of functional studies in *D. melanogaster* to test the causality of genes and variants identified in arthropods for which possibilities for functional analyses are currently limited or non-existent.

67. Grbić M, Van Leeuwen T, Clark RM, Rombauts S, Rouzé P, Grbić V, Osborne EJ, Dermauw W, Thi Ngoc PC, Ortego F, et al.: **The genome of *Tetranychus urticae* reveals herbivorous pest adaptations**. *Nature* 2011, **479**:487–492.

68. Jain M, Olsen HE, Paten B, Akeson M: **The Oxford Nanopore MinION: delivery of nanopore sequencing to the genomics community**. *Genome Biol* 2016, **17**:239.

69. Rhoads A, Au KF: **PacBio sequencing and its applications**. *Genomics Proteomics Bioinformatics* 2015, **13**:278–289.

70. Kingan S, Heaton H, Cudini J, Lambert C, Baybayan P, Galvin B, Durbin R, Korlach J, Lawniczak M: **A High-Quality De novo Genome Assembly from a Single Mosquito Using PacBio Sequencing**. *Genes* 2019, **10**:62.

• In this recent study, the authors were able to use a small amount of genomic DNA isolated from a single *Anopheles coluzzii* individual to produce a high-quality genome assembly by PacBio Single Molecule, Real-Time (SMRT) sequencing. This is an advance as standard implementations of the method require large DNA inputs, which can be a challenge, or even an impossibility in practice, for many small-bodied insects and mites.

**Figure 1** (1.5 column width figure)

Experimental designs for BSA genetic mapping. Illustrations depict spider mites, but the methods are generally applicable where crosses can be performed and derived, segregating populations can be expanded in controlled settings in the laboratory, greenhouses, or field settings (e.g., in cage enclosures). (**A**) Parental strains with contrasting genotypes (chromosomes are indicated as blue or red rectangles) and phenotypes (gray shading) are crossed to produce an F2 population (or backcross population, not shown) harboring recombinant chromosomes (**B**). To allow additional recombination events to accrue, populations can also be propagated for additional generations (**C**). In the traditional BSA design, bulks are collected at a defined endpoint (**B** or **C;** red lines with arrows indicate the step at which the phenotype of interest is selected). The bulks consist of individuals with contrasting extremes in the phenotype of interest, e.g., visual differences like pigmentation. Alternatively, fitness differences in response to a selective agent can be assessed, as for a pesticide treatment, in which case "unselected" and resistant "selected" bulks are prepared. The specific scenario illustrated is for a hypothetical case of monogenic pesticide resistance. A variant of the traditional BSA design involves selection across multiple generations (**D**). Either way, DNA is prepared from bulk samples for genotyping, including by high-throughput, short-read sequencing as indicated (**E**). Read alignments to a reference genome sequence are used for discovery of markers and assessment of allele frequencies in sliding windows. In the case of monogenic recessive inheritance, a single fixation event at and nearby the causal variant in the selected bulk is observed (**F**; AF, allele frequency). To account for systematic deviations in allele frequencies in populations independent of the trait of interest (e.g., as can happen in the case of purging of deleterious alleles, selection for alleles favorable in a laboratory environment, or as a result of segregation distortion), a comparison of allele frequencies in replicates of selected relative to unselected populations is typically performed (see also Figure 2).

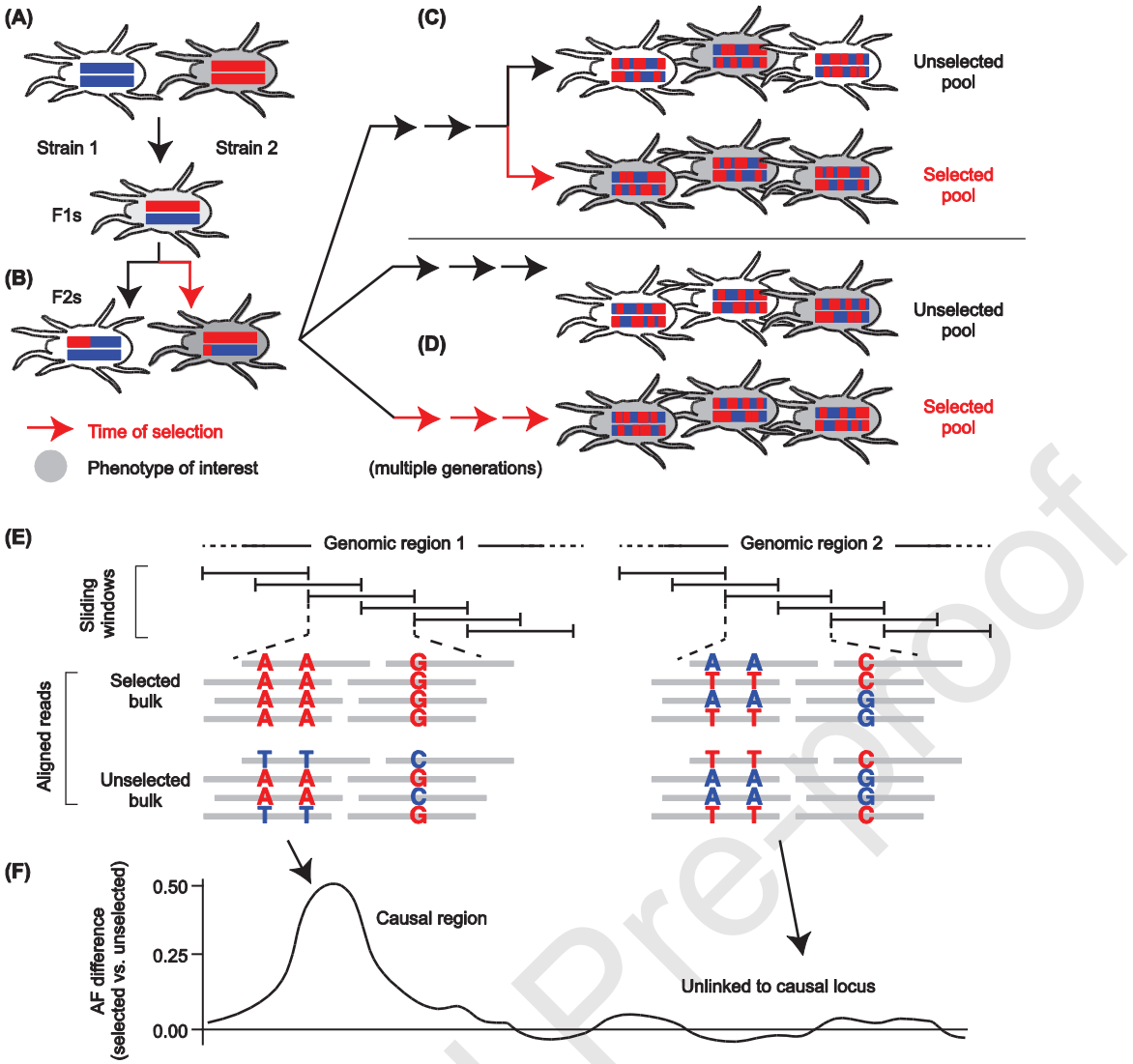



**Figure 2** (one column width figure)

Example allele frequencies of replicate *T. urticae* populations under selection by the pesticide spirodiclofen in a study that applied selection over many generations. Shown are the raw allele frequencies of three representative replicates of unselected populations (blue lines) and spirodiclofen-selected populations (red lines) as assessed in sliding windows (500 kb with a 25 kb offset). The data used to construct the plots are from Wybouw *et al.* [27] (the experimental design was after Figure 1A,B,D). As plotted, vertical deflections indicate increases in the frequency of alleles coming from the spirodiclofen-resistant parental strain. Systematic differences in allele frequencies between the paired unselected and selected populations (gray shading) indicate two QTL [27], and are indicated by vertical dashed lines. Note that several regions of fixation (or near fixation) are observed even in unselected populations (black arrows), potentially reflecting the purging of deleterious alleles (but see the Figure 1

legend for other possibilities). The code used to plot this figure was adapted from Wybouw *et al.* [27] and Snoeck *et al.* [28], and has been made available on Github (https://github.com/rmclarklab/BSA).

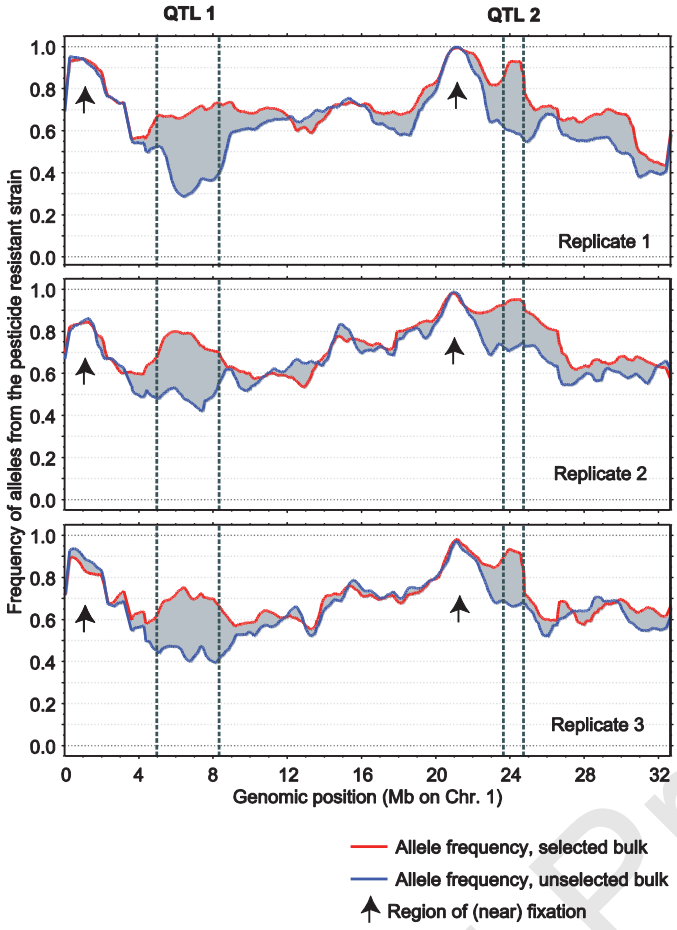
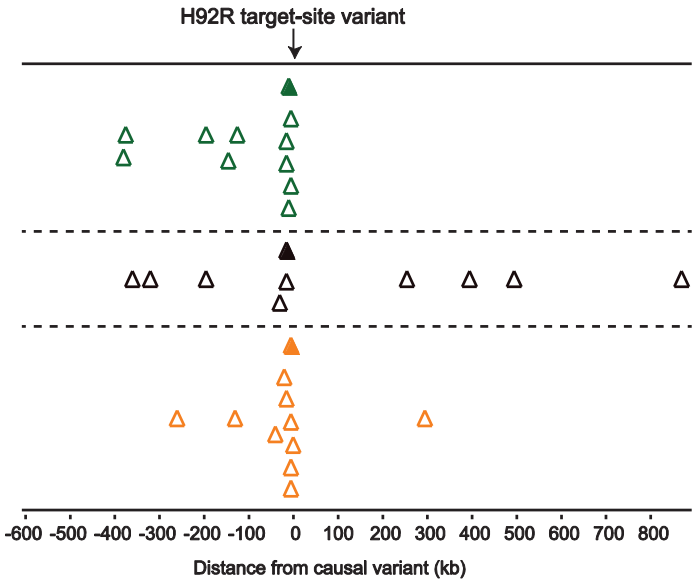



**Figure 3** (one column width figure)

BSA mapping resolution for a pesticide resistance QTL. Data from nine or ten selected populations from Snoeck *et al.* [28] in *T. urticae* for each of three pesticides (fenpyroximate, pyridaben, and tebufenpyrad) was reanalyzed along with control populations (see Figure 2 legend for methods; 75 kb windows were used with 5 kb offsets; experimental design after Figure 1A,B,D). The three pesticides are Mitochondrial Electron Transport Inhibitors of complex I (METI-Is), for which the histidine-to-arginine change at position 92 (denoted H92R) in a gene encoding a subunit of NADH:ubiquinone oxidoreductase associates with target-site resistance [28]. For each pesticide, the BSA peaks calculated by combining all replicates into a single analysis (solid triangles) were within tens of kb of the causal variant; this was true for some replicates also (open triangles), but in a moderate number of cases, peaks were much farther from the causal variant (especially for selection by pyridaben, for which the populations may have undergone a bottleneck during the propagation steps [28]). Therefore, while BSA approaches can provide high mapping resolution, relying on a small number of replicates can potentially be misleading.




Fenpyroximate selections:
BSA replicate peaks (n = 10) BSA peak, all replicates
Pyridaben selections:
BSA replicate peaks (n = 9) BSA peak, all replicates
Tebufenpyrad selections:
BSA replicate peaks (n = 10) BSA peak, all replicates