

Identification of loci that cause phenotypic variation in diverse species with the reciprocal hemizyosity test

David L. Stern

Janelia Research Campus, Howard Hughes Medical Institute, Ashburn, VA 20147, USA

The reciprocal hemizyosity test is a straightforward genetic test that can positively identify genes that have evolved to contribute to a phenotypic difference between strains or between species. The test involves a comparison between hybrids that are genetically identical throughout the genome except at the test locus, which is rendered hemizygous for alternative alleles from the two parental strains. If the two reciprocal hemizygotes display different phenotypes, then the two parental alleles must have evolved. New methods for targeted mutagenesis will allow application of the reciprocal hemizyosity test in many organisms. This review discusses the principles, advantages, and limitations of the test.

From candidate genes to causality

Natural phenotypic variation is often distributed continuously and results from the combined effects of variation at multiple loci and environmental influences [1]. This is true for phenotypic variation within and between natural populations and species [2], between strains of domesticated species [3,4], and for variation among individuals that contributes to disease risk [5]. Therefore, geneticists have long hoped to identify the loci underlying such ‘quantitative’ traits to better understand evolution, domestication, and disease.

Considerable progress in genomics and in the statistical methods required for analysis of genetic crosses and segregating populations has facilitated identification of genomic regions that contribute to quantitative traits – so-called quantitative trait loci (QTLs) – in many organisms, but the individual loci underlying these QTLs has eluded identification for most species [2,3,6]. In some cases, QTLs have been dissected with further mapping experiments, but this work has rarely been straightforward and – outside yeast [7] – dissection down to the individual gene level remains the exception. Usually, genetic linkage mapping studies provide lists of candidate genes within QTL regions that may be causal [8]. In addition, genome-wide association mapping, divergence mapping, and admixture mapping in some cases can provide high-resolution mapping of

candidate loci contributing to quantitative traits [9–12]. However, in most cases these genes have remained candidates. Promotion of candidate genes to the level of causal loci usually requires extensive further analysis of gene function, which has not been available or straightforward in many non-model species.

Three genetic tests can provide strong evidence that a candidate gene harbors variation that causes phenotypic variation. The ‘gold standard’ is to perform homologous recombination to precisely replace a genomic region with the orthologous genomic region from a different strain or species (Figure 1A). Homologous recombination is challenging to perform in most species, although the recent development of CRISPR–Cas9-based genome editing technologies may make this approach more practical in many species [13,14]. Nonetheless, this experiment requires considerable prior knowledge of the function of genomic regions, which may not be available in many species with complex genomes, and the experiment is likely to remain laborious or impossible for tests of large genomic regions.

Transgenesis also has been used to demonstrate causality of allelic differences between strains or species (Figure 1B). For studies of species differences, this experiment often involves transferring a DNA region from multiple species into a single species in which transgenesis is straightforward [15–17]. This experimental design provides some control over differences in the *trans*-regulatory environment, but it is not as ideal as precise reciprocal replacement by homologous recombination (Box 1).

The third test, the reciprocal hemizyosity test [18–21], overcomes all of the limitations of transgenesis and is simpler to perform than homologous recombination, although it has limitations of its own. In this test, reciprocal crosses are performed between one wild type strain and one strain carrying a null allele for the candidate gene. If the reciprocal hybrids – which are hemizygous for the candidate gene from the two parental strains – possess different phenotypes, then it can be concluded that the candidate locus contributes to the difference between the parental strains. The reciprocal hemizyosity test requires only the generation of null alleles of the candidate gene in both strains or species of interest and the ability to cross the strains or species. It is also possible to generate the allele-specific mutations directly in hybrids [22]. This is a quantitative test; the difference between reciprocal hemizygotes represents the quantitative contribution of the

Corresponding author: Stern, D.L. (sternd@janelia.hhmi.org).

Keywords: reciprocal hemizyosity test; natural variation; evolution.

0168-9525/

© 2014 Elsevier Ltd. All rights reserved. <http://dx.doi.org/10.1016/j.tig.2014.09.006>

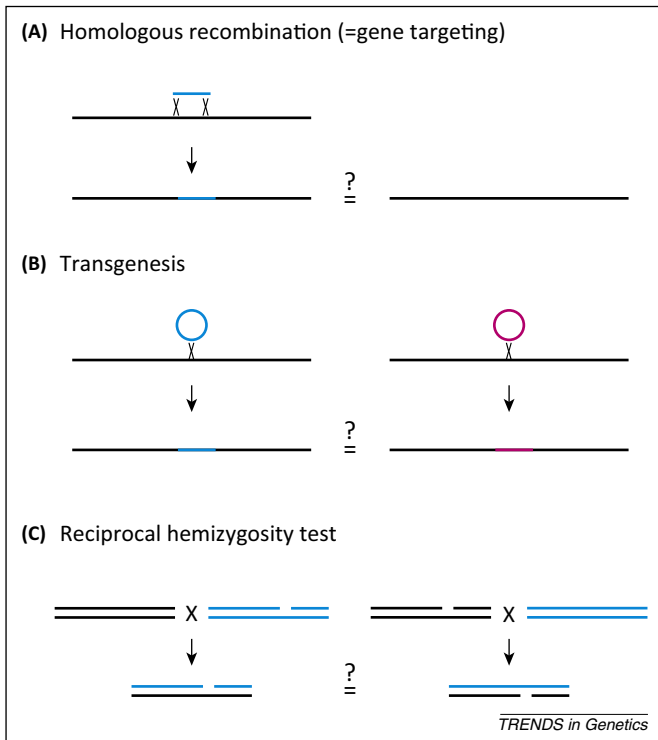


Figure 1. Homologous recombination, reciprocal transgenesis, and the reciprocal hemizygosity test. **(A)** Homologous recombination involves direct replacement of a gene region with homologous DNA from another organism. The organism carrying the homologous DNA can be compared with an organism carrying the original genome to determine whether this gene region contributes to a phenotypic difference. The blue and black lines represent a linear DNA strand and a chromosome, respectively. **(B)** Transgenic test of homologous genetic regions in a common genetic background. The blue and red circles represent plasmids carrying the gene to be inserted. The black lines represent chromosomes. **(C)** The reciprocal hemizygosity test requires crosses between strains in which a null mutation or deletion has been generated. The two homologous chromosomes are represented as parallel black and blue lines. The F1 offspring resulting from reciprocal crosses are genetically identical at all loci except at the location of the mutation, where they are hemizygous for alternative parental alleles. This provides a test of the contribution of this candidate region to a phenotypic difference between strains.

alternative alleles to the difference between strains. Comparison of the difference between the reciprocal hemizygotes with the difference between the parental strains that carry the null alleles provides an estimate of the fraction of the evolved difference resulting from evolution of the candidate gene. The test is insensitive to dosage effects and, because only hemizygotes are generated, it can be used to test essential genes. The test can also be performed on multiple loci simultaneously to detect interactions between evolved loci [23–25]. The test does not require any knowledge of the function of the candidate gene nor a detailed understanding of the structure of the locus. The test therefore serves as a useful early step to quickly rule a candidate gene in or out as the causal locus underlying phenotypic variation.

The reciprocal hemizygosity test has been used primarily to test candidate genes identified initially by mapping experiments. However, for many phenotypic characteristics, a collection of candidate genes often can be generated based on many other sources of information. The reciprocal hemizygosity test is sufficiently efficient that, depending on the research goals, it may be advantageous to skip preliminary mapping experiments and test all candidate genes [26].

Box 1. Limitations of transgenic assays

Over the past few decades, inter-strain and inter-species transgenesis experiments have been used to test for evolved differences in genes (e.g., [15–17,57,58]). However, there are at least six issues that should be considered when performing such assays. First, transgenesis of DNA into non-native genomic locations can generate novel and/or variable patterns of gene expression [59]. Second, for the study of large genes it may be difficult to include the entire relevant DNA region for transgenesis [60]. These two issues may cause a transgene to fail to rescue completely, and in some cases it may be challenging to distinguish this effect from possible evolutionary differences [61]. Third, transgenes may disrupt activity of neighboring genes in the target genome [62]. Fourth, if the transgenes are used to perform functional rescue assays (rather than only to assay expression), then the transgene should be tested in a genetic background that is homozygous null for the relevant locus in the host genome. This may limit which host species can be used for functional transgenic assays. Finally, it is often not clear how to interpret the effects of transgenes in the absence of other information on the expected effect size of the candidate genes. If the comparison strains or species can be crossed, then it is advisable to perform a linkage study first to estimate the effect size for the candidate gene. In the absence of this information, transgenic rescue experiments essentially are sign tests of no effect versus some effect, which provides weak evidence for the identification of causal loci underlying evolutionary differences. Inference is further compromised because the multiple experimental limitations of transgenic assays discussed above may, on their own, generate small magnitude effects. Without knowledge of the expected magnitude of effects, even very small magnitude effects may be interpreted (or misinterpreted) as important [16]. For all of these reasons, transgenic assays should be interpreted with caution and they may fall out of favor in coming years with rapid advances in direct genome manipulation allowing widespread use of the reciprocal hemizygosity test and homologous recombination.

The reciprocal hemizygosity test is distinct from, and should not be confused with, the quantitative complementation test (Box 2). The quantitative complementation test involves a comparison of alleles from two genetic backgrounds against a third genetic background containing a mutation in a candidate gene or region and a fourth background containing a wild type allele [27–29]. Because of the increased complexity of genetic backgrounds compared with the reciprocal hemizygosity test, it is likely that in most cases uncontrolled non-focal loci contribute to the observed effects, leading to the high type I error rate associated with this test [30,31]. Additional limitations of the quantitative complementation test probably also inflate the type II error rate (Box 2). It is strongly recommended that investigators avoid the quantitative complementation test in favor of the reciprocal hemizygosity test.

Although homologous recombination and transgenic approaches are established technologies for multiple model organisms [32], the reciprocal hemizygosity test appears not to be well known outside of the yeast genetics community [7,20,21,28,33–35]. This is a shame, because the test is both powerful and broadly applicable in principle. The availability of new efficient tools for genome manipulation across a broad range of species, including the use of CRISPR–Cas9 for targeting mutagenesis and homologous recombination [14,13], means that the test can be applied to identify the causal loci underlying phenotypic variation in a wide range of taxa. Previous uses of the reciprocal hemizygosity test are reviewed below to illustrate the types of questions that can be addressed, and limitations

Box 2. Limitations of the quantitative complementation test

The quantitative complementation test has been proposed as a method to identify genes that have evolved between strains [3,27,29]. The test is derived from the classical complementation test, which is a powerful method to assign alleles to complementation groups (loci) [63]. In the classical complementation test, two alleles are considered to affect the same locus if the phenotype of a *trans*-heterozygote resembles the homozygous phenotype of at least one of the separate alleles. Interpretation of the complementation test is straightforward when both alleles encode loss of function and are recessive. Interpretation of complementation test results is fraught when these conditions are not met [63].

The quantitative complementation test, following this logic, attempts to determine whether natural variants of a single locus (*Q* and *q*) can be identified as alleles of a known locus [8,27,29]. The test is performed by separately crossing the two variants to a null allele for a candidate locus (*Null*). As a control for background genetic effects, both alleles (*Q* and *q*) are crossed also to a chromosome carrying a 'wild type' allele (*WT*). Failure to complement (allelism to the candidate locus) is revealed if the contrast between the two *trans*-heterozygotes carrying the null allele is greater than the contrast between the two *trans*-heterozygotes carrying the wild type allele:

$$(Q/Null - q/Null) > (Q/WT - q/WT)$$

In practice, this deviation is detected as a significant interaction term in an analysis of variance (ANOVA). The underlying assumption is that the two natural variants have more pronounced phenotypic effects in the hemizygous state than they do when paired with a wild type allele, implying that the wild type allele buffers the effects of the variant alleles [28].

There are several practical limitations of the quantitative complementation test. First, as the test is commonly performed, it is likely to yield many false positives [30]. This is caused most often by epistasis with other variable loci in the cross [30,31]. For example, in *Drosophila*, mutations are usually maintained over balancer chromosomes. Both the mutant and balancer chromosome may therefore accumulate deleterious alleles and modifier alleles of the target mutation. Quantitative complementation tests that use these stocks are, therefore, tests of both the target mutation and of many unknown mutations, which can generate complex patterns of epistasis and generate false positive results. Second, the quantitative complementation test may generate false negative results if the alleles do not behave as expected by the assumed genetic model. For example, if the effects of all functional alleles in the crosses (*Q*, *q*, and *WT*) are strictly additive, then the test will not reject the null hypothesis, even if the variant alleles are allelic to the test locus.

and experimental design issues are discussed that should be considered when extending the test to diverse species.

Previous uses of the reciprocal hemizygosity test

The reciprocal hemizygosity test has been applied most widely to various strains of the yeast, *Saccharomyces cerevisiae* (Table 1), in part because it has long been straightforward to generate targeted gene deletions in yeast. Application of the reciprocal hemizygosity test has dramatically improved our understanding of the genetic causes for quantitative variation in yeast, as illustrated by the 40 studies in yeast listed in Table 1. These studies have led to the identification of many loci in yeast that vary for ecologically and economically important traits. For example, Steinmetz *et al.* [19], who named the test, used the reciprocal hemizygosity test to dissect a single QTL region for high-temperature growth into three closely linked causal loci. Similarly, the test has been used to identify multiple small-effect loci contributing to low glycerol and high ethanol yield in yeast fermentation [36]. The test has also been used to identify both large-effect and closely linked small-effect alleles in wild populations that contribute to sporulation efficiency [37,38]. The test is a particularly powerful way of dissecting the contributions of closely linked causal loci, as demonstrated in a study that identified four linked genes contributing to sporulation efficiency [24].

The reciprocal hemizygosity test has also been used to study evolutionary variation in several other species. For example, it was employed to determine whether a candidate gene, the homeobox-containing gene *Ultrabithorax* (*Ubx*), contributed to a morphological difference between the fruit fly species *Drosophila melanogaster* and *D. simulans* [18]. Reciprocal crosses between the parental strains and newly-induced *Ubx* mutant strains in each species demonstrated that a substantial proportion of the difference in the patterning of trichomes on the legs between the species was caused by evolution of *Ubx*. This experiment was practical in 1998 only because null alleles could be generated at *Ubx* at a reasonable frequency using irradiation, and because these alleles are dominant, which

simplified screening. Until recently, the difficulty of making null alleles at other candidate loci has prevented use of this test to examine other traits in these and other species.

This *Drosophila* study illustrates one of the powerful advantages of the reciprocal hemizygosity test for evolutionary studies. Crosses between different species – such as between these *Drosophila* species – often yield sterile offspring, which prevents further recombination mapping to identify QTLs and evolving genes. The reciprocal hemizygosity test, however, requires only the generation of F1 hybrids. This simple fact makes the test applicable to a huge range of species pairs for which classical genetic analysis is impossible.

Studies in two other species have applied the reciprocal hemizygosity test with modified methods. One study in mice employed reciprocal hemizygotes with randomized genetic backgrounds to study obesity-related traits [39]. A second study used RNAi to perform allele-specific knock-down of strain-specific alleles to identify a gene contributing to mosquito (*Anopheles gambiae*) resistance to malaria parasites [40].

Until recently, yeast has held a major advantage over other organisms for the application of the reciprocal hemizygosity test. The introduction of TALENs [41] and the CRISPR–Cas9 system [14,13] for targeted mutagenesis, however, will now allow the reciprocal hemizygosity test to be used in a vast array of species, including many non-model organisms. Many aspects of morphology, physiology, and behavior differ between closely related species [42], and simple mutagenesis combined with the ability to assay F1 offspring in reciprocal hemizygosity tests will allow the identification of many genes that contribute to these species differences.

Future applications and potential limitations of the reciprocal hemizygosity test

The reciprocal hemizygosity test is poised to provide a huge push forward in the identification of the genes that contribute to natural variation in many species, and it is important that experiments are planned with a clear

Table 1. Studies employing the reciprocal hemizyosity test sorted by date illustrate that application of the test is accelerating

Taxa	Trait	Date	Refs
<i>Drosophila melanogaster</i> and <i>D. simulans</i>	Patterning of trichomes on legs	1998	[18]
<i>Saccharomyces cerevisiae</i> strains	High-temperature growth	2002	[19]
<i>Mus musculus</i> strains	Hepatic lipase activity	2004	[39]
<i>S. cerevisiae</i> strains	Sporulation efficiency	2005	[54]
<i>S. cerevisiae</i> strains	High-temperature growth	2006	[23]
<i>S. cerevisiae</i> strains	Sporulation efficiency	2006	[24]
<i>S. cerevisiae</i> strains	High-temperature growth	2008	[25]
<i>Anopheles gambiae</i> strains	Resistance to malaria parasites – RNAi test	2009	[40]
<i>S. cerevisiae</i> strains	Sporulation efficiency	2009	[37]
<i>S. cerevisiae</i> strains	Telomere length	2009	[64]
<i>S. cerevisiae</i> strains	Drug sensitivity	2009	[65]
<i>S. cerevisiae</i> strains	Alkali stress	2010	[66]
<i>S. cerevisiae</i> strains	Heat stress	2011	[67]
<i>S. cerevisiae</i> strains	High-temperature growth	2011	[56]
<i>S. cerevisiae</i> strains	Morphological variation	2011	[68]
<i>S. cerevisiae</i> strains	Fermentation rate	2011	[69]
<i>S. cerevisiae</i> strains	Colony color and size	2012	[30]
<i>S. cerevisiae</i> strains	Oenological traits	2012	[70]
<i>S. cerevisiae</i> strains	Sporulation efficiency	2012	[38]
<i>S. cerevisiae</i> lab and pathogenic strain	High-temperature growth	2012	[71]
<i>S. cerevisiae</i> strains	Production of wine aroma compounds	2012	[72]
<i>S. cerevisiae</i> strains	High ethanol-tolerance	2013	[73]
<i>S. cerevisiae</i> strains	High-temperature growth	2013	[55]
<i>S. cerevisiae</i> strains	Glycerol/ethanol ratio	2013	[36]
<i>S. cerevisiae</i> strains	Growth rate in multiple media and colony size and shape	2013	[22]
<i>S. cerevisiae</i> strains	Variation in nitrogen source use	2013	[26]
<i>S. cerevisiae</i> strains	Glycerol/ethanol ratio	2013	[74]
<i>S. cerevisiae</i> strains	Maximal ethanol accumulation capacity	2013	[75]
<i>S. cerevisiae</i> strains	Ammonium-toxicity resistance	2013	[76]
<i>S. cerevisiae</i> strains	High iron-toxicity resistance	2013	[77]
<i>S. cerevisiae</i> strains	Growth in glycerol	2013	[78]
<i>S. cerevisiae</i> strains	Growth response to stress	2013	[79]
<i>S. paradoxus</i> strains	Morphological and growth variation	2013	[80]
<i>S. cerevisiae</i> strains	Response to nutrient starvation	2014	[81]
<i>S. cerevisiae</i> strains	Ethanol tolerance	2014	[82]
<i>S. cerevisiae</i> strains	Lag-phase duration	2014	[83]
<i>S. cerevisiae</i> strains	Nitrogen metabolism	2014	[84]
<i>S. cerevisiae</i> strains	Fermentation under nitrogen limitation	2014	[85]
<i>S. cerevisiae</i> strains	Hydrogen sulfide production	2014	[86]
<i>S. cerevisiae</i> strains	Transcriptomic response to ethanol	2014	[87]
<i>S. cerevisiae</i> strains	Colony morphology	2014	[88]
<i>S. cerevisiae</i> strains	Protein level variation	2014	[89]
<i>S. cerevisiae</i> strains	Response to stress during bioethanol fermentation	2014	[90]

understanding of the advantages and limitations of this method. There are several obvious requirements for application of the reciprocal hemizyosity test. First, the test requires the ability to perform controlled crosses of the focal strains or species that yield offspring that are viable, at least to the desired developmental stage. Second, the test requires mutations at the same candidate locus in both strains or species. If a species can be bred in the laboratory and embryos can be injected with DNA or RNA, then it is likely that the reciprocal hemizyosity test can be employed. Given that we will likely see an explosion in the use of the reciprocal hemizyosity test with the advent of precision genome editing, I first describe important limitations of the test that are specific to species with genomes of greater complexity than the yeast genome.

The precise design of mutations for the test is discussed in a separate section below.

Compared to multicellular organisms, the yeast genome is relatively simple, with small, compact genes and no sex chromosomes. When considering applications of the reciprocal hemizyosity test to other organisms, six potential limitations should be considered.

First, the reciprocal hemizyosity test cannot test for the effects of loci on the sex chromosome in the heterogametic sex (for example, the X chromosome in male fruit flies) because the test requires diploidy. X-linked candidate genes can be tested in the homogametic sex, except in mammals, where X-chromosome inactivation makes the homogametic sex effectively hemizygous for the X chromosome [43].

Second, tests of autosomal loci in the heterogametic sex may be influenced differentially by alternative alleles on the sex chromosome when crosses are performed in different directions to generate the reciprocal hemizygotes. This can be controlled for easily by ensuring that the sex chromosomes are identical in the reciprocal hemizygotes, for example by assaying only the homogametic sex or by comparing heterogametic individuals only from crosses performed in the same direction.

Third, in haplodiploid species (such as honeybees and *Nasonia* wasps), the reciprocal hemizygosity test cannot be performed in the haploid sex (males).

Fourth, some null alleles are dominant lethal and cannot be employed in the reciprocal hemizygosity test.

Fifth, mutations in genes required for viability and fertility may be challenging to maintain in non-model organisms. One approach is to generate mutations by targeted insertion of a dominant visible marker gene, such as that encoding GFP. Another approach is to track the mutation in heterozygous individuals each generation with molecular markers. Alternatively, it may be possible to use CRISPR–Cas9 to generate ‘balancer chromosomes’, inversions that suppress recombination and that contain a gratuitous dominant visible marker and a lethal allele. Deleterious mutations could then be maintained indefinitely in a *trans*-heterozygous state with this balancer. Balancer chromosomes are the ‘trick’ that has allowed the flourishing of *Drosophila* genetics over the past century, and it may be possible to exploit this approach in a wider range of species [44]. By contrast, it is probably best to perform the reciprocal hemizygosity test soon after new alleles are generated to minimize the evolution of linked modifier alleles that can bias the results of any genetic test.

Finally, *trans* interactions between alleles on homologous chromosomes may bias the results of the reciprocal hemizygosity test. These effects are best known from studies of *Drosophila*, in part because in Diptera homologous chromosomes are synapsed in somatic cells. It is possible that this is a concern only for studies of Diptera. Nevertheless, because many reciprocal hemizygosity tests will likely be performed using fly species in coming years, and because it is possible that these effects are more general, I provide a short review of these effects. There are two major classes of effects that are known to influence expression of alleles on homologous chromosomes: transvection and *trans*-splicing.

Transvection is the regulation of alleles on homologous chromosomes by enhancers located in *trans* [45]. For example, when a wild type allele is paired with a null allele, transvection may allow regulation of the wild type allele by enhancers located on the chromosome carrying the null allele. Transvection, then, will tend to equalize the observed expression driven by alternative alleles in a reciprocal hemizygosity test. While it has long been thought that transvection was limited to Diptera, similar effects have been detected also in fungi [46], mice [47,48], humans [49], and plants [50]. Studies of transvection have revealed that enhancer regions preferentially regulate the promoter located in *cis*; removal of a promoter located in *cis* can drive the enhancer to more strongly regulate the promoter in *trans* [51]. This fact has obvious implications for design of a

reciprocal hemizygosity test. Most importantly, if a deletion is employed to make a null allele, then care should be taken to preserve the native promoter. This issue is discussed further below.

Trans-splicing, the generation of a mature mRNA transcript by splicing of the products from homologous chromosomes, has also been described in *Drosophila* [52]. In principle, *trans*-splicing could generate false negative results in the reciprocal hemizygosity test if the causal evolutionary differences reside in an exon. For example, if the causal mutation resides 3' of the null mutation generated for the test, then *trans*-splicing could bypass the causal mutation and tend to equalize the effects of the two alleles. Although this effect is likely to be rare, it is important to remain aware that this *trans* effect may act at some loci.

Design considerations for the reciprocal hemizygosity test

Assuming a pair of study strains or species can be crossed, the next obvious question is ‘what type of mutation should be generated?’ In yeast, typically the entire coding region is deleted, from the start to the stop codon. However, yeast genes are compact, with small *cis*-regulatory regions usually located directly adjacent to genes, and there is no evidence that transvection acts in yeast. By contrast, multicellular organisms have complex *cis*-regulatory regions distributed both adjacent to, within, and between exons [53]. Of greater concern is that the *cis*-regulatory information for some genes resides in the introns of neighboring genes. Thus, a deletion of one gene may actually alter function of multiple non-target genes. Therefore, although it may not matter precisely what type of null mutation is generated in yeast, the specific type of mutation employed in most other species should be considered carefully.

There are five major classes of mutations that can be considered for the reciprocal hemizygosity test: (i) small mutations that generate protein null alleles, (ii) deletion of all of the coding exons, (iii) deletion of the ‘entire’ locus, (iv) deletion of multiple loci, and (v) deletion of specific *cis*-regulatory modules. Each class of mutations has its own advantages and disadvantages.

Point mutations or small deletions that generate protein null alleles are likely to produce the most reliable and precise results. These mutations have the advantage that they are unlikely to disrupt any of the regulatory information for the target gene or neighboring genes. In addition, small mutations are unlikely to disrupt the promoter, and this will reduce the likelihood that transvection can influence the test. For many species, other advantages of making a small protein null mutation include the lack of requirement to understand the potentially complex *cis*-regulatory architecture for each gene and the relative ease of generating such mutations. The major drawback of this approach is that only one gene can be tested per mutation. A second potential drawback is that a small mutation within the coding region does not protect entirely against the potential for transvection. Another caveat is that some mutations in genes with alternative splice variants may render one isoform null while leaving other isoforms intact.

Deletions from the start to the stop codon are used universally in yeast and might seem, at first glance, to provide a favorable reagent for the reciprocal hemizyosity test. In most other species, however, the disadvantages of these deletions are likely to outweigh any advantages over a small mutation that generates a protein null. As mentioned above, inadvertent deletion of the *cis*-regulatory information for neighboring genes can result in unintentional tests of multiple loci, generating false positives. For example, Deutschbauer and Davis [54] reported that a test of one gene in their study of yeast sporulation may be a false positive because the deletion probably interferes with the regulation of a neighboring known causal gene. Finally, as discussed above, deletions that remove the promoter may enhance transvection, which will tend to equalize the effects of homologous alleles and which may generate false negative results.

Deletions that remove the entire locus, including all of the relevant *cis*-regulatory information, are another option for the reciprocal hemizyosity test. A major advantage of this approach is that removal of the regulatory region for the target gene will protect against transvection. However, these deletions may still remove *cis*-regulatory information used by neighboring genes and generate false positives. A further disadvantage of whole-locus deletion is that functional delineation of the full extent of a locus – including all of the relevant regulatory DNA – can be a major undertaking and may not be practical for most species.

Deletions that remove multiple loci provide a way to accelerate discovery of the evolved loci within a genomic region that contains multiple candidate genes [55,56]. Because this experiment would normally only be a prelude to mutagenesis of individual genes, it does not suffer from the disadvantages associated with deletion of a single gene. A second approach that may improve throughput is reciprocal hemizyosity scanning [22] which employs a pool of multiple single-gene deletions that are tested simultaneously. In this experiment, a population of hybrids is generated, exposed to selection, and allele frequencies are examined after selection. In practice, this method has yielded many false positives [22], probably due to the generation of off-target chromosomal abnormalities associated with the generation of the targeted deletions. It is likely that new methods of generating deletions may improve the specificity of these deletion reagents and make reciprocal hemizyosity scanning a viable option.

Finally, in principle, the reciprocal hemizyosity test could be performed by deletion of individual *cis*-regulatory modules. If a regulatory region contains mutations that alter gene function, then a reciprocal hemizyosity test using deletions of this region will reveal the contribution of these variants to a phenotypic difference. There are at least two advantages to this experiment. First, this test provides a much more precise estimate of the gene regions that have contributed to evolution than does a protein null. Second, some null mutations generate dominant lethality, and targeted deletion of enhancers allows tests of genes that cannot otherwise be used in the reciprocal hemizyosity test. There are several disadvantages to this approach. First, the relevant *cis*-regulatory region must first be identified. Second, this test may generate a false negative

result if the relevant variation is distributed across multiple *cis*-regulatory regions and any one region is insufficient on its own to cause an altered phenotype.

In summary, there is no one-size-fits-all mutation for the reciprocal hemizyosity test, and the choice of the specific class of mutations employed depends on the specific biological question being asked.

Concluding remarks

The reciprocal hemizyosity test provides a robust method for determining whether a candidate gene (or candidate *cis*-regulatory module) contributes to an evolved phenotypic difference between strains or species. There are a few limitations associated with application of the test, and the specific biological question being asked will determine the optimal mutation that should be employed. In addition, the results of the test should always be considered in combination with complementary sources of data, such as patterns and levels of gene expression and other information on gene function. Recent improvements in targeted mutagenesis are likely to make the reciprocal hemizyosity test applicable to a broad range of species, leading to a substantial improvement in our understanding of the genetic causes of evolution, agricultural crops, and disease-related traits.

Acknowledgments

I thank Scott Hawley, Mark Rebeiz, Thomas Turner, and three anonymous referees for helpful comments on the manuscript.

References

- 1 Falconer, D.S. and Mackay, T.F.C. (1996) *Introduction to Quantitative Genetics*, Addison Wesley Longman
- 2 Slate, J.O.N. (2005) Quantitative trait locus mapping in natural populations: progress, caveats and future directions. *Mol. Ecol.* 14, 363–379
- 3 Georges, M. (2007) Mapping, fine mapping, and molecular dissection of quantitative trait loci in domestic animals. *Annu. Rev. Genomics Hum. Genet.* 8, 131–162
- 4 Ross-Ibarra, J. (2005) Quantitative trait loci and the study of plant domestication. *Genetica* 123, 197–204
- 5 Plomin, R. et al. (2009) Common disorders are quantitative traits. *Nat. Rev. Genet.* 10, 872–878
- 6 Paran, I. and Zamir, D. (2003) Quantitative traits in plants: beyond the QTL. *Trends Genet.* 19, 303–306
- 7 Fay, J.C. (2013) The molecular basis of phenotypic variation in yeast. *Curr. Opin. Genet. Dev.* 23, 672–677
- 8 Mackay, T.F.C. (2001) Quantitative trait loci in *Drosophila*. *Nat. Rev. Genet.* 2, 11–20
- 9 Winkler, C.A. et al. (2010) Admixture mapping comes of age. *Annu. Rev. Genomics Hum. Genet.* 11, 65–89
- 10 Shriner, D. et al. (2011) Mapping of disease-associated variants in admixed populations. *Genome Biol.* 12, 223
- 11 Crawford, J.E. and Nielsen, R. (2013) Detecting adaptive trait loci in nonmodel systems: divergence or admixture mapping? *Mol. Ecol.* 22, 6131–6148
- 12 Carlson, C.S. et al. (2004) Mapping complex disease loci in whole-genome association studies. *Nature* 429, 446–452
- 13 Sander, J.D. and Joung, J.K. (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355
- 14 Charpentier, E. and Doudna, J.A. (2013) Rewriting a genome. *Nature* 495, 50–51
- 15 Jeong, S. et al. (2008) The evolution of gene regulation underlies a morphological difference between two *Drosophila* sister species. *Cell* 132, 783–793
- 16 Cretokos, C.J. et al. (2008) Regulatory divergence modifies limb length between mammals. *Genes Dev.* 22, 141–151

- 17 Frankel, N. *et al.* (2012) Conserved regulatory architecture underlies parallel genetic changes and convergent phenotypic evolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, 20975–20979
- 18 Stern, D.L. (1998) A role of Ultrabithorax in morphological differences between *Drosophila* species. *Nature* 396, 463–466
- 19 Steinmetz, L.M. *et al.* (2002) Dissecting the architecture of a quantitative trait locus in yeast. *Nature* 416, 326–330
- 20 Christians, J.K. and Keightley, P.D. (2002) Genetic architecture: dissecting the genetic basis of phenotypic variation identifying genes that influence phenotypic variation. *Curr. Biol.* 12, 415–416
- 21 Darvasi, A. and Pisanté-Shalom, A. (2002) Complexities in the genetic dissection of quantitative trait loci. *Trends Genet.* 18, 489–491
- 22 Wilkening, S. *et al.* (2014) An evaluation of high-throughput approaches to QTL mapping in *Saccharomyces cerevisiae*. *Genetics* 196, 853–865
- 23 Sinha, H. *et al.* (2006) Complex genetic interactions in a quantitative trait locus. *PLoS Genet.* 2, e13
- 24 Ben-Ari, G. *et al.* (2006) Four linked genes participate in controlling sporulation efficiency in budding yeast. *PLoS Genet.* 2, e195
- 25 Sinha, H. *et al.* (2008) Sequential elimination of major-effect contributors identifies additional quantitative trait loci conditioning high-temperature growth in yeast. *Genetics* 180, 1661–1670
- 26 Beltran, G. *et al.* (2013) Genetic basis of variations in nitrogen source utilization in four wine commercial yeast strains. *PLoS ONE* 8, e67166
- 27 Turner, T.L. (2014) Fine-mapping natural alleles: quantitative complementation to the rescue. *Mol. Ecol.* 23, 2377–2382
- 28 Georges, M. (2012) Impact of high-throughput genotyping and sequencing on the identification of genes and variants underlying phenotypic variation in domestic cattle. In *Bovine Genomics* (Womack, J.E., ed.), pp. 234–258, Wiley-Blackwell
- 29 Long, A.D. *et al.* (1996) Genetic interactions between naturally occurring alleles at quantitative trait loci and mutant alleles at candidate loci affecting bristle number in *Drosophila melanogaster*. *Genetics* 144, 1497–1510
- 30 Kim, H.S. *et al.* (2012) A noncomplementation screen for quantitative trait alleles in *Saccharomyces cerevisiae*. *G3 (Bethesda)* 2, 753–760
- 31 Service, P.M. (2004) How good are quantitative complementation tests? *Sci. Aging Knowledge Environ.* 2004, pe13
- 32 Segal, D.J. and Meckler, J.F. (2013) Genome engineering at the dawn of the golden age. *Annu. Rev. Genomics Hum. Genet.* 14, 135–158
- 33 Swinnen, S. *et al.* (2012) Genetic mapping of quantitative phenotypic traits in *Saccharomyces cerevisiae*. *FEMS Yeast Res.* 12, 215–227
- 34 Liti, G. and Louis, E.J. (2012) Advances in quantitative trait analysis in yeast. *PLoS Genet.* 8, e1002912
- 35 Parts, L. (2014) Genome-wide mapping of cellular traits using yeast. *Yeast* 31, 197–205
- 36 Hubmann, G. *et al.* (2013) Identification of multiple interacting alleles conferring low glycerol and high ethanol yield in *Saccharomyces cerevisiae* ethanolic fermentation. *Biotechnol. Biofuels* 6, 87
- 37 Gerke, J. *et al.* (2009) Genetic interactions between transcription factors cause natural variation in yeast. *Science* 323, 498–501
- 38 Lorenz, K. and Cohen, B.A. (2012) Small- and large-effect quantitative trait locus interactions underlie variation in yeast sporulation efficiency. *Genetics* 192, 1123–1132
- 39 Farahani, P. *et al.* (2004) Reciprocal hemizyosity analysis of mouse hepatic lipase reveals influence on obesity. *Obes. Res.* 12, 292–305
- 40 Blandin, S. and *et al.* (2009) Dissecting the genetic basis of resistance to malaria parasites in *Anopheles gambiae*. *Science* 326, 147–150
- 41 Joung, J.K. and Sander, J.D. (2013) TALENs: a widely applicable technology for targeted genome editing. *Nat. Rev. Mol. Cell Biol.* 14, 49–55
- 42 Orgogozo, V. and Stern, D.L. (2009) How different are recently diverged species? – more than 150 phenotypic differences have been reported for the *D. melanogaster* species subgroup. *Fly (Austin)* 3, 117
- 43 Avner, P. and Heard, E. (2001) X-chromosome inactivation: counting, choice and initiation. *Nat. Rev. Genet.* 2, 59–67
- 44 Hentges, K.E. and Justice, M.J. (2004) Checks and balancers: balancer chromosomes to facilitate genome annotation. *Trends Genet.* 20, 252–259
- 45 Duncan, I.W. (2002) Transvection effects in *Drosophila*. *Annu. Rev. Genet.* 36, 521–556
- 46 Aramayo, R. and Metzberg, R.L. (1996) Meiotic transvection in fungi. *Cell* 86, 103–113
- 47 Herman, H. *et al.* (2003) Trans allele methylation and paramutation-like effects in mice. *Nat. Genet.* 34, 199–202
- 48 Rassoulzadegan, M. *et al.* (2002) Transvection effects involving DNA methylation during meiosis in the mouse. *EMBO J.* 21, 440–450
- 49 Ashe, H.L. *et al.* (1997) Intergenic transcription and transinduction of the human beta-globin locus. *Genes Dev.* 11, 2494–2509
- 50 Grant-Downton, R.T. and Dickinson, H.G. (2004) Plants, pairing and phenotypes – two’s company? *Trends Genet.* 20, 188–195
- 51 Geyer, P.K. *et al.* (1990) Tissue-specific transcriptional enhancers may act in trans on the gene located in the homologous chromosome: the molecular basis of transvection in *Drosophila*. *EMBO J.* 9, 2247–2256
- 52 Mongelard, F. *et al.* (2002) Trans-splicing as a novel mechanism to explain interallelic complementation in *Drosophila*. *Genetics* 160, 1481–1487
- 53 Davidson, E.H. (2006) *The Regulatory Genome: Gene Regulatory Networks in Development and Evolution*, Academic Press
- 54 Deutschbauer, A.M. and Davis, R.W. (2005) Quantitative trait loci mapped to single-nucleotide resolution in yeast. *Nat. Genet.* 37, 1333–1340
- 55 Yang, Y. *et al.* (2013) QTL analysis of high thermotolerance with superior and downgraded parental yeast strains reveals new minor QTLs and converges on novel causative alleles involved in RNA processing. *PLoS Genet.* 9, e1003693
- 56 Cubillos, F. and *et al.* (2011) Assessing the complex architecture of polygenic traits in diverged yeast populations. *Mol. Ecol.* 20, 1401–1413
- 57 Laurie-Ahlberg, C.C. and Stam, L.F. (1987) Use of P-element-mediated transformation to identify the molecular basis of naturally occurring variants affecting Adh expression in *Drosophila melanogaster*. *Genetics* 115, 129–140
- 58 Yoon, H.S. and Baum, D.A. (2004) Transgenic study of parallelism in plant morphological evolution. *Proc. Natl. Acad. Sci. U.S.A.* 101, 6524–6529
- 59 Henikoff, S. (1992) Position effect and related phenomena. *Curr. Opin. Genet. Dev.* 2, 907–912
- 60 Cande, J. *et al.* (2014) Looking under the lamp post: neither fruitless nor doublesex have evolved to generate divergent male courtship in *Drosophila*. *Cell Rep.* 8, 363–370
- 61 Campesan, S. *et al.* (2001) The nonA gene in *Drosophila* conveys species-specific behavioral characteristics. *Genetics* 158, 1535–1543
- 62 Bellen, H.J. *et al.* (2004) The BDGP gene disruption project: single transposon insertions associated with 40% of *Drosophila* genes. *Genetics* 167, 761–781
- 63 Hawley, R.S. and Gilliland, W.D. (2006) Sometimes the result is not the answer: the truths and the lies that come from using the complementation test. *Genetics* 174, 5–15
- 64 Liti, G. *et al.* (2009) Segregating YKU80 and TLC1 alleles underlying natural variation in telomere properties in wild yeast. *PLoS Genet.* 5, e1000659
- 65 Kim, H.S. and Fay, J.C. (2009) A combined-cross analysis reveals genes with drug-specific and background-dependent effects on drug sensitivity in *Saccharomyces cerevisiae*. *Genetics* 183, 1141–1151
- 66 Romano, G.H. *et al.* (2010) Different sets of QTLs influence fitness variation in yeast. *Mol. Syst. Biol.* 6, 346
- 67 Parts, L. *et al.* (2011) Revealing the genetic structure of a trait by sequencing a population under selection Revealing the genetic structure of a trait by sequencing a population under selection. *Genome Res.* 21, 1131–1138
- 68 Lee, H.N. *et al.* (2011) Natural variation in CDC28 underlies morphological phenotypes in an environmental yeast isolate. *Genetics* 188, 723–730
- 69 Ambroset, C. *et al.* (2011) Deciphering the molecular basis of wine yeast fermentation traits using a combined genetic and genomic approach. *G3 (Bethesda)* 1, 263–281
- 70 Salinas, F. *et al.* (2012) The genetic basis of natural variation in oenological traits in *Saccharomyces cerevisiae*. *PLoS ONE* 7, e49640
- 71 Fraser, H.B. *et al.* (2012) Polygenic cis-regulatory adaptation in the evolution of yeast pathogenicity. *Genome Res.* 22, 1930–1939
- 72 Steyer, D. *et al.* (2012) QTL mapping of the production of wine aroma compounds by yeast. *BMC Genomics* 13, 573
- 73 Swinnen, S. *et al.* (2012) Identification of novel causative genes determining the complex trait of high ethanol tolerance in yeast

- using pooled-segregant whole-genome sequence analysis. *Genome Res.* 22, 975–984
- 74 Hubmann, G. *et al.* (2013) Quantitative trait analysis of yeast biodiversity yields novel gene tools for metabolic engineering. *Metab. Eng.* 17, 68–81
- 75 Pais, T.M. *et al.* (2013) Comparative polygenic analysis of maximal ethanol accumulation capacity and tolerance to high ethanol levels of cell proliferation in yeast. *PLoS Genet.* 9, 1–18
- 76 Reisser, C. *et al.* (2013) Genetic basis of ammonium toxicity resistance in a sake strain of yeast: a Mendelian case. *G3 (Bethesda)* <http://dx.doi.org/10.1534/g3.113.005884>
- 77 Lee, H.N. *et al.* (2013) Divergence of iron metabolism in wild Malaysian yeast. *G3 (Bethesda)* 3, 2187–2194
- 78 Swinnen, S. *et al.* (2013) Re-evaluation of glycerol utilization in *Saccharomyces cerevisiae*: characterization of an isolate that grows on glycerol without supporting supplements. *Biotechnol. Biofuels* 6, 157
- 79 Cubillos, F. and *a et al.* (2013) High-resolution mapping of complex traits with a four-parent advanced intercross yeast population. *Genetics* 195, 1141–1155
- 80 Roop, J.I. and Brem, R.B. (2013) Rare variants in hypermutable genes underlie common morphology and growth traits in wild *Saccharomyces paradoxus*. *Genetics* 195, 513–525
- 81 Bergström, A. *et al.* (2014) A high-definition view of functional genetic variation from natural yeast genomes. *Mol. Biol. Evol.* 31, 872–888
- 82 Duitama, J. *et al.* (2014) Improved linkage analysis of quantitative trait loci using bulk segregants unveils a novel determinant of high ethanol tolerance in yeast. *BMC Genomics* 15, 207
- 83 Zimmer, A. *et al.* (2014) QTL Dissection of lag phase in wine fermentation reveals a new translocation responsible for *Saccharomyces cerevisiae* adaptation to sulfite. *PLoS ONE* 9, e86298
- 84 Jara, M. *et al.* (2014) Mapping genetic variants underlying differences in the central nitrogen metabolism in fermenter yeasts. *PLoS ONE* 9, e86533
- 85 Brice, C. *et al.* (2014) A genetic approach of wine yeast fermentation capacity in nitrogen-starvation reveals the key role of nitrogen signaling. *BMC Genomics* 15, 495
- 86 Huang, C. *et al.* (2014) MET2 affects production of hydrogen sulfide during wine fermentation. *Appl. Microbiol. Biotechnol.* 98, 7125–7135
- 87 Lewis, J.A. *et al.* (2014) Genetic architecture of ethanol-responsive transcriptome variation in *Saccharomyces cerevisiae* strains. *Genetics* <http://dx.doi.org/10.1534/genetics.114.167429>
- 88 Taylor, M.B. and Ehrenreich, I.M. (2014) Genetic interactions involving five or more genes contribute to a complex trait in yeast. *PLoS Genet.* 10, e1004324
- 89 Parts, L. *et al.* (2014) Heritability and genetic basis of protein level variation in an outbred population. *Genome Res.* 24, 1363–1370
- 90 Greetham, D. *et al.* (2014) The genetic basis of variation in clean lineages of *Saccharomyces cerevisiae* in response to stresses encountered during bioethanol fermentations. *PLoS ONE* 9, e103233