

## EVOLUTIONARY GENETICS

# An essential cell cycle regulation gene causes hybrid inviability in *Drosophila*

Nitin Phadnis,<sup>1†</sup> EmilyClare P. Baker,<sup>2\*</sup> Jacob C. Cooper,<sup>1</sup> Kimberly A. Frizzell,<sup>1</sup> Emily Hsieh,<sup>2</sup> Aida Flor A. de la Cruz,<sup>2</sup> Jay Shendure,<sup>3,4</sup> Jacob O. Kitzman,<sup>3,5</sup> Harmit S. Malik<sup>2,6†</sup>

Speciation, the process by which new biological species arise, involves the evolution of reproductive barriers, such as hybrid sterility or inviability between populations. However, identifying hybrid incompatibility genes remains a key obstacle in understanding the molecular basis of reproductive isolation. We devised a genomic screen, which identified a cell cycle–regulation gene as the cause of male inviability in hybrids resulting from a cross between *Drosophila melanogaster* and *D. simulans*. Ablation of the *D. simulans* allele of this gene is sufficient to rescue the adult viability of hybrid males. This dominantly acting cell cycle regulator causes mitotic arrest and, thereby, inviability of male hybrid larvae. Our genomic method provides a facile means to accelerate the identification of hybrid incompatibility genes in other model and nonmodel systems.

Genetic crosses between *Drosophila melanogaster* females and males from its closest sister species, *D. simulans*, produce only adult hybrid F<sub>1</sub> females (1, 2). These unisexual broods are a result of hybrid F<sub>1</sub> male inviability between these species, which manifests during larval stages of development. Despite decades of investigation, the genetic basis of this hybrid F<sub>1</sub> male inviability remains incompletely resolved (3, 4). A series of x-ray mutagenesis experiments previously revealed that a complex interaction between the *D. melanogaster* X-chromosome and dominant alleles from the *D. simulans* second and third chromosomes is necessary to kill hybrids (5, 6). The isolation of hybrid rescue strains that produce viable hybrid F<sub>1</sub> males led to the identification of two causal elements of this hybrid incompatibility: *Hybrid male rescue* (*Hmr*) on the *D. melanogaster* X-chromosome (7, 8) and *Lethal hybrid rescue* (*Lhr*) on the *D. simulans* second chromosome (9, 10). The absence of either *Lhr<sup>sim</sup>* or *Hmr<sup>mel</sup>* results in viable hybrid males (fig. S1). However, *D. melanogaster* males that carry transgenic copies of *D. simulans* *Lhr* are viable despite carrying both the *Hmr<sup>mel</sup>* and *Lhr<sup>sim</sup>* incompatible alleles (9). These results suggest that the presence of at least one additional unidentified hybrid incompatibility gene is necessary to cause hybrid male inviability.

Traditional genetic approaches have failed to identify this missing hybrid incompatibility gene for several reasons. First, hybrid sterility and inviability between *D. melanogaster* and *D. simulans* hinder recombination-based methods for gene identification. Second, genetic disruptions in *D. melanogaster* do not assist in identifying this gene because it is a dominantly acting *D. simulans* factor. Third, the lack of efficient balancer chromosomes in *D. simulans* prevents the construction and maintenance of mutation-accumulation lines that could help identify this missing incompatibility gene. Finally, all known naturally occurring hybrid rescue alleles are mutations of either *Hmr* or *Lhr*; no new rescue alleles have been identified that may correspond to a third gene. Together, these roadblocks have prevented the identification of this missing hybrid incompatibility gene.

Because no null alleles for the missing *D. simulans* hybrid incompatibility gene have been isolated from natural populations, we speculated that—in contrast to *Hmr* and *Lhr*—this gene might be essential for viability. We reasoned that the complex epistatic interaction underlying hybrid F<sub>1</sub> male inviability is analogous to a multi-component toxin; reconstitution of this toxin requires the simultaneous presence of all components. Under this model, hybrid inviability does not occur when even one of the components or hybrid incompatibility genes is missing (e.g., loss of either *Lhr<sup>sim</sup>* or *Hmr<sup>mel</sup>* rescues hybrid males). Extending this analogy, we sought other genes whose ablation results in viable hybrid males using a simple genomics-based approach (Fig. 1A).

We mutagenized 55,000 *D. simulans* males by feeding adults with ethyl methane sulfonate (EMS) and crossed these males to *D. melanogaster* females. All resulting progeny inherit one mutagenized complement of the *D. simulans* genome

and one intact complement from *D. melanogaster*. When *D. simulans* sperm carrying null mutations at any F<sub>1</sub> hybrid incompatibility gene fertilize *D. melanogaster* eggs, the resulting hybrid male progeny are predicted to be viable. This strategy allows us to survey mutations in all *D. simulans* genes that may be involved in the F<sub>1</sub> hybrid incompatibility, even those in essential genes; however, haploinsufficient genes (i.e., genes that require two copies for viability) would not be sampled.

We recovered 32 viable hybrid F<sub>1</sub> males from these crosses (compared with >300,000 hybrid F<sub>1</sub> females). Of these, 26 males were the result of a nondisjunction event that led to their inheriting a *D. simulans* X-chromosome (11); these males were viable, as shown previously (2). Further tests confirmed that we had recovered six independently produced rescue hybrid F<sub>1</sub> males, each of which is viable because of mutations at a locus different from *Lhr<sup>sim</sup>* (11).

Because rescue hybrid F<sub>1</sub> males isolated from these crosses are sterile, they cannot be used in genetic crosses to map the causal gene. Instead, we performed high-throughput sequencing to obtain whole-genome sequences of each of the six independently derived rescue hybrid males and of both parental strains. We then compared the *D. simulans*-derived component of the genomes of rescue hybrid F<sub>1</sub> males to the unmutagenized *D. simulans* parental strain. This allowed us to identify all new mutations in each of the rescue males (11) (table S1 and fig. S2). As expected, most of the EMS-induced mutations were point substitutions (Fig. 1B). However, we also identified two large partially overlapping deletions, which mapped to the *D. simulans*-derived chromosomal arm 3R (Fig. 1B and fig. S3). Each of the six rescue males carried between 600 and 1200 new mutations, as expected, on the basis of the random mutagenesis strategy. Only one *D. simulans* gene, however, was disrupted across all six rescue hybrid males (Fig. 1C). This gene was *Suppressor of Killer-of-prune* [*Su(Kpn)*]-*glutathione-S-transferase-containing FLYWCH zinc finger protein* (*gfzf*) (we refer to this as *gfzf*) (12, 13).

The *gfzf* gene encodes two alternative transcripts. The longer transcript encodes a polypeptide with four FLYWCH zinc finger domains and one glutathione-S-transferase (GST) domain, whereas the shorter transcript encodes a polypeptide with only the GST domain. The *D. simulans* allele of *gfzf* (i.e., *gfzf<sup>sim</sup>*) incurred unique mutations (two nonsense, one frameshift, two deletions, and one missense mutation in a highly conserved residue) in each of the six rescue hybrid F<sub>1</sub> males (Fig. 1D and table S2). Four of these mutations only disrupt the longer of the two alternate transcripts encoded by *gfzf* (Fig. 1D and table S2). These results suggest that the longer *gfzf<sup>sim</sup>* transcript is involved in hybrid incompatibility. None of the rescue hybrid F<sub>1</sub> males we collected had mutations in the *Lhr* gene, which suggested that our genetic screen did not achieve saturation. We attribute this to the fact that the coding sequence of *Lhr<sup>sim</sup>* [1188 base pairs (bp)] is smaller and may present a smaller mutagenesis target than *gfzf<sup>sim</sup>* (3117 bp).

<sup>1</sup>Department of Biology, University of Utah, Salt Lake City, UT 84112, USA. <sup>2</sup>Basic Sciences Division, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA. <sup>3</sup>Genome Sciences, University of Washington, Seattle, WA 98195, USA. <sup>4</sup>Howard Hughes Medical Institute, University of Washington, Seattle, WA 98195, USA. <sup>5</sup>Department of Human Genetics, University of Michigan, Ann Arbor, MI 48109, USA. <sup>6</sup>Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA 98109, USA.

\*Present address: Department of Genetics, University of Wisconsin-Madison, Madison, WI 53706, USA. †Corresponding author. E-mail: nitin.phadnis@utah.edu (N.P.); hsmalik@fhcr.org (H.S.M.)

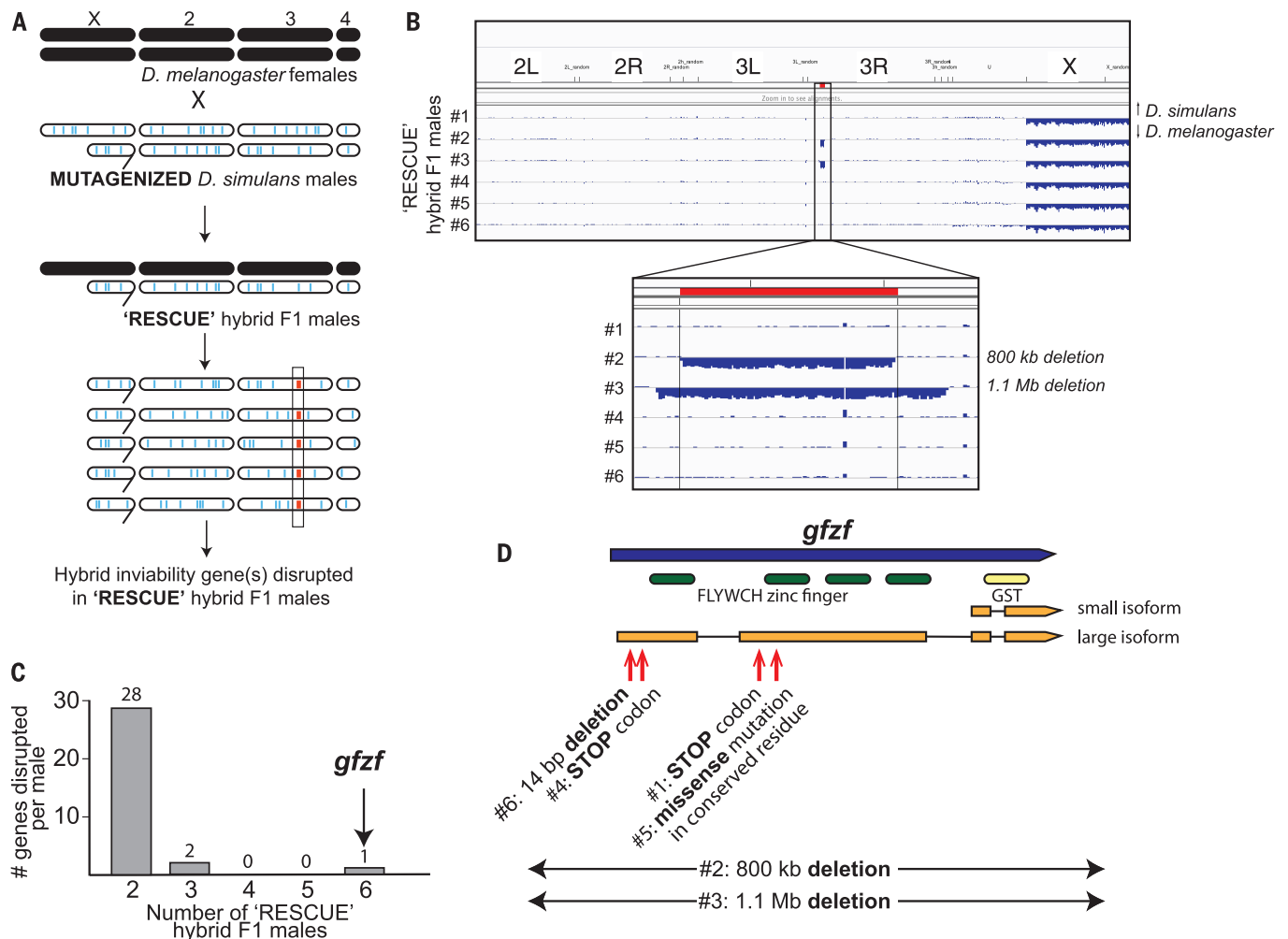
Consistent with our predictions (6), *gfzf<sup>sim</sup>* resides on the *D. simulans* third chromosome and is essential for viability (13). To circumvent the difficulty of testing the contribution of an essential gene in hybrid inviability, we depleted (knocked down) the expression of the *gfzf<sup>sim</sup>* longer transcript in F<sub>1</sub> hybrids using RNA interference (RNAi) knockdown constructs (*pValium20-gfzf<sup>sim</sup>*) that target only *gfzf<sup>sim</sup>* but not *gfzf<sup>mel</sup>* (11) (figs. S4 and S5). We produced transgenic *D. melanogaster* strains that carry these constructs under the control of the inducible promoter upstream activating sequence (UAS), inserted on the *D. melanogaster* X-chromosome (fig. S6).

We crossed these transgenic flies to a heterozygous *D. melanogaster* strain carrying a *CyO* balancer and a ubiquitously expressing *GAL4* driver *P{Act5C-GAL4}* on the second chromosome. This cross produces two types of daughters.

The first set inherits the *CyO* balancer chromosome but not the *Actin5C-GAL4* driver and, therefore, does not express the knockdown construct. When these *D. melanogaster* females are crossed to *D. simulans* males, the resulting hybrid F<sub>1</sub> males are inviable, as expected, because they do not express the knockdown construct (“RNAi off,” Fig. 2A). The second set of daughters inherits one copy each of the *Act5C-GAL4* driver and the RNAi construct. In crosses between these *D. melanogaster* females and *D. simulans* males, one out of four hybrid F<sub>1</sub> sons inherit both the RNAi construct and the *Act5C-GAL4* driver and therefore robustly express the RNAi construct. We found that these hybrid F<sub>1</sub> males are viable (“RNAi on,” Fig. 2A and figs. S7 to S9). Thus, knocking down the expression of only the long transcript of *gfzf<sup>sim</sup>* in hybrid F<sub>1</sub> males is sufficient to reverse hybrid inviability. These re-

sults confirm that *gfzf<sup>sim</sup>* is the missing hybrid incompatibility gene.

In contrast to our results with *gfzf<sup>sim</sup>* knockdown, we found that disrupting *gfzf<sup>mel</sup>* does not rescue the viability of F<sub>1</sub> hybrid males (table S3). Thus, allelic differences between *gfzf<sup>mel</sup>* and *gfzf<sup>sim</sup>* are important for hybrid inviability as are both *Hmr* and *Lhr*, with the limitation of comparing results between *gfzf<sup>sim</sup>* knockdown and *gfzf<sup>mel</sup>* disruption. Because positive selection likely resulted in the divergent functional properties of *Hmr* and *Lhr* orthologs (9, 14), we tested whether *gfzf* had also been subject to positive selection. We obtained *gfzf* sequences from 9 *D. melanogaster* and 13 *D. simulans* strains (table S4). Using a McDonald-Kreitman test and an outgroup species *D. yakuba*, we found that an excess of fixed nonsynonymous changes had occurred leading up to the hybrid inviability-associated *gfzf<sup>sim</sup>*,



**Fig. 1. A genomics screen identifies *gfzf<sup>sim</sup>* as a hybrid incompatibility gene.**

(A) We mutagenized *D. simulans* males (new mutations shown in blue) and crossed them with *D. melanogaster* females. When a *D. simulans* sperm carrying a mutation at a hybrid incompatibility gene fertilizes a *D. melanogaster* egg, a viable “rescue” hybrid F<sub>1</sub> male is produced. Sequencing the genomes of multiple “rescue” hybrid males identify the causative hybrid incompatibility gene mutated across these rescue males (shown in red and outlined). (B) Single-fly genome sequencing of all “rescue” hybrid males allows assignment

of new mutations (including large deletions in two of the males) to the *D. simulans*-derived component of hybrid genomes. (C) A single gene, *gfzf<sup>sim</sup>*, is mutated across all six “rescue” hybrid F<sub>1</sub> males. The x axis represents the number of genes mutated across all six males, any five males, and so on. The y axis represents the number of genes mutated across these males. (D) *gfzf* encodes two alternative transcripts. The larger transcript encodes FLYWCH zinc finger domains along with a GST domain, whereas the shorter transcript encodes only the GST domain.

especially in the FLYWCH zinc finger domains (fig. S10). In contrast, we found no evidence for positive selection along the *D. melanogaster* lineage.

Although our results demonstrate the role of *gfzf<sup>sim</sup>* in hybrid male inviability between *D. melanogaster* and *D. simulans*, previous studies have found that *gfzf<sup>mel</sup>* also affects dominant genetic incompatibility between strains of *D. melanogaster* (13). Indeed, *gfzf<sup>mel</sup>* plays an essential role in potentiating inviability seen in crosses between *D. melanogaster* females homozygous for the eye color mutation *prune* (*pn*), and *D. melanogaster* males carrying *Killer-of-prune* (*Kpn*) (13) [hence, the name *Su(Kpn)*, or *Suppressor of Killer of prune*]. The essential, dominant role of *gfzf* in lethal incompatibilities within and between species suggests that there may be limited genetic paths to the evolution of dominant lethal incompatibilities.

The ability of *gfzf* in mediating dominant lethal incompatibilities may stem from its role in the

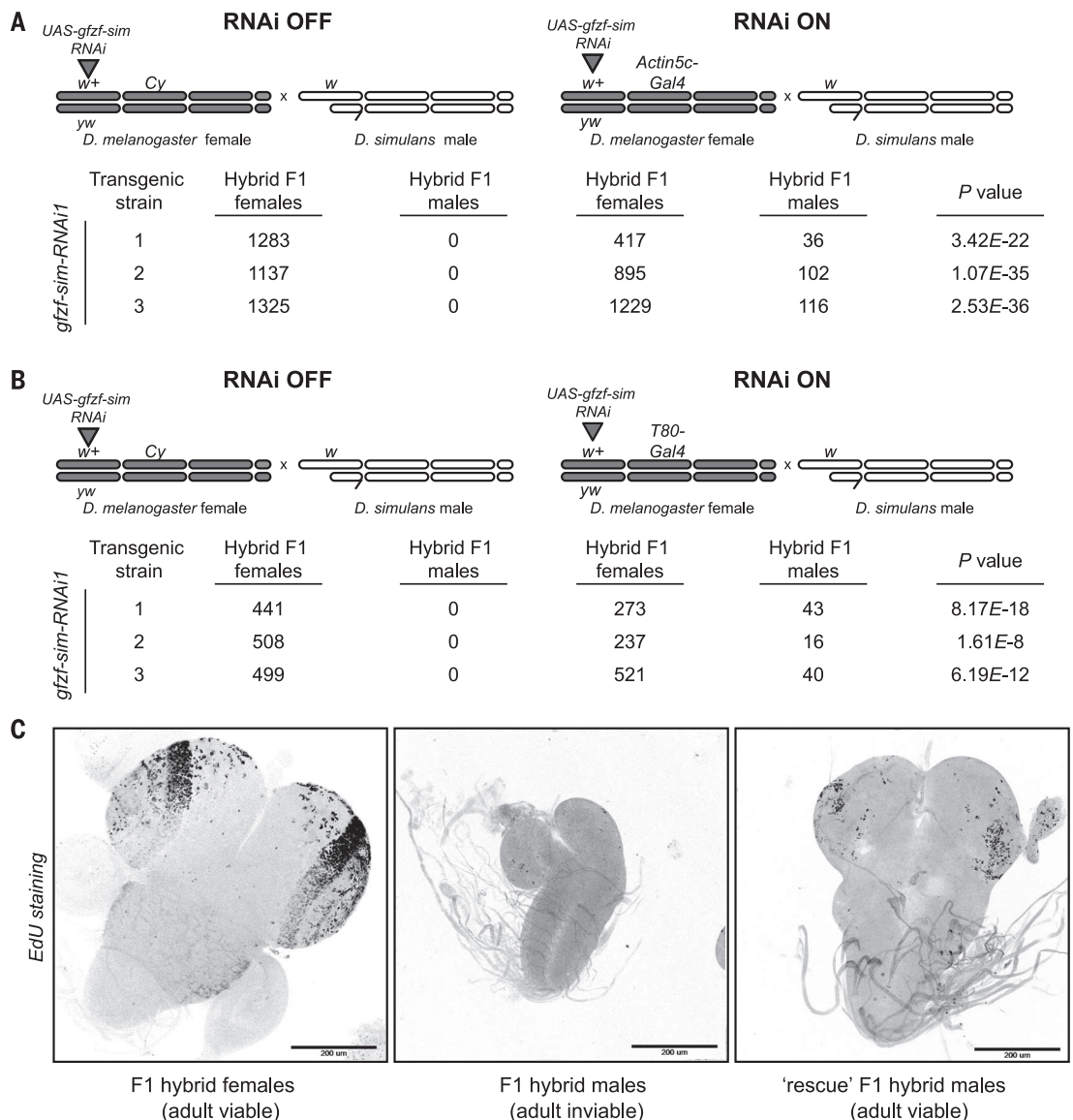
DNA damage-induced G<sub>2</sub>-M cell cycle-checkpoint mechanism, where it can potentiate the dE2F2/RBF pathway to block cell proliferation (15, 16). In contrast, *gfzf* has also shown to be required for cell proliferation by transcriptionally regulating the RAS/MAPK pathway (17). Despite its essential role in both cell cycle arrest and regulation of cell proliferation, the precise molecular function of *gfzf* is still uncharacterized. Moreover, the biological consequence of *gfzf<sup>sim</sup>* activity on hybrid male viability is unknown. Nevertheless, the developmental timing and consequences of either *gfzf* deficiency or *gfzf*-mediated dominant lethality are suggestive of a common mechanism that manifests in the larval-pupal transition. Larval tissues in *Drosophila* mostly consist of polyploid cells, whereas the larval nervous system and imaginal discs are composed of diploid cells. During pupation, the polyploid tissues are degraded, and the diploid imaginal discs proliferate to produce the adult body form. Individuals that lack proper imaginal discs can survive and

continue to grow as larvae but die during the larval-pupal transition. Notably, homozygous *gfzf<sup>mel</sup>* null mutants, *gfzf<sup>mel</sup>-pn-Kpn* males, and *gfzf<sup>sim</sup>*-expressing hybrid males, all lack imaginal discs and die as larvae (13, 18, 19). This phenotype of lethality during the larval-pupal transition along with an absence of imaginal discs is diagnostic of dysfunction in cell cycle-regulation mechanisms (20).

We hypothesized that the *gfzf<sup>sim</sup>*-associated hybrid male lethality was due to cell proliferation defects in hybrid larvae. We therefore drove the expression of our *gfzf<sup>sim</sup>* knockdown construct in hybrid F<sub>1</sub> males with a *T80-GAL4* driver, which is expressed ubiquitously in early embryonic stages but specifically in the nervous system and imaginal discs in late larval stages (21). We found that *T80-GAL4* mediated *gfzf<sup>sim</sup>* knockdown robustly rescued the viability of hybrid F<sub>1</sub> males to adulthood (Fig. 2B and fig. S11). This result suggests that the primary defect in hybrid F<sub>1</sub> males produced in *D. melanogaster* × *D. simulans* crosses

**Fig. 2. Knockdown of *gfzf<sup>sim</sup>* rescues cell proliferation defects and restores hybrid male viability.** (A) No hybrid males are recovered in crosses where the

*pValium20-gfzf<sup>sim</sup>* RNAi construct is not expressed (no GAL4 driver, “RNAi OFF”). In crosses between *D. simulans* males and *D. melanogaster* females carrying one copy each of *pValium20-gfzf<sup>sim</sup>* and a ubiquitously expressing Actin5C-GAL4 driver, one out of four possible hybrid male progeny inherit both *pValium20-gfzf<sup>sim</sup>* and the Actin5C-GAL4 driver (“RNAi ON”) and produce viable F<sub>1</sub> hybrid male progeny. *P* values were calculated using Fisher’s exact test. (B) RNAi knockdown of *gfzf<sup>sim</sup>* by a T80-GAL4 driver, more specific to larval neuroblasts and imaginal discs, successfully restores the viability of F<sub>1</sub> male hybrids. (C) EdU staining shows the diminutive larval brains and cell proliferation defects in “inviable” hybrid males compared with viable F<sub>1</sub> hybrid female larvae. These cell proliferation defects are also partially rescued in hybrid males upon *gfzf<sup>sim</sup>* knockdown.



may be in diploid tissue proliferation. Indeed, previous studies on larval brains have shown both cell cycle arrest and profound mitotic defects in hybrid F<sub>1</sub> male larvae. These larvae also display diminutive imaginal discs and reduced larval brain sizes due to cell cycle arrest (22, 23). Using 5-ethynyl-2'-deoxyuridine (EdU) to track DNA synthesis in proliferating cells, we found that cell proliferation is restored in larval brains from hybrid F<sub>1</sub> males upon *gfzf<sup>sim</sup>* knockdown, which indicates a relief from cell cycle arrest (Fig. 2C and fig. S12). Thus, *gfzf<sup>sim</sup>* knockdown relieves both cell cycle arrest and hybrid F<sub>1</sub> male inviability. Together, these results support the model that *gfzf* is a cell cycle regulator of diploid tissues in larvae. Furthermore, they implicate the arrest of cell proliferation as the cause of hybrid F<sub>1</sub> male inviability at this late-larval stage of *Drosophila* development.

Although Hmr and Lhr physically interact with each other (9), there is no evidence of a direct physical interaction between *gfzf* and either Hmr or Lhr. Both Hmr and Lhr proteins localize to centromeres and pericentric heterochromatin, where they play a role in mitotic chromosome segregation (24) and the suppression of transposable elements (25). These findings have led to a model in which incompatibility between *Lhr<sup>sim</sup>* and *Hmr<sup>mel</sup>* and their expression levels may cause dysfunction at centromeres or pericentric heterochromatin (24). Although the molecular nature of this dysfunction is still unclear, we speculate that the direct engagement of *gfzf<sup>sim</sup>* arrests the proliferation of dysfunctional diploid imaginal discs, which leads to hybrid inviability. Under this scenario, ablation of *Lhr<sup>sim</sup>* or *Hmr<sup>mel</sup>* removes the primary dysfunction, whereas ablation of *gfzf<sup>sim</sup>* removes the cell cycle arrest. Alternatively, *gfzf<sup>sim</sup>* may act indirectly by contributing to the sensitization of the hybrid genetic background, making it susceptible to the defects caused by the incompatibility between *Lhr<sup>sim</sup>* and *Hmr<sup>mel</sup>* that leads to hybrid

inviability. In both scenarios, removal of any one of these three genes would restore hybrid viability. Thus, the same checkpoints that normally ensure the correction of mitotic errors may be also responsible for the inviability of hybrid males in the *D. melanogaster* × *D. simulans* interspecies cross.

The discovery of hybrid rescue genes, with mutations that reverse hybrid sterility or inviability, has meaningfully advanced our understanding of the genetic mechanisms that underlie the evolution of reproductive isolation during or following speciation. The identification of *gfzf*, in particular, emphasizes the role of cell cycle-regulation mechanisms in the evolution of hybrid incompatibilities (22, 23) and the complex epistatic interactions that underlie dominant hybrid incompatibilities in F<sub>1</sub> hybrids. Our genomics-based approach may also allow mapping of genes that underlie hybrid incompatibilities and other phenotypes even when they lie within chromosomal inversions, which often impedes their precise genetic identification. Although this method requires that there be a single incompatibility separating two species, recently diverged species are likely to meet this criterion (26). Our approach may help accelerate the discovery of genes and genetic mechanisms underlying hybrid dysfunction in multiple taxa, shedding light on how reproductive isolation evolves.

#### REFERENCES AND NOTES

1. L. S. Quackenbush, *Science* **32**, 183–185 (1910).
2. A. H. Sturtevant, *Genetics* **5**, 488–500 (1920).
3. D. A. Barbash, *Genetics* **186**, 1–8 (2010).
4. W. B. Provine, *Genetics* **129**, 1–5 (1991).
5. H. J. Muller, G. Pontecorvo, *Nature* **146**, 199–200 (1940).
6. G. Pontecorvo, *J. Genet.* **45**, 51–66 (1943).
7. D. A. Barbash, D. F. Siino, A. M. Tarone, J. Roote, *Proc. Natl. Acad. Sci. U.S.A.* **100**, 5302–5307 (2003).
8. P. Hutter, M. Ashburner, *Nature* **327**, 331–333 (1987).
9. N. J. Brideau *et al.*, *Science* **314**, 1292–1295 (2006).
10. T. K. Watanabe, *Jpn. J. Genet.* **54**, 325–331 (1979).
11. Materials and methods are available as supplementary materials on Science Online.
12. M.-S. Dai, X.-X. Sun, J. Qin, S. M. Smolik, H. Lu, *Gene* **342**, 49–56 (2004).
13. E. Provost *et al.*, *Genetics* **172**, 207–219 (2006).
14. S. Maheshwari, J. Wang, D. A. Barbash, *Mol. Biol. Evol.* **25**, 2421–2430 (2008).
15. A. M. Ambrus, V. I. Rasheva, B. N. Nicolay, M. V. Frolov, *Genetics* **183**, 79–92 (2009).
16. S. Kondo, N. Perrimon, *Sci. Signal.* **4**, rs1 (2011).
17. D. Ashton-Beaucage *et al.*, *PLOS Biol.* **12**, e1001809 (2014).
18. L. Sanchez, A. Dubendorfer, *Roux's Arch. Dev. Biol.* **192**, 48 (1983).
19. T. Seiler, R. Nothiger, *Experientia* **30**, 709 (1974).
20. M. Gatti, B. S. Baker, *Genes Dev.* **3**, 438–453 (1989).
21. B. J. Loveall, D. L. Deitcher, *BMC Dev. Biol.* **10**, 92 (2010).
22. B. J. Bolkan, R. Booker, M. L. Goldberg, D. A. Barbash, *Genetics* **177**, 2233–2241 (2007).
23. H. A. Orr, L. D. Madden, J. A. Coyne, R. Goodwin, R. S. Hawley, *Genetics* **145**, 1031–1040 (1997).
24. A. W. Thomae *et al.*, *Dev. Cell* **27**, 412–424 (2013).
25. P. R. V. Satyaki *et al.*, *PLoS Genet.* **10**, e1004240 (2014).
26. N. Phadnis, *Genetics* **189**, 1001–1009 (2011).

#### ACKNOWLEDGMENTS

We thank T. Levin, M. Levine, M. Patel, Naina Phadnis, B. Ross, and S. Zanders for their comments on the manuscript and members of the Malik, Peichel, and Phadnis labs for discussions. This work was supported by a Howard Hughes Medical Institute–Life Sciences Research Foundation (HHMI-LSRF) postdoctoral fellowship (N.P.); a Mario Cappechi–endowed assistant professorship (N.P.); an NIH Developmental Biology Training Grant 5T32 HD0741 (J.C.C.); graduate research fellowship DGE-071824 from the NSF (J.O.K.); grants from the NIH: R01 GM115914 (N.P.), HG006283 (J.S.), and R01 GM074108 (H.S.M.); and especially a grant from the Mathers Foundation (H.S.M.). J.S. and H.S.M. are investigators of HHMI. The data reported in this paper are tabulated in the supplementary materials and archived at the following databases: primary read data at the Sequence Read Archive under accession number PRJNA290665 and *gfzf* sequences at GenBank under accession numbers KU064780 to KU064795.

#### SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/350/6267/1552/suppl/DC1  
Materials and Methods  
Tables S1 to S4  
Figs. S1 to S12  
References (27–32)

8 June 2015; accepted 13 November 2015  
10.1126/science.aac7504