

Drosophila melanogaster and *D. simulans* rescue strains produce fit offspring, despite divergent centromere-specific histone alleles

A Sainz^{1,3}, JA Wilder^{2,3}, M Wolf² and H Hollocher¹

¹Department of Biological Sciences, University of Notre Dame, Notre Dame, IN 46556, USA; ²Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08544, USA

The interaction between rapidly evolving centromere sequences and conserved kinetochore machinery appears to be mediated by centromere-binding proteins. A recent theory proposes that the independent evolution of centromere-binding proteins in isolated populations may be a universal cause of speciation among eukaryotes. In *Drosophila* the centromere-specific histone, Cid (centromere identifier), shows extensive sequence divergence between *D. melanogaster* and the *D. simulans* clade, indicating that centromere machinery incompatibilities may indeed be involved in reproductive isolation and speciation. However, it is presently unclear whether the adaptive evolution of Cid was a cause of the divergence between these species, or merely a product of postspeciation adaptation in the separate lineages. Furthermore, the extent to which divergent centromere

identifier proteins provide a barrier to reproduction remains unknown. Interestingly, a small number of rescue lines from both *D. melanogaster* and *D. simulans* can restore hybrid fitness. Through comparisons of *cid* sequence between nonrescue and rescue strains, we show that *cid* is not involved in restoring hybrid viability or female fertility. Further, we demonstrate that divergent *cid* alleles are not sufficient to cause inviability or female sterility in hybrid crosses. Our data do not dispute the rapid divergence of *cid* or the coevolution of centromeric components in *Drosophila*; however, they do suggest that *cid* underwent adaptive evolution after *D. melanogaster* and *D. simulans* diverged and, consequently, is not a speciation gene.

Heredity (2003) 91, 28–35. doi:10.1038/sj.hdy.6800275

Keywords: *Drosophila*; speciation; centromere identifier; rescue

Introduction

Broadly applicable models for the origin of species are fundamental to understanding widespread evolutionary phenomena. The process of speciation often involves the development of reproductive isolation in the form of hybrid sterility and inviability. Recently, coevolution between centromere DNA sequences and centromere-binding proteins has been proposed as a mechanism for causing reproductive isolation, such that 'speciation is an inevitable consequence of centromere evolution (Henikoff *et al.*, 2001a)'. At the crux of this model is the idea that as centromere DNA sequences change, centromere-binding proteins must adapt to remain compatible with the new centromeres. These components must function together effectively in order to facilitate exact chromosome separation during cell division, a process that leads to the independent and rapid coevolution of centromeres and centromere-binding proteins in genetically nonmixing populations (Henikoff *et al.*, 2001a). It has been proposed that this coevolution will ultimately lead to incompatibilities in the form of inviability or sterility in

hybrids from populations with divergent centromere machinery (Henikoff *et al.*, 2001a). Corroboration for this 'centromere-drive' model appears to exist among *Drosophila*, where the main centromere-binding protein, Cid (centromere identifier), has been demonstrated to have undergone strong and recurrent adaptive evolution, presumably in response to changes in centromeric satellite sequences (Malik and Henikoff, 2001; Malik *et al.*, 2002). However, the centromere-drive mechanism, which has the potential to act as a universal cause of divergence in almost any incipient eukaryotic species, has yet to be demonstrated to have any direct influence on naturally occurring patterns of reproductive isolation.

The evolution of centromere machinery is driven by the need for precise chromosome segregation. This process requires proper assembly and operation of the kinetochore, a structure comprised of regulatory and associated proteins and centromere-specific histones that forms at the centromere DNA. While the kinetochore apparatus is highly conserved (Tyler-Smith and Florida, 2000; Kitagawa and Hieter, 2001), centromere satellite sequences have ubiquitous rapid evolution (Haaf and Willard, 1997; Murphy and Karpen, 1998). The contrast between these slowly and rapidly evolving portions of the centromere has been referred to as the 'centromere paradox' (Henikoff *et al.*, 2001a). Centromere-specific histones are responsible for facilitating the interaction between the stable kinetochore components and the

Correspondence: H Hollocher, 148 Galvin Life Sciences, University of Notre Dame, Notre Dame, IN 46556, USA.

E-mail: Hope.Hollocher.1@nd.edu

³Contributed equally to this paper

Received 4 September 2002; accepted 18 February 2003

highly variable centromere sequences (Henikoff *et al*, 2001a). As such, centromeric histones are thought to be coevolving in response to changes in the rapidly evolving centromere DNA sequences. Without this form of coevolution, chromosomes may fail to segregate properly, thus disrupting cell division.

In addition to their role maintaining functional histone-centromere relationships and proper cell division, Henikoff *et al* (2001a) also propose that centromeric histones may experience additional selective pressure due to chromosomal competition during female meiosis. In organisms with asymmetric female meiosis, chromosomes may compete for inclusion into the single meiotic product (of the four produced) that becomes an oocyte. If there are multiple alleles of the centromere-specific histone, then the one with the strongest centromere relationship will be preferentially included over those with weaker interactions, providing a powerful source of positive selection on the centromeric histones. In males, selection may also exist to minimize centromeric imbalances that might lead to nondisjunction during sperm production. Since the coevolution of centromeres and their corresponding histones will follow separate evolutionary trajectories in populations with limited gene flow, Henikoff *et al* (2001a) suggest that the centromeric components may develop incompatibilities sufficient to cause reproductive isolation through both hybrid inviability and hybrid sterility.

To examine the role of centromere evolution in speciation, we must look at the divergence of centromere-specific histones in closely related species and the relationship of these proteins to patterns of reproductive isolation. In *Drosophila* the critical centromere-specific histone gene *cid* has been sequenced from a number of species and is thought to have been evolving adaptively in numerous lineages for at least 25 million years (Malik *et al*, 2002). A dramatic recent round of selection has

driven the adaptive divergence of this gene in the sister taxa *Drosophila simulans* and *D. melanogaster* (Malik and Henikoff, 2001). The two domains of the *cid* gene that have undergone strong adaptive evolution, the N-terminal tail and the histone fold, are both putative DNA-binding regions (Malik and Henikoff, 2001). The fact that the *cid* locus is most rapidly evolving in these critical functional domains underscores the potential importance of coevolution between Cid and centromeric DNA sequences (Malik and Henikoff, 2001). However, these data do not necessarily support the hypothesis that the coevolution of centromeric histones and centromeric DNA is responsible for the development of reproductive isolation between *D. melanogaster* and *D. simulans*.

In order to test the relationship between reproductive isolation and *cid*, we have further examined the evolution of *cid* in a survey of numerous strains of *D. melanogaster* and *D. simulans*. Initial studies of these species by Malik and Henikoff (2001) focused on strains that exhibit reproductive isolation in the form of both inviable and sterile hybrids, depending on the direction of the cross (see Table 1). However, several 'rescue strains' of each species exist that restore hybrid viability and female fertility when crossed in specific combinations, as shown in Table 1 (Lachaise *et al*, 1986; Hutter *et al*, 1990; Sawamura *et al*, 1993; Davis *et al*, 1996). In the majority of these strains, the genetic basis of hybrid rescue is very poorly understood (Hollocher, 1998; Hollocher *et al*, 2000). While most of these strains have been previously described, here we present three new African *D. simulans* lines with unique rescue phenotypes (a strain that completely rescues female fertility, a partial fertility rescue strain, and a strain that refuses to be rescued by a *D. melanogaster* inviability rescue strain).

We have sequenced the *cid* locus from each of the nine rescue and partial strains included in Table 2, as well as

Table 1 Rescue and nonrescue phenotypes of *D. melanogaster*/*D. simulans* hybrids

Cross (female × male)	Female F1 hybrids	Male F1 hybrids
Wild type		
<i>D. simulans</i> × <i>D. melanogaster</i>	Inviability (embryonic)	Viable; sterile
<i>D. melanogaster</i> × <i>D. simulans</i>	Viable; sterile	Inviability (larval)
Rescue		
<i>D. simulans</i> × <i>In(1)AB, f-M1</i> (mel)	Viable ; sterile → fertile ^a	Viable; sterile
<i>In(1)AB</i> (mel) × <i>D. simulans</i>	Viable; sterile → fertile ^a	Viable ; sterile
C167.4 (sim) × <i>In(1)AB, f-M1</i> (mel)	Viable; fertile	Viable; sterile
<i>In(1)AB</i> (mel) × C167.4 (sim)	Viable; fertile	Viable ; sterile
MAZ1 (sim) × <i>In(1)AB, f-M1</i> (mel)	Viable; fertile	Viable; sterile
<i>In(1)AB</i> (mel) × MAZ1 (sim)	Viable; fertile	Viable ; sterile
C167.4 (sim) × Melbourne (mel)	Viable ; sterile	Viable; sterile
<i>D. simulans</i> × Tai 255.1 (mel)	Viable ; sterile	Viable; sterile
Mhr (sim) × <i>D. melanogaster</i>	Viable ; sterile	Viable; sterile
Refuses rescue		
VF11 (sim) × <i>In(1)AB, f-M1</i> (mel)	Low viability ; sterile	Viable; sterile
<i>In(1)AB</i> (mel) × VF11 (sim)	Viable; sterile	Viable ; sterile

Bold type indicates a rescue strain and the phenotype that the strain is rescuing. Underline indicates a strain that refuses to be rescued and the phenotype that is not being rescued. Data in table from Lachaise *et al* (1986), Hutter *et al* (1990), Sawamura *et al* (1993), Davis *et al* (1996), Carracedo *et al* (2000), and presented in this paper.

^aFemale hybrid offspring cover a spectrum ranging from sterile to fertile depending on the strain of *D. simulans* used in the cross.

Table 2 *D. melanogaster* and *D. simulans* strains

Species	Strain	Rescue	Partial rescue	Nonrescue	Rescue refusal
<i>D. melanogaster</i>	<i>In(1)AB</i> and <i>In(1)AB, f-M1</i>	X (viability)			
<i>D. melanogaster</i>	Melbourne	X (viability)			
<i>D. simulans</i>	<i>mhr</i>	X (viability)			
<i>D. melanogaster</i>	Tai 255.1	X (fertility)			
<i>D. simulans</i>	C167.4	X (fertility)			
<i>D. simulans</i>	MAZ1	X (fertility)			
<i>D. simulans</i>	<i>vermillion</i>		X (fertility)		
<i>D. simulans</i>	Oxnard		X (fertility)		
<i>D. simulans</i>	MAZ6		X (fertility)		
<i>D. melanogaster</i>	Oregon R			X ^a	
<i>D. melanogaster</i>	Antigua			X	
<i>D. simulans</i>	Tsimbazaza			X	
<i>D. simulans</i>	Wild type			X	
<i>D. simulans</i>	VF11				X ^b

Data in the table from Lachaise *et al* (1986), Hutter *et al* (1990), Sawamura *et al* (1993), Davis *et al* (1996), and presented in this paper.

^aRescues female viability only at low temperatures (18°C).

^bRefuses to be rescued, has low viability of females when *In(1)AB, f-M1* is the father and has no sterility rescue.

from several nonrescue and rescue-refusal strains. Since the model of centromere-driven speciation predicts that reproductive isolation may be a product of incompatibilities between centromeres and centromeric histones (*Cid* in the case of *Drosophila*), we have examined whether hybrid inviability or sterility rescue appears to be related to *cid* sequence. In addition to the evidence of recent adaptive evolution of *cid* among *D. simulans* and *D. melanogaster*, patterns of hybrid inviability and sterility also indicate that centromeric incompatibilities may play a role in the reproductive isolation of these two species. Hollocher *et al* (2000) have shown that hybrid female sterility between *D. simulans* and *D. melanogaster* appears to be caused by defects in mitotic control during early oogenesis. This defective phenotype is predicted by the centromere-drive model of speciation, which hypothesizes that improper chromosomal segregation can disrupt cell division in hybrids with mismatched centromeric machinery (Henikoff *et al*, 2001a). Thus, in addition to *Cid* acting as a possible barrier to reproduction between these species, it is possible that the rescue phenotype is caused by *Cid* alleles that allow for increased hybrid fitness. These alleles may be ancestral, occurring in the populations by introgression or incomplete lineage sorting, or novel *Cid* alleles may have evolved that are compatible with divergent centromere structures. Alternatively, *Cid* alleles in the rescue strains may not be different from nonrescue strains, indicating that *Cid* is not involved in hybrid rescue, and, more importantly, that the presence of divergent centromeric machinery does not preclude the production of viable and fertile offspring.

Materials and methods

Drosophila strains

The *D. simulans* strains C167.4, Tsimbazaza, S-17 *vermillion*, and Oxnard and the *D. melanogaster* strains *In(1)AB*, *In(1)AB, f-M1/C(1)M4,y* (a recombinant derivative of *In(1)AB* with males that carry the forked bristle,

abbreviated as *In(1)AB, f-M1* in this paper), Antigua, and Oregon-R are all described in Davis *et al* (1996). The *D. simulans* strain wild type is described in Johnson and Wu (1992). The *D. simulans* strains MAZ1, MAZ6, and VF11 are from our laboratory stocks and were collected by Andrew Dubill in Zimbabwe (Dubill, 1996). Finally, the *D. melanogaster* strains Tai 255.1 and Melbourne and the *D. simulans* strain *mhr* were kindly provided by John Roote. Tai 255.1 is described in Lachaise *et al* (1986), Melbourne is described in Davis *et al* (1996), and *mhr* is described in Sawamura *et al* (1993).

Survey of *D. simulans* strains

We examined the *D. simulans* strains MAZ1, MAZ6, and VF11 in order to identify novel strains with rescue phenotypes of interest. Male flies from survey lines (including MAZ1 and VF11) and males from the inviability rescue line *D. melanogaster In(1)AB, f-M1* were collected as virgins and aged 5–7 days. Females from survey lines (including MAZ1, MAZ6, and VF11) and females from the inviability rescue line *In(1)AB* were collected as virgins and 20–30 were immediately mated with 60–80 aged males in vials with the cotton pushed halfway down on instant *Drosophila* medium (Carolina Biological Supply). F1 hybrids were collected as virgins. The males were aged 5–7 days and their testes were dissected. Males were considered sterile if their testes were atrophied and did not contain any active sperm. Females were aged 6–7 days and their ovaries were dissected. Females with ovaries that contained normal, fully developed eggs were considered fertile. Females with ovaries lacking germline or eggs at stage 8 or less were considered sterile.

Rescue and nonrescue strains

Certain combinations of rescue strains produce hybrids that have increased viability and/or fertility above the levels observed in crosses between nonrescue strains of *D. melanogaster* and *D. simulans* (Table 1). These mutations generally fall into two classes: those that restore the

viability of hybrid offspring, and those that restore fertility. Of the inviability rescue lines, the *D. melanogaster* strain *In(1)AB* (and the recombinant *In(1)AB, f-M1* strain) is the only one to rescue viability of both sexes (Hutter et al, 1990). Other strains rescue female viability; however, we limit our analysis to the lines Tai 255.1 (*D. melanogaster*), Melbourne (*D. melanogaster*), and *mhr* (*D. simulans*). In each of these strains, the rescue mutations are either unmapped (Tai 255.1 and Melbourne) or map to the region of chromosome 2 containing the *cid* gene (Lachaise et al, 1986; Sawamura et al, 1993; Davis et al, 1996; Carracedo et al, 2000). In two of the known lines where male viability is rescued (*Hmr* and *Lhr*), the mutations responsible have been mapped outside of the region containing *cid* and are excluded from this study (Watanabe, 1979; Hutter et al, 1990).

Of the fertility rescue lines, two *D. simulans* strains, C167.4 (Davis et al, 1996) and MAZ1 (data presented in this paper), have the ability to restore hybrid female fertility to normal levels, but this occurs only when crossed with the *D. melanogaster* rescue strain *In(1)AB, f-M1* males.

Other *D. simulans* lines form a continuum of fertility when crossed to viability rescue lines. For example, the strains *vermillion* and Oxnard produce approximately 50% females with eggs, when crossed as the mother to *In(1)AB, f-M1* males, while other strains produce substantially lower proportions (Davis et al, 1996). Also, MAZ6, when crossed with *In(1)AB, f-M1* males, produces a high number of fertile females but each ovary has less than five eggs (data presented in this paper). In all of the partial rescue crosses, the number of eggs per hybrid female is typically very low, indicating that the fertility rescue of these strains does not match that of the rescue strains C167.4 or MAZ1. In our sample we include C167.4 and MAZ1, for their remarkable fertility rescue ability, as well as other *D. simulans* lines that are known to fall along both the upper and lower margins of the fertility continuum. We also include the *D. simulans* strain VF11, which is unique in having no fertility rescue when crossed with *In(1)AB* females and refuses rescue of female viability when crossed with *In(1)AB, f-M1* males (data presented in this paper).

DNA preparation, cloning, and sequencing

DNA was extracted from single male flies from each strain using the squish technique of Gloor et al (1993). Given the inbred nature of the hybrid rescue lines, the fact that many were founded from only a single wild-caught female, and that most are standard laboratory strains that have been maintained in culture for hundreds of generations, it is unlikely that our choice of a single fly from each strain in any way influences the data. Amplification of the *cid* gene was performed using the primers Cid-upstream and Cid-downstream, as described in Malik and Henikoff (2001). Amplification products were cloned into the PCR2.1 vector using an Invitrogen TOPO TA Cloning Kit. A single clone from each strain was sequenced using an Applied Biosystems Big Dye Terminator Kit and an ABI 377 automated sequencer. The sequencing primers were CidmidRev and CidmidFor, again as described in Malik and Henikoff (2001). Sequences collected in this study were pooled with sequences from the nonrescue

strains examined by Malik and Henikoff (2001) in all analyses.

Sequence analysis

Sequences were aligned using the CLUSTAL W module (Thompson et al, 1994) of the program BioEdit 5.0.9 (Hall, 1999). Both neighbor-joining and maximum parsimony trees of nucleotide sequences were constructed using the PHYLIP phylogenetic inference package (Felsenstein, 2001). The neighbor-joining tree was estimated using Jukes–Cantor genetic distances. Confidence was assessed using 1000 bootstraps for each tree. Estimates of *Cid* sequence diversity in each species were calculated using the program DNAsp (Rozas and Rozas, 1999). All sequences have been deposited in Genbank under accession numbers AY126929–AY126942.

Results

Identification of rescue, partial rescue, and rescue-refusal *D. simulans* strains

Our survey of multiple *D. melanogaster* and *D. simulans* lines identified three new African *D. simulans* lines of interest: MAZ1, MAZ6, and VF11. MAZ1 females crossed with the *D. melanogaster* inviability rescue line *In(1)AB, f-M1* produced viable females. Furthermore, 97.3% of the female ovaries dissected had normal, fully developed eggs. In addition, 64.7% of the females examined had greater than 20 eggs per ovary, which is approximately the number present in nonhybrid female ovaries (Hollocher et al, 2000). The reciprocal cross of *In(1)AB* × MAZ1 produced females with normal eggs but most ovaries contained 10 eggs or less. Based on these data (shown in Table 3a and b), we consider MAZ1 a full rescue line because of the high levels of fertility seen in most of the female progeny. Interestingly, the only other *D. simulans* full rescue line that has been identified, C167.4, shows slightly lower levels of female fertility when crossed with *In(1)AB, f-M1* males (88%) and much lower levels when crossed with *In(1)AB* females (42%; Davis et al, 1996).

Only 32.0% of the MAZ6 × *In(1)AB, f-M1* hybrid progeny were females, indicating a somewhat reduced level of hybrid viability compared to full rescue crosses. Although 90.5% of the hybrid female ovaries contained normal eggs, each ovary contained at most five fully developed eggs. Based on these data (shown in Table 3a and b) we consider MAZ6 to be a partial rescue line because while most of the hybrid females were fertile their ovaries contained a dramatically depleted number of eggs. The other previously identified partial rescue *D. simulans* strains Oxnard and *vermillion* had lower levels of female fertility when crossed with *In(1)AB, f-M1* males (56 and 57% respectively), but both produced some female hybrids with more than five normal eggs per ovary (Davis et al, 1996).

VF11 females crossed with *In(1)AB, f-M1* males produced a very low number of female progeny (8.6%) and all of them had atrophied ovaries that lacked germline. The reciprocal cross of *In(1)AB* × VF11 produced female hybrids that had atrophied ovaries without germline as well, although male viability was restored. We classify VF11 as a rescue-refusal strain because female viability is not restored when it is crossed with the inviability rescue strain *In(1)AB, f-M1* (Table 3). Not

Table 3 Hybrid viability analysis of MAZ1, MAZ6, and VF11 cross progeny

	MAZ1 × <i>In(1)AB, f-M1</i>	<i>In(1)AB</i> × MAZ1	MAZ6 × <i>In(1)AB, f-M1</i>	VF11 × <i>In(1)AB, f-M1</i>	<i>In(1)AB</i> × VF11
(a)					
No. of crosses	4	4	5	4	2
Male F1	262	200	327	222	218
Female F1	183	241	154	21	184
% Male F1	58.9%	45.4%	68.0%	91.4%	54.2%
% Female F1	41.1%	54.6%	32.0%	8.6%	45.8%
(b)					
Females dissected	150	64	74	18	43
No germline/atrophied	0	6 (9.4%)	4 (5.4%)	18 (100%)	40 (93.0%)
Up to stage 8 eggs	4 (2.7%)	6 (9.4%)	3 (4.1%)	0	3 (7.0%)
10 or less normal eggs	16 (10.7%)	30 (46.9%)	67 (90.5%) ^a	0	0
10–20 normal eggs	33 (22.0%)	18 (28.1%)	0	0	0
20 or more normal eggs	97 (64.6%)	4 (6.2%)	0	0	0

All crosses are written female × male. Data for male F1 progeny in (b) are not listed as all were sterile in every cross.

^aAll females had five normal eggs or less per ovary.

only does VF11 resist the rescue of female viability, but also sterility is ubiquitous in the female progeny of both crosses (Table 3a and b).

Genealogy of *cid* sequences

Phylograms showing the neighbor-joining and maximum parsimony trees for the *cid* sequences are shown in Figure 1. These trees recover extremely similar topologies, with identical clades receiving high bootstrap support in both trees. From these data it is clear that *cid* sequences group by taxa, regardless of the rescue phenotype of the individual strain. Three clear clades are visible: the first containing *cid* from *D. tesseleri*, which is relatively distantly related to *D. melanogaster* and *D. simulans*; a second clade containing all *D. melanogaster* *cid* sequences; and a third clade containing all *D. simulans* *cid* sequences, as well as those from *D. sechellia* and *D. mauritiana*. The sequence from *D. sechellia* is paraphyletic to the *D. simulans* *cid* sequences, while *D. mauritiana* is weakly supported as ancestral to *D. simulans*.

While there is relatively little structure within the *D. melanogaster* clade of *cid* sequences, the *D. simulans* clade contains two internal clades that have high bootstrap support. The first of these contains the strains C167.4, *vermillion*, and Oxnard. The second clade contains MAZ1, VF11, sim00, sim01, and sim09.

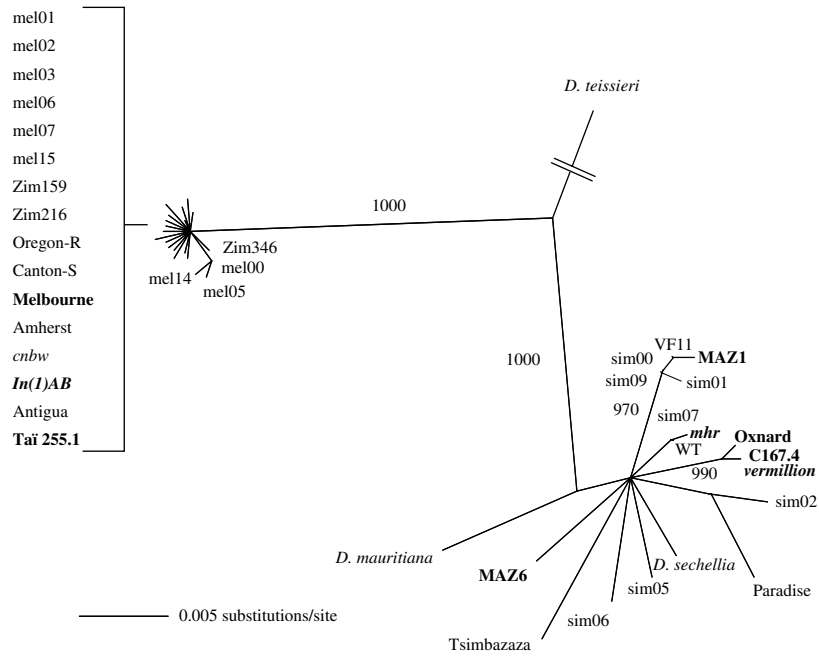
Levels of polymorphism at the *cid* locus were dramatically different between *D. simulans* and *D. melanogaster*. *D. simulans* had a relatively high level of variability, with 39 segregating sites and a nucleotide heterozygosity (π) value of 0.0163 (SD = 0.0013). In contrast, *D. melanogaster* had a very low level of nucleotide variability at the *cid* locus, with only 14 segregating sites and a π value of 0.0046 (SD = 0.0008). This difference in variability was also apparent at the amino-acid level. While the *D. simulans* alleles contained 13 nonsynonymous polymorphisms, those from *D. melanogaster* contained only two. Furthermore, these two nonsynonymous changes in *D. melanogaster* were confined to only a single strain, Tai 255.1.

Discussion

Our results show that there are no significant differences between *cid* alleles in rescue versus nonrescue strains in either *D. simulans* or *D. melanogaster*. In both species, a phylogenetic analysis of the data (Figure 1) clearly indicates that rescue strains are most similar to conspecific nonrescue strains. Thus, we can definitively rule out the hypothesis that any form of rescue of hybrid fitness is caused by the introgression of *cid* alleles across the *D. simulans*/*D. melanogaster* species boundary. Additionally, it is apparent that rescue lines do not contain novel *cid* alleles that are dramatically divergent from either of the nonrescue groups. In such a case, the rescue strains, or some subset of them, might form a third clade distinct from either *D. melanogaster* or *D. simulans*. Since each rescue strain is grouped in the same clade as the nonrescue strains with extremely high bootstrap support, it is very unlikely that hybrid rescue is caused by the emergence of a novel *cid* allele that is divergent from either *D. melanogaster* or *D. simulans*.

While phylogenetic analysis detects major patterns of divergence between sequence groups, it may be insensitive to finer genetic changes, perhaps those affecting only a few amino acids, which may alter the function of the Cid protein. In the case of *D. melanogaster*, it is very unlikely that *cid* alleles from rescue strains have any functional differences from nonrescue strains. The few differences that are unique to either *In(1)AB* or Melbourne are restricted to silent sites. Tai 255.1, on the other hand, differs from the remainder of the *D. melanogaster* strains by two adjacent amino-acid replacements. While we cannot rule out the possible functional consequences of these two amino-acid changes based on our present data, the fact that they are restricted to the Tai 255.1 strain indicates that they are not changes that generally cause the rescue of inviability by *D. melanogaster*. Furthermore, the observed changes in the amino-acid sequence of Cid in Tai 255.1 do not occur in the functional DNA binding domains of the Cid protein, as we might predict if the changes were responsible for the observed rescue phenotype.

a Neighbor-Joining



b Maximum Parsimony

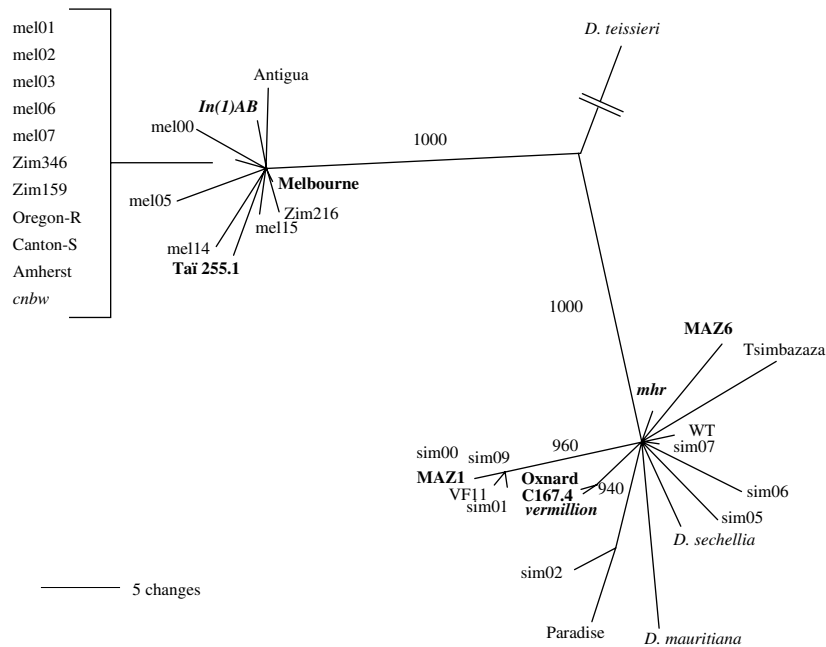


Figure 1 The neighbor-joining and the maximum parsimony phylograms of the *cid* sequences indicate that rescue strains (shown in bold) of *D. melanogaster* and *D. simulans* do not form separate clades from nonrescue strains. Bootstrap support (out of 1000) for species groupings, regardless of rescue phenotype, is very high. These data indicate that hybrid rescue is not caused by *cid* alleles introgressing across species boundaries, or by adaptation of *cid* to form a novel rescue clade. Furthermore, the data show that *cid* is much less variable in *D. melanogaster* than *D. simulans*, and that *cid* from members of the *D. simulans* species cluster (which includes *D. mauritiana* and *D. sechellia*) cannot be differentiated.

The case in *D. simulans* is somewhat more complex, because *cid* in this species is much more variable than in *D. melanogaster*. The 13 replacement polymorphisms that we observed in *D. simulans* have frequencies ranging from singleton sites to those that are present in 50% of the sample, indicating that much natural amino-acid

variation is currently segregating in the *D. simulans* population. Importantly, this variation is partitioned widely among both rescue and nonrescue strains. For instance, the sequences MAZ1 and VF11 form part of a tight clade within *D. simulans* and differ by only two amino acids. However, these two strains lie at opposite

ends of the rescue phenotype spectrum, with MAZ1 being a complete rescue line and VF11 refusing to be rescued by *D. melanogaster* *In(1)AB*, *f-M1* males. In contrast, C167.4, which has a very similar rescue phenotype to MAZ1, differs from this latter strain by six amino acids. Thus, it appears that no single mutation is adding to the rescue ability of the *D. simulans* strains. Furthermore, it is unlikely that any of the replacement changes are themselves responsible for rescue in their respective lines. If so, independent mutations within *Cid* would be responsible for some, but not all, of the rescue genotypes.

In addition to demonstrating that *Cid* is not involved in hybrid rescue between *D. simulans* and *D. melanogaster*, our data also show that divergent *Cid* alleles from these species are compatible in hybrid offspring. While hybrids from the rescue strains do not have the same fitness levels as conspecific crosses, rescue crosses do produce progeny that are viable and female progeny that are fully fertile. Thus, the coming together of divergent *Cid* alleles from the parental strains does not preclude hybrid offspring from having relatively high fitness levels, despite the contrary predictions of the centromere-driven speciation model (Henikoff *et al*, 2001a). This model asserts that crosses between populations with divergent centromeric machinery will experience centromeric imbalances sufficient to cause both inviability and sterility. Furthermore, the model predicts that these problems will most likely be manifested in the heterogametic sex, since it contains the most dissimilar pair of centromeres (those on the sex chromosomes), thus providing a potential explanation for Haldane's rule. However, our data from the rescue strains indicate that fully viable male and female offspring can result, although they have potentially incompatible components. Crosses of female *D. melanogaster* to male *D. simulans* normally produce inviable male progeny. However, male *D. melanogaster* from line *In(1)AB*, *f-M1* crossed to *D. simulans* produce fully viable males, despite the divergent *Cid* sequences of the parents. Likewise, viability in the reciprocal cross, which normally produces inviable females, is rescued by a variety of rescue strain crosses, again despite potentially incompatible centromeric machinery. Since males become inviable at the larval stage and inviability of the females is embryonic (Hadorn, 1961), it is possible that inviability is caused by a separate genetic mechanism in each direction. However, our data indicate that centromeric incompatibilities do not adequately explain hybrid inviability in either of these *Drosophila* crosses.

In the case of hybrid sterility, centromeric incompatibilities do not provide a suitable explanation for the restoration of fertility in females, but the case is still ambiguous in males. Since none of the rescue lines successfully restore male fertility, it is possible that the divergent centromeres play a role. This prospect is intriguing, especially because Henikoff *et al* (2001a) predict that male fertility in *Drosophila* should be more affected by centromeric incompatibilities than females. However, the data of Malik and Henikoff (2001) indicate that *Cid* from the closely related species *D. simulans*, *D. mauritiana*, and *D. sechellia* is nearly identical, even in putative DNA binding regions, in spite of the fact that these species produce completely sterile hybrid males (David *et al*, 1974; Lachaise *et al*, 1986). Studies focusing

on the genetic basis of reproductive isolation among these species suggest that many epistatic genes are involved in both female and male sterility (Cabot *et al*, 1994; Davis *et al*, 1994; Palopoli and Wu, 1994; Hollocher and Wu, 1996). The pattern of multiple loci appearing to be involved in male sterility appears to even be the case among taxa where speciation is not yet complete. In the case of the incipient Bogota and USA subspecies of *D. psuedoobscura*, at least four loci that interact epistatically have been determined to cause hybrid male sterility (Orr and Irving, 2001). Thus, while we cannot rule out centromeric incompatibility as a cause of male sterility, if it exists, it appears to be mediated through numerous additional genetic mechanisms.

The genetic components of reproductive isolation are best characterized between closely related, recently diverged, sister taxa. It is in these systems that the actual genetic mechanisms of reproductive isolation will be most easily separable from divergence that is merely consequential to isolation. In the case of *D. simulans* and *D. melanogaster*, many genes have been subject to adaptive divergence in one or both lineages (eg Eanes *et al*, 1993; Begun and Whitley, 2000; Begun *et al*, 2000; Schmidt *et al*, 2000; Verrelli and Eanes, 2000), only a handful of which (if any) are likely to have been a cause of isolation, rather than a product of postspeciation adaptation. In the case of *Cid*, adaptive evolution has clearly caused much of the divergence between *D. simulans* and *D. melanogaster*, affecting both the rescue and nonrescue lines (Malik and Henikoff, 2001). Nonetheless, because our data show there to be no apparent incompatibilities between *Cid* from *D. melanogaster* and *D. simulans*, the adaptive evolution of *Cid* appears to be a product of divergence and not a cause of speciation.

Centromeric histone and centromere incompatibilities were purported to be a general mechanism of speciation; however, *Cid* incompatibilities are not sufficient to cause reproductive isolation between *D. melanogaster* and *D. simulans*. Thus, the role of *cid* as a candidate 'speciation gene' appears to be rather limited. While we cannot rule out the centromere-drive model still acting in this species pair through other genes that may play a role in alleviating centromeric imbalances (Henikoff *et al*, 2001b), no gene except for *cid* has been shown to have such a central role in centromere function and to be adaptively divergent among species. The divergence of *cid* in these taxa demonstrates that *cid* coevolves with changing centromere sequences. However, we see no indication that hybrids suffer any loss of fitness as a result of incompatibilities between the divergent components of their centromere machinery.

Acknowledgements

We thank John Roote for supplying the lines Tai 255.1, Melbourne, and *Mhr*. We also thank the Evolutionary Discussion Group at the University of Notre Dame and Manuel Sainz for stimulating thought on this subject. Special thanks to Nora Besansky and members of her laboratory for use of their sequencing facility and Kristin Hennessy for laboratory assistance. This work was funded by a National Science Foundation award to Hope Hollocher, the Clare Boothe Luce Program of the Henry Luce Foundation, and the College of Science at the University of Notre Dame.

References

- Begun DJ, Whitley P (2000). Adaptive evolution of *relish*, a *Drosophila* NF-kappaB/IkappaB protein. *Genetics* **154**: 1231–1238.
- Begun DJ, Whitley P, Todd BL, Waldrip-Dail HM, Clark AG (2000). Molecular population genetics of male accessory gland proteins in *Drosophila*. *Genetics* **156**: 1879–1888.
- Cabot EL, Davis AW, Johnson NA, Wu CI (1994). Genetics of reproductive isolation in the *Drosophila simulans* clade: complex epistasis underlying hybrid male sterility. *Genetics* **137**: 175–189.
- Carracedo MC, Asenjo A, Casares P (2000). Location of *Shfr*, a new gene that rescues hybrid female viability in crosses between *Drosophila simulans* females and *D. melanogaster* males. *Heredity* **84**: 630–638.
- David J, Lemeunier F, Tsacas L, Bocquet C (1974). Hybridation d'une nouvelle espece, *Drosophila mauritiana*, avec *D. melanogaster* et *D. simulans*. *Ann Genet* **17**: 235–241.
- Davis AW, Noonburg EG, Wu CI (1994). Evidence for complex genetic interactions between conspecific chromosomes underlying hybrid female sterility in the *Drosophila simulans* clade. *Genetics* **137**: 191–199.
- Davis AW, Roote J, Morley T, Sawamura K, Herrmann S, Ashburner M (1996). Rescue of hybrid sterility in crosses between *D. melanogaster* and *D. simulans*. *Nature* **380**: 157–159.
- Dubill AJ (1996). The Evolution and Biogeography of Sexual Isolation in Southern African *Drosophila*. BA Thesis, Princeton University.
- Eanes WF, Kirchner M, Yoon J (1993). Evidence for adaptive evolution of the *G6pd* gene in the *Drosophila melanogaster* and *Drosophila simulans* lineages. *Proc Natl Acad Sci USA* **90**: 7475–7479.
- Felsenstein J (2001). *Phylip: phylogeny inference package*. Release 3.6 (alpha2). Seattle.
- Gloor GB, Preston CR, Johnson-Schlitz DM, Nassif NA, Phillis RW, Benz WK et al (1993). Type I repressors of P element mobility. *Genetics* **135**: 81–95.
- Haaf T, Willard HF (1997). Chromosome-specific alpha-satellite DNA from the centromere of chimpanzee chromosome 4. *Chromosoma* **106**: 226–232.
- Hadorn E (1961). Zur autonomie und phasenzpezifitat der letalitat von bastarden zwischen *Drosophila melanogaster* und *Drosophila simulans*. *Rev Suisse Zool* **68**: 197–207.
- Hall T (1999). BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucl Acids Symp Ser* **41**: 95–98.
- Henikoff S, Ahmad K, Malik HS (2001a). The centromere paradox: stable inheritance with rapidly evolving DNA. *Science* **293**: 1098–1102.
- Henikoff S, Ahmad K, Malik HS (2001b). Speciation and centromere evolution – response. *Science* **294**: 2479–2480.
- Hollocher H (1998). Reproductive isolation in *Drosophila*: how close are we to untangling the genetics of speciation? *Curr Opin Genet Dev* **8**: 709–714.
- Hollocher H, Agopian K, Waterbury J, O'Neill RW, Davis AW (2000). Characterization of defects in adult germline development and oogenesis of sterile and rescued female hybrids in crosses between *Drosophila simulans* and *Drosophila melanogaster*. *J Exp Zool* **288**: 205–218.
- Hollocher H, Wu CI (1996). The genetics of reproductive isolation in the *Drosophila simulans* clade: X vs autosomal effects and male vs female effects. *Genetics* **143**: 1243–1255.
- Hutter P, Roote J, Ashburner M (1990). A genetic basis for the inviability of hybrids between sibling species of *Drosophila*. *Genetics* **124**: 909–920.
- Johnson NA, Wu CI (1992). An empirical test of the meiotic drive models of hybrid sterility: sex-ratio data from hybrids between *Drosophila simulans* and *Drosophila sechellia*. *Genetics* **130**: 507–511.
- Kitagawa K, Hieter P (2001). Evolutionary conservation between budding yeast and human kinetochores. *Nat Rev Mol Cell Biol* **2**: 678–687.
- Lachaise D, David J, Lemeunier F, Tsacas L, Ashburner M (1986). The reproductive relationships of *Drosophila sechellia* with *D. mauritiana*, *D. simulans*, and *D. melanogaster* from the Afrotropical region. *Evolution* **40**: 262–271.
- Malik HS, Henikoff S (2001). Adaptive evolution of *Cid*, a centromere-specific histone in *Drosophila*. *Genetics* **157**: 1293–1298.
- Malik HS, Vermaak D, Henikoff S (2002). Recurrent evolution of DNA-binding motifs in the *Drosophila* centromeric histone. *Proc Natl Acad Sci USA* **99**: 1449–1454.
- Murphy TD, Karpen GH (1998). Centromeres take flight: alpha satellite and the quest for the human centromere. *Cell* **93**: 317–320.
- Orr HA, Irving S (2001). Complex epistasis and the genetic basis of hybrid sterility in the *Drosophila pseudoobscura* Bogota–USA hybridization. *Genetics* **158**: 1089–1100.
- Palopoli MF, Wu CI (1994). Genetics of hybrid male sterility between *Drosophila* sibling species: a complex web of epistasis is revealed in interspecific studies. *Genetics* **138**: 329–341.
- Rozas J, Rozas R (1999). DnaSP version 3: an integrated program for molecular population genetics and molecular evolution analysis. *Bioinformatics* **15**: 174–175.
- Sawamura K, Taira T, Watanabe TK (1993). Hybrid lethal systems in the *Drosophila melanogaster* species complex I. The maternal hybrid rescue (*mhr*) gene of *Drosophila simulans*. *Genetics* **133**: 299–305.
- Schmidt PS, Duvernell DD, Eanes WF (2000). Adaptive evolution of a candidate gene for aging in *Drosophila*. *Proc Natl Acad Sci USA* **97**: 10861–10865.
- Thompson JD, Higgins DG, Gibson TJ (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* **22**: 4673–4680.
- Tyler-Smith C, Florida G (2000). Many paths to the top of the mountain: diverse evolutionary solutions to centromere structure. *Cell* **102**: 5–8.
- Verrelli BC, Eanes WF (2000). Extensive amino acid polymorphism at the *Pgm* locus is consistent with adaptive protein evolution in *Drosophila melanogaster*. *Genetics* **156**: 1737–1752.
- Watanabe TK (1979). A gene that rescues the lethal hybrids between *Drosophila melanogaster* and *D. simulans*. *Jpn J Genet* **54**: 325–331.