

The Nuclear Component of a Cytonuclear Hybrid Incompatibility in *Mimulus* Maps to a Cluster of Pentatricopeptide Repeat Genes

Camille M. Barr¹ and Lila Fishman

Division of Biological Sciences, University of Montana, Missoula, Montana 59812

Manuscript received August 4, 2009

Accepted for publication November 16, 2009

ABSTRACT

Characterizing the genetic and molecular basis of hybrid incompatibilities is a first step toward understanding their evolutionary origins. We fine mapped the nuclear restorer (*Rf*) of cytoplasm-dependent anther sterility in *Mimulus* hybrids by identifying and targeting regions of the *Mimulus guttatus* genome containing large numbers of candidate pentatricopeptide repeat genes (PPRs). The single Mendelian locus *Rf* was first isolated to a 1.3-cM region on linkage group 7 that spans the genome's largest cluster of PPRs, then split into two tightly linked loci (*Rf1* and *Rf2*) by <10 recombination events in a large ($N = 6153$) fine-mapping population. Progeny testing of fertile recombinants demonstrated that a dominant *M. guttatus* allele at each *Rf* locus was sufficient to restore fertility. Each *Rf* locus spans a physical region containing numerous PPRs with high homology to each other, suggesting recent tandem duplication or transposition. Furthermore, these PPRs have higher homology to restorers in distantly related taxa (petunia and rice) than to PPRs elsewhere in the *Mimulus* genome. These results suggest that the cytoplasmic male sterility (CMS)–PPR interaction is highly conserved across flowering plants. In addition, given our theoretical understanding of cytonuclear coevolution, the finding that hybrid CMS results from interactions between a chimeric mitochondrial transcript that is modified by *Rf* loci identified as PPRs is consistent with a history of selfish mitochondrial evolution and compensatory nuclear coevolution within *M. guttatus*.

INTRINSIC postzygotic reproductive isolation caused by hybrid incompatibilities contributes to the origin and maintenance of species boundaries in many taxa (COYNE and ORR 1989, 2004; RIESEBERG and WILLIS 2007). However, because the production of sterile or inviable offspring appears inherently maladaptive, the evolution of hybrid incompatibilities has historically been difficult to explain. The Dobzhansky-Muller (D-M) model, in which interactions among heterospecific alleles cause sterility or inviability only in hybrids (BATESON 1909; DOBZHANSKY 1937; MULLER 1942; BOWMAN *et al.* 1992), provides a theoretical solution to this long-standing evolutionary problem. The key strength of this model is its generality: alleles that interact to cause low hybrid fitness need not have negative effects in their native genetic background, and so may evolve by drift or natural selection. Abundant empirical evidence exists for D-M incompatibilities as a major source of hybrid breakdown in many different taxa (WITTBRODT *et al.* 1989; TING *et al.* 1998; FISHMAN and WILLIS 2001; BARBASH *et al.* 2003, 2004; PRESGRAVES *et al.* 2003; WU and TING 2004; NOOR and FEDER 2006). Recently, the identification of individual D-M loci and interacting

pairs has also begun to provide insight into the evolutionary forces that drive the spread of incompatible alleles within populations/species (ORR 2005; MASLY *et al.* 2006; PRESGRAVES and STEPHAN 2007; BIKARD *et al.* 2009; PHADNIS and ORR 2009; TANG and PRESGRAVES 2009). One striking finding is that antagonistic coevolution within species, rather than the independent fixation of alleles that fortuitously interact only in hybrids, may be a common source of D-M incompatibilities (WITTBRODT *et al.* 1989; FRANK 1991; HURST and POMIANKOWSKI 1991; TING *et al.* 1998; BARBASH *et al.* 2003; PRESGRAVES *et al.* 2003; BRIDEAU *et al.* 2006; MASLY *et al.* 2006; PHADNIS and ORR 2009; TANG and PRESGRAVES 2009). Study of the molecular basis of hybrid incompatibilities in diverse systems is necessary to assess whether this is a general pattern and is also an important step toward assessing the role such incompatibilities play in maintaining species barriers.

Incompatibilities between uniparentally inherited cytoplasmic genes and biparentally inherited nuclear genes (henceforth, cytonuclear incompatibilities) appear to be common (TIFFIN *et al.* 2001; TURELLI and MOYLE 2007; LOWRY *et al.* 2008) and may be particularly likely to involve a history of molecular coevolution within species. Such cytonuclear incompatibilities may arise through a variety of coevolutionary pathways. First, the high degree of interaction between the protein products of nuclear and mitochondrial genes may necessitate coordinated evolution in response to environmental

Supporting information is available online at <http://www.genetics.org/cgi/content/full/genetics.109.108175/DC1>.

¹Corresponding author: Division of Biological Sciences, 32 Campus Dr., University of Montana, Missoula, MT 59812.
Email: camille.barr@mso.umt.edu

conditions (RAND *et al.* 2004). For example, in both the marine copepod *Tigriopus californicus* (BURTON 1990; WILLETT and BURTON 2004; WILLETT and BERKOWITZ 2007; ELLISON and BURTON 2008; WILLETT 2008) and the yeasts *Saccharomyces cerevisiae* and *S. bayanus* (LEE *et al.* 2008), mismatch between nuclear and mitochondrial genes results in reduced performance of electron transport in hybrids. There is evidence that the incompatibility in the yeasts is driven by adaptive divergence of mitochondrial genes and nuclear coevolution (LEE *et al.* 2008). Environmental selection has also been invoked in cytonuclear hybrid incompatibilities between sunflower taxa (SAMBATTI *et al.* 2008). Second, high mitochondrial mutation rates in animals (WOLFE *et al.* 1987; PALMER and HERBON 1988), coupled with low effective population sizes (BIRKY 2001), may allow mildly deleterious cytotypes to rise in frequency and favor nuclear alleles with compensatory effects (RAND *et al.* 2004). This mechanism has been suggested to drive the cytonuclear incompatibilities in *Nasonia* wasps in which hybrids experience reduced performance of electron transport complexes that involve nuclear and mitochondrial components (ELLISON *et al.* 2008; NIEHUIS *et al.* 2008). Finally, asymmetry in the inheritance of nuclear and cytoplasmic elements may generate the conditions for selfish cytoplasmic evolution and driven nuclear coevolution (HURST *et al.* 1996; BIRKY 2001; BURT and TRIVERS 2006). Maternally inherited endosymbionts that spread by disrupting male function, such as *Wolbachia* in insects, frequently cause sterility or inviability in interspecific hybrids (BORDENSTEIN *et al.* 2001; BORDENSTEIN and WERREN 2007). Similarly, there is an inherent conflict of interest between mitochondrial and nuclear genes that makes their antagonistic coevolution a likely source of hybrid incompatibility, particularly hybrid male sterility (COSMIDES and TOOBY 1981; FRANK 1989; HURST *et al.* 1996).

In plants, cytonuclear incompatibilities are both theoretically well understood (LEWIS 1941; COSMIDES and TOOBY 1981; GOUYON and COUVET 1987; FRANK 1989) and, in crop taxa, molecularly well characterized (SCHNABLE and WISE 1998; HANSON and BENTOLILA 2004). From the theoretical perspective, it is clear that a mitochondrial variant causing male sterility (cytoplasmic male sterility or CMS) in an otherwise hermaphroditic plant should spread rapidly through a population (LEWIS 1941; FRANK 1989). However, this creates strong selection for nuclear restorer alleles (*Rf*), which are predicted to jointly fix with the CMS mitochondrial type under broad conditions (FRANK 1989; HURST *et al.* 1996). Thus, hybrids between pairs of hermaphroditic populations/species with different CMS-*Rf* genes may frequently reveal otherwise hidden male sterility.

As predicted, CMS is widespread throughout the plant kingdom, as indicated by the frequent appearance of asymmetric male sterility in hybrid crosses (LASER and LERSTEN 1972; KAUL 1988). Because of its value in

the commercial production of hybrid seed (HAVEY 2004), the genetics and molecular biology of CMS have been heavily studied in crop taxa, and both the CMS and *Rf* genes have been identified in a number of cases (SCHNABLE and WISE 1998; HANSON and BENTOLILA 2004). In all cases, the CMS loci are chimeric genes, generated by structural rearrangements of the mitochondrial genome, which code for novel protein products (HANSON and BENTOLILA 2004). These genes share little-to-no homology among taxa, suggesting multiple, novel origins of CMS (HANSON and BENTOLILA 2004).

In contrast, *Rf* genes appear highly similar; all but one of those identified belong to the pentatricopeptide repeat (PPR) family (CUI *et al.* 1996; BENTOLILA *et al.* 2002; BROWN *et al.* 2003; DESLOIRE *et al.* 2003; KAZAMA and TORIYAMA 2003; KOIZUKA *et al.* 2003; KOMORI *et al.* 2004; LURIN *et al.* 2004; KLEIN *et al.* 2005). This large gene family appears to function predominantly by modifying mitochondrial transcripts (MEIERHOFF *et al.* 2003; LURIN *et al.* 2004; O'TOOLE *et al.* 2008), in concordance with what we know about CMS and the mechanisms of fertility restoration (IWABUCHI *et al.* 1993; BROWN 1999; HANSON and BENTOLILA 2004). In *Arabidopsis*, the PPR gene family has a cluster of duplicated genes that are homologous to the restorer from crop species, suggesting repeated cycles of duplication indicative of selfish evolution (GEDDY and BROWN 2007). However, because the molecular basis of restorer loci is understood only from domesticated crop taxa, it has been difficult to assess whether they have a history of selfish evolution.

Here, we describe the fine mapping and molecular characterization of the genomic region containing the *Rf* locus involved in a cytonuclear incompatibility between two species of monkeyflower, *Mimulus guttatus* and *M. nasutus*. In this system, a cytonuclear incompatibility causes anther sterility in hybrids with the *M. guttatus* cytoplasm, and the *Rf* segregates as a single locus in F₂ progeny (FISHMAN and WILLIS 2006). Recently, the mitochondrial CMS locus has been preliminarily identified as a novel chimeric open reading frame (ORF) generated through a rearrangement of the gene coding for the oxidative phosphorylation protein *NAD6* (CASE and WILLIS 2008). This novel gene appears to be fixed within the hermaphroditic *M. guttatus* population in which it was found, but its distribution is geographically restricted (CASE and WILLIS 2008). Thus, this system provides an excellent opportunity to understand both molecular mechanism and evolutionary history of a CMS–restorer interaction causing interspecific hybrid incompatibility.

In the present study, we used a “candidate gene-family” approach to fine map the *M. guttatus* *Rf* locus and characterize the genomic region surrounding it. We used bioinformatics tools based on the *Arabidopsis*/crop literature to identify regions of the draft *M. guttatus* genome sequence rich in PPR genes and then used

previous and new genetic maps to identify one PPR-rich region that was coincident with the location of *Rf* on linkage group 7 (LG7) of the hybrid genome. We then genotyped a large fine-mapping population ($N > 6000$) to isolate *Rf* to a region of 1.3 cM and break it up into two tightly linked loci. Both loci reside in a genomic region containing 17 PPRs, which cluster phylogenetically with restorer loci from other taxa but appear to be relatively recently duplicated.

MATERIALS AND METHODS

Study system: The yellow monkey flowers of the *M. guttatus* species complex are a morphologically diverse but largely interfertile group of wildflowers with their center of diversity in western North America (VICKERY 1978). *M. guttatus*, the most common species in the complex, has large, insect-pollinated flowers and is predominantly outcrossing (WILLIS 1993). Routine selfing appears to have evolved several times in the group (WU *et al.* 2008). The most widespread of the selfing taxa, *M. nasutus*, produces reduced, often cleistogamous flowers. Where *M. guttatus* and *M. nasutus* co-occur, potential pre-mating barriers to hybridization include differences in microhabitat and flowering time (MARTIN and WILLIS 2007) as well as differences in floral morphology (RITLAND and RITLAND 1989), pollen production (FENSTER and CARR 1997), and pollen tube growth (DIAZ and MACNAIR 1999) associated with their different mating systems. Despite these barriers to mating, hybrids are observed at low frequency at many sympatric sites, and there is evidence of recent local introgression at nuclear loci (SWEIGART and WILLIS 2003).

Specific populations and lines: Previously, we identified a cytonuclear incompatibility in hybrids between an inbred line of *M. guttatus* from Iron Mountain, Oregon (IM62) and a *M. nasutus* line from Sherar's Falls, Oregon (SF5) (FISHMAN and WILLIS 2006). In F_2 hybrids with the *M. guttatus* cytoplasm (F_2G), anther sterility segregates at $\sim 25\%$. The reciprocal cross shows no anther sterility, although there is an independent nuclear–nuclear incompatibility causing pollen inviability (SWEIGART *et al.* 2006). This pattern is consistent with a single Mendelian restorer locus with a dominant *M. guttatus* *Rf* allele. In a first generation backcross to *M. nasutus* (CSB1 population), the *Rf* locus was mapped to a region at the end of linkage group 7 bounded by the marker MgSTS574a (FISHMAN and WILLIS 2006).

Genome scan for PPRs: To target our fine mapping of *Rf*, we identified regions of the genome with high PPR density by running a hidden Markov model using programs in the HMMER package (<http://hmmer.janelia.org/>) and the consensus PPR model developed by LURIN *et al.* (2004) for *Arabidopsis thaliana*. Because PPR genes are numerous and highly diverse in both sequence and numbers of motifs, simple BLAST algorithms are not sufficient for their complete identification in a genome. Programs in the HMMER package allow the development of a profile of the likelihood of each amino acid in each position of the PPR motif and the searching of databases using this profile. The PPR profile used was developed by running the program hmmbuild on an alignment of 2357 *A. thaliana* PPR motifs (LURIN *et al.* 2004). We then ran the program hmmsearch to search our Mimulus sequence database (originally on the 4X genome build to develop markers for fine mapping and then later on the 7X build for comparative purposes) for similarity to this PPR profile. At the time of this work, the genome consisted of a large number of internally contiguous, but unassembled,

tracts of sequence called “scaffolds.” To identify PPR genes (each of which consists of a number of PPR motifs), we used a protocol similar to that of LURIN *et al.* (2004), identifying and removing hits that were 200 nucleotides away from any other motif (orphan motifs).

Marker development and genotyping: To genetically map PPR-rich scaffolds that were not already anchored to existing *M. guttatus* complex maps with gene-based MgSTS markers (<http://www.mimulusevolution.org>), we developed new intron-spanning markers. For a target region, we identified putative intron-containing genes and likely intron positions by aligning the genomic sequence to a database of dicot cDNA using the program GeneSeqer (USUKA *et al.* 2000) (<http://www.plantgdb.org>). Primers were designed in exon sequence to span putative introns, tested for amplification in both parental species, and screened for informative differences in amplicon length between the parental lines.

Coarse mapping: We used a bulk segregant approach to screen all informative MgSTS markers (<http://www.mimulusevolution.org>), as well as newly designed markers from PPR-rich scaffolds, for association with fertility restoration. The screening panel consisted of six pools of three sterile CSB1 individuals, presumed to be *M. nasutus* homozygotes (*rf/rf*), each. Potentially linked markers were then mapped in the full CSB1 population ($N = 192$). Marker amplification and genotyping protocols followed those used previously (FISHMAN and WILLIS 2005, 2006).

Fine mapping: For fine mapping of *Rf*, we introgressed the *M. guttatus* *Rf* allele, along with the IM62 CMS mitochondrion, into a *M. nasutus* nuclear background, by backcrossing two fertile CSB1 hybrids to the SF parental line. This selection and backcross process was repeated for two more generations to form CSB4 lines carrying the IM62 cytoplasm, heterozygous at loci involved with male fertility restoration, and 93.75% homozygous for *M. nasutus* alleles elsewhere. Representatives of the two independent CSB4 lines were selfed to form a large segregating population (CSBG; $N = 8448$) for fine mapping. At maturity, we phenotyped all individuals, collected leaf and bud tissue, and extracted genomic DNA using a standard CTAB/chloroform extraction protocol (see FISHMAN and WILLIS 2005 for details).

The entire CSBG population ($N = 6153$) was genotyped at the two flanking MgSTS markers and three markers from the target scaffold, and then a subset of informative recombinants was genotyped at three additional markers designed by walking through the *Rf* region scaffolds. Genotypes were scored automatically using GeneMapper 3.7 (Applied Biosystems), with additional hand scoring where necessary. Primer sequences for all mapping markers are in supporting information, Table S1.

Phylogenetic analysis of PPRs: To examine the homology of the Mimulus *Rf* candidate loci with other *Rf* loci, we compared the phylogenetic relationships of the amino acid sequences of select PPRs from Mimulus, Arabidopsis, petunia, rice, radish, and sorghum. To compare the sequences of the whole PPR gene, which includes the N- and C-terminal sequences, we identified ORFs using the probabilistic gene-identification program Genscan with the Arabidopsis parameter matrix (<http://genes.mit.edu/GENSCAN.html>, BURGE and KARLIN 1997) and with the ORF prediction program found in the *M. guttatus* genome GBrowse (www.mimulusevolution.org); we confirmed that these ORFs were PPRs by using BLAST (ALTSCHUL *et al.* 1990). We included 16 of the 17 PPRs in the regions spanning *Rf1* and *Rf2* (one of the PPRs was too small to include in this alignment) as well as 2–4 PPRs from the 4 LG7-unlinked scaffolds with the top HMMER-predicted PPRs (scaffolds 4, 8, 11, and 16). Three of these were 13–16 motif PPRs, and we included one additional 13–16 mer from the *M. guttatus*

genome. We also included the *Rf* sequences from petunia (BENTOLILA *et al.* 2002), rice (KAZAMA and TORIYAMA 2003; KOMORI *et al.* 2004), radish (DESLOIRE *et al.* 2003), and sorghum (KLEIN *et al.* 2005), five Arabidopsis *Rf* homologs on chromosome 1 (GEDDY and BROWN 2007), and 5 non-*Rf* homologous PPR sequences from *A. thaliana* (www.arabidopsis.org). Because we are focusing on the relationship between *M. guttatus* *Rf* candidates and *Rf* genes from other taxa, which vary between 13 and 16 motifs, to generate a data set with a sufficient amount of information for accurate tree prediction, we restricted our phylogeny to PPRs that had 7–20 motifs. We aligned all the PPRs with ClustalW.

For tree building, two independent analyses of MCMC were run as implemented in Mr. Bayes assuming a WAG + G + F model of sequence evolution as selected by ProtTest (ABASCAL *et al.* 2005). Analyses were metropolis coupled, with three heated chains and one cold chain (from which trees were sampled). Analyses were run for 1×10^6 generations, with trees sampled every 100 generations. Convergence of the analyses on stationarity (*i.e.*, when the chain has become independent of its starting position and is sampling trees in a manner that reflects the posterior probability distribution) was evaluated by the average standard deviation of split frequencies of the two analyses, and the first 25% of trees sampled were discarded as burn-in.

RESULTS

Identification of PPR-rich genomic regions: We identified 5635 PPR motifs that, after omitting orphan motifs, clustered into 789 groups representing putative PPR genes. The actual number of PPR-coding genes is likely smaller (and potentially much smaller) (see DISCUSSION). The distribution of PPRs among genome scaffolds was highly skewed (Figure 1), with a few scaffolds containing large clusters of PPRs and most (regardless of scaffold size) containing none. We were able to genetically map all scaffolds with >10 PPRs by identifying previously mapped MgSTS markers (www.mimulusevolution.org) on each scaffold or by designing and mapping new gene-based markers (Table S1). The distribution of PPRs with restorer-like motif numbers (13–16 motifs) was even more skewed, with a third (29/95) occurring on just three scaffolds (Table 1). Two of those scaffolds (s14 and s97) mapped to LG7 near markers associated with the *M. guttatus* *Rf* locus (FISHMAN and WILLIS 2006) and became the focus of our fine-mapping efforts.

Coarse mapping: Using a bulk-segregant approach to screen all previously unmapped MgSTS markers, we found one (MgSTS331) associated with *Rf* in the CSB1 population. A single crossover in this population ($N=192$) placed MgSTS331 distal to *Rf* on LG7.

Fine mapping: The CSBG fine-mapping population segregated, as expected, $\sim 1/4$ (23%; 1428/6153) anther-sterile individuals. All individuals were genotyped at markers known to flank the *Rf* locus from coarse mapping in the CSB1 population (MgSTS331, this study; MgSTS574a, FISHMAN and WILLIS 2006), as well as three newly designed markers from the two LG7 scaffolds rich in restorer-like PPRs (s97: 191_27; s14: 191_45 and 1_36).

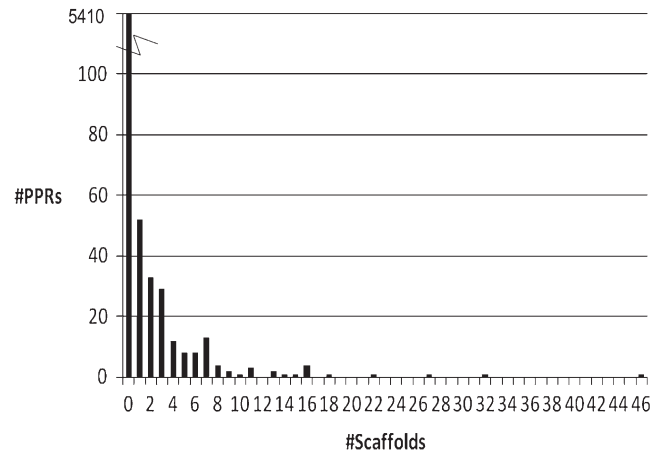


FIGURE 1.—Distribution of PPR genes on scaffolds of the *Mimulus guttatus* genome, with a very large majority of scaffolds containing no PPRs and a small number containing many. The actual number of PPR-coding genes in the *M. guttatus* genome is likely smaller than reported here due to spurious motif identification by HMMER and artificial splitting of single genes; however, we closely examined most of the PPRs on scaffold 14 (the rightmost bar) and it is clear that this scaffold is extremely PPR rich. While variation in number of PPRs among scaffolds is partially explained by scaffold size ($r^2 = 0.42$, $P < 0.000$), the majority of the variation observed in PPR number is not explained by scaffold size, and scaffold 14 falls as a clear outlier.

All sterile individuals were SF *M. nasutus* (*rf/rf*) homozygotes at both 191_45 and 191_27, but not at the flanking markers MgSTS331 and 1_36, narrowing the *Rf* region to 2.15 cM (265 crossovers/12,340 meioses) between the latter markers (Figure 2). These recombinants were genotyped at three additional markers on the two target scaffolds (s14: 14_172, 271_89; s97: 97_329) to orient the scaffolds and more finely map *Rf* within each scaffold.

TABLE 1

The top 10 genome scaffolds with *Rf*-like PPRs (13–16 motifs) in rank order

Scaffold	LG	13–16 motif PPRs				
		(sum)	13	14	15	16
14	7	20	2	18	0	0
8	6	6	2	4	0	0
97	7	3	1	2	0	0
24	8	2	2	0	0	0
11	8	2	1	1	0	0
16	6	2	1	0	0	1
4	13	2	1	1	0	0
37	12	2	1	0	0	1
87	10	2	1	0	0	1
29	7	2	1	1	0	0

Other than scaffold 14 (see MATERIALS AND METHODS), all scaffolds were anchored to linkage groups using publicly available linkage maps of MgSTS markers (www.mimulusevolution.org) and the draft *M. guttatus* genome browser [Mimulus Genome Project, U.S. Department of Energy (DOE) Joint Genome Institute].

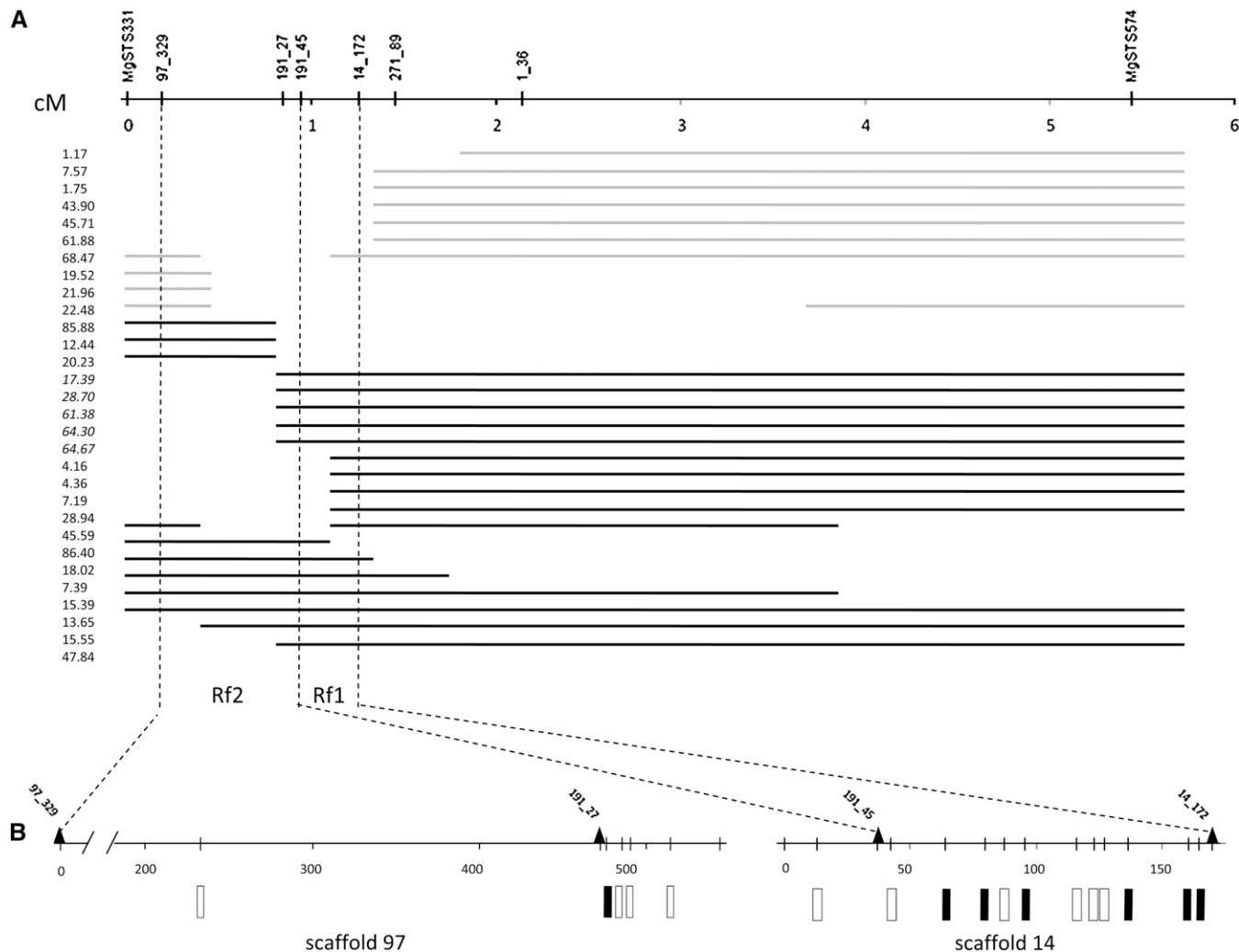


FIGURE 2.—Genetic and physical maps of the *Rf* loci and candidate PPR genes in *Mimulus*. (A) Genetic map showing informative recombinants. Horizontal bars represent regions of heterozygosity for CSBG individuals. Shaded bars indicate male-fertile plants and solid bars indicate male-sterile plants. Restoration of fertility maps to a region of ~ 1.3 cM. We refer to these loci as *Rf1* and *Rf2*. (B) Physical map of the region containing *Rf1* and *Rf2* showing the location of markers (triangles) and PPR genes (bars). Solid bars indicate those PPRs that are mitochondrially targeted as identified in the program PREDOTAR (SMALL *et al.* 2004), and open bars indicate those that are not mitochondrially targeted.

As shown in Figure 2, recombination within this region suggests that there are two very tightly linked *Rf* loci, each with a dominant *M. guttatus* allele capable of restoring fertility. First, we found that all individuals heterozygous at either 191_45 or 191_27 were fertile. This alone could be due to a single locus in the >1 -cM interval between these markers. However, there were also some individuals that were *M. nasutus* homozygotes at both 191_45 and 191_27, but were nonetheless fertile. To test whether this was due to the action of multiple restorer loci or to incomplete penetrance, we examined segregation ratios in the selfed progeny of representative lines. Lines homozygous *M. nasutus* at 191_45, but with heterozygous introgressions only to the left or right both segregated $\sim 1/4$ steriles (total $N = 8$, Table 2). These data are consistent with the existence of two *Rf* loci, one (*Rf1*) in the 0.32-cM interval bounded by

191_45 and 14_172 and the other (*Rf2*) in the 0.75-cM interval between 97_329 and 191_45. Hybrid anther sterility occurs only when both loci are homozygous for *M. nasutus* alleles.

In our original screen of the CSBG population, we also found 68 individuals that were *M. nasutus* homozygotes across the region containing *Rf1* and *Rf2* but were fertile. This represents $<5\%$ of the expected sterile genotypes. Barring double recombinants in the interval between 191_45 and 1_36, possible explanations for this are a *M. guttatus* introgression at a third locus sufficient for fertility restoration or incomplete penetrance of the CMS. Because segregation ratios in the CSBG and previous mapping populations did not differ significantly from single-locus Mendelian expectations, any third locus would have to be tightly linked to *Rf1* and *Rf2* and thus would segregate in the selfed progeny of these

TABLE 2
The frequencies of sterile individuals segregating in
progeny testing of selfed CSBG individuals
heterozygous at either *Rf1* or *Rf2*

Family	% steriles	<i>N</i>	X^2 <i>P</i> -value
20.23	0.30	27	0.91
85.88	0.11	55	0.02 ^a
12.44	0.25	61	0.94
4.16	0.21	184	0.17
4.36	0.18	108	0.08
7.19	0.24	84	0.80
28.94	0.26	188	0.74
45.59	0.29	35	0.63
28.70	0.22	76	0.60
61.38	0.18	38	0.35
17.39	0.26	38	0.85
64.67	0.29	45	0.55
64.30	0.15	55	0.07

Frequencies do not significantly differ from a model of a single dominant Mendelian locus, except with family 85.88, suggesting that both *Rf1* and *Rf2* are each sufficient to restore fertility.

^a Family 85.88 had ambiguous anther morphologies indicating the presence of other loci that modify male fertility; we scored individuals as sterile only if they had very clear morphology, which is likely the cause of the significant deviation from 25% sterile.

individuals. However, we saw no steriles in the progeny of three tested fertile individuals that were *M. nasutus* homozygotes across the entire region between markers MgSTS331 and 1_36. This result is intriguing and worthy of further investigation.

Molecular content of *Rf1* and *Rf2* regions: In the final draft of the *M. guttatus* (IM62) genome assembly, the region containing *Rf1* and *Rf2* spans two genome scaffolds. *Rf1* is bounded by two markers on s14, whereas *Rf2* is bounded by markers on s14 and s97 (Figure 2). To assess the occurrence of PPRs and other potential candidate genes, we examined the gene content of both regions using the gene prediction program GENSCAN (BURGE and KARLIN 1997). In the 133.4-kb region containing *Rf1*, we found 12 PPR proteins, 11 of which have 13–16 PPR motifs, and 8 of which had mitochondrially targeted transit peptides as identified with the program PREDOTAR (SMALL *et al.* 2004) and thus are likely candidates for *Rf1* (Figure 2).

Assessing candidates for *Rf2* was more complicated, as we cannot determine the gene content of the interval between scaffolds. However, in the previous 6X draft *M. guttatus* genome assembly, markers 191_27 and 191_45 were separated by only 108.7 kb on a single scaffold. Thus, we infer that the current gap between s14 and s97 most likely reflects difficulty in assembling this repetitive genomic region rather than a large intervening genomic region. Consistent with this interpretation, the genetic distance between 191_27 and 191_45 (0.18 cM)

is near the expectation for markers ~100 kb apart, given a local estimate of 320–765 kb/cM in the CSBG mapping population. Assembly issues were also indicated by the fact that markers on the far end of scaffold 97 (past 191_27, opposite 97_329) consistently mapped to another linkage group (data not shown). Therefore, for *Rf2*, we examined the gene content of the 544.7-kb region of s97 from 97_329 to 71.7 kb past 191_27 (where homology to the contiguous scaffold in the 6X assembly was lost) and the 37 kb of s14 distal to 191_45. In this region, we found six PPR proteins, three of which had PPR motif numbers between 13 and 16, and only one of which had a mitochondrially targeted transit peptide (Figure 2).

In addition to PPRs, we located 16 and 84 other ORFs in the *Rf1* and *Rf2* regions, respectively. Most of these genes were transposable elements, and none were obvious functional candidates for cytoplasm-dependent anther sterility.

We used a Bayesian tree-building algorithm to examine phylogenetic relationships among *Mimulus* PPR genes from the *Rf1/Rf2* region, *Mimulus* PPRs (some with 13–16 PPR domains) from elsewhere in the genome, as well as *Arabidopsis* PPRs from both categories and known *Rf* alleles from diverse crops. The 16 PPR candidate loci for *Rf1* and *Rf2* formed a clear clade, which is sister to the *Rf* loci of petunia and rice as well as the radish *Rf0* and *Arabidopsis* *Rf* homologs (Figure 3). Most other *Mimulus* PPRs group with non-*Rf* PPRs from other taxa, along with the *Rf* for sorghum. This pattern suggests a history of duplication in the *Mimulus* *Rf* region, but also indicates that *Rf1* and *Rf2* are likely to share significant amino acid sequence similarity with restorer loci in distantly related systems.

DISCUSSION

Characterizing the genetic and molecular basis of hybrid incompatibilities is a first step toward understanding their evolutionary origins. We fine mapped the *Rf* of cytoplasm-dependent anther sterility in *Mimulus* hybrids by targeting regions of the *M. guttatus* genome containing large numbers of candidate PPRs. The single Mendelian locus *Rf* was first isolated to a 1.3-cM region on linkage group 7 that spans the genome's largest cluster of PPRs, then split into two tightly linked loci (*Rf1* and *Rf2*) by <10 recombination events in our large ($N = 6153$) fine-mapping population. Progeny testing of fertile recombinants demonstrated that a dominant *M. guttatus* allele at each *Rf* locus was sufficient to restore hybrid CMS alone. Each *Rf* locus spans a physical region containing multiple PPRs with high homology to each other and to restorer loci in distantly related systems. These results suggest that crop plant CMS provides an excellent model for the molecular basis of cytonuclear hybrid male sterility in hermaphroditic plants and that the CMS–PPR interaction is extremely conserved across

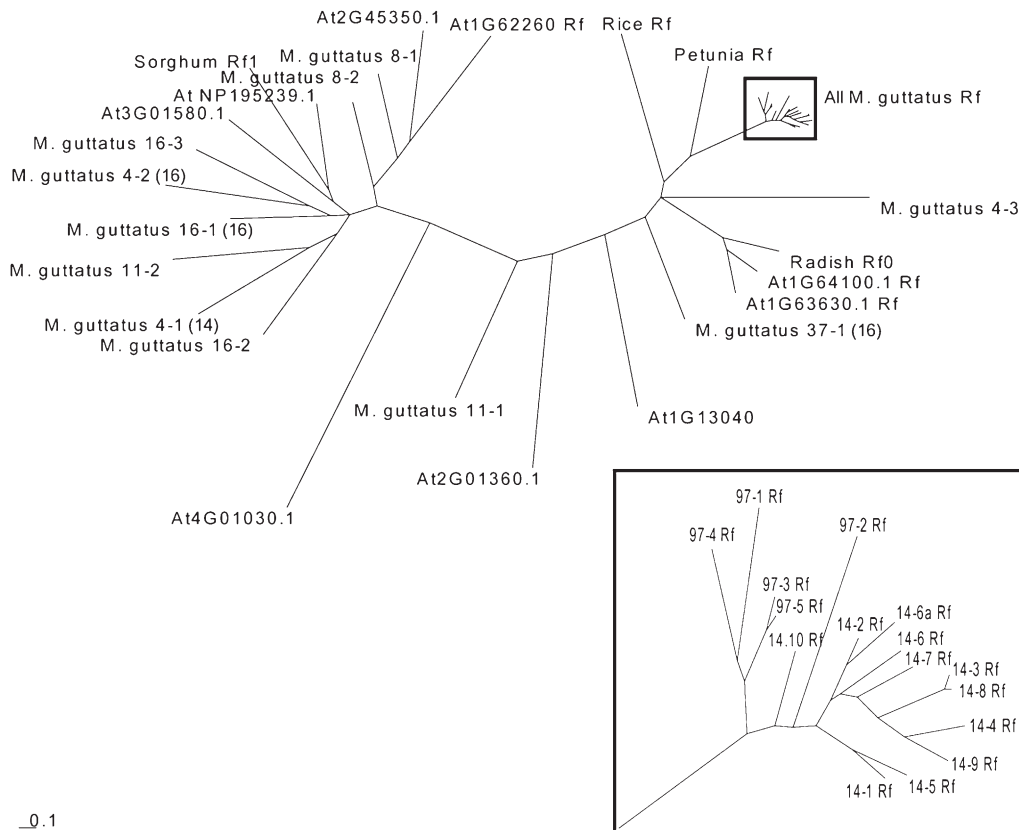


FIGURE 3.—Bayesian unrooted tree showing the relationships between *M. guttatus* Rf-linked PPRs, known *Rf* from crop taxa, Arabidopsis *Rf* homologs, and non-*Rf* PPRs from *M. guttatus* and other taxa. At, *Arabidopsis thaliana*. Numbers following *M. guttatus* PPRs refer to the scaffold followed by a PPR-specific number. Parentheses for *M. guttatus* PPRs indicate that the PPR has 13–16 motifs, and the number refers to the number of motifs. The inset expands the *M. guttatus* clade that contains all 16 (of the 17 total) PPRs used in this analysis within the region containing *Rf1/Rf2*.

flowering plants. In addition, given our theoretical understanding of the selfish basis of cytonuclear coevolution with the expectation of genetic arms race (FRANK 1989; BURT and TRIVERS 2006), the finding that hybrid CMS results from interactions between a novel chimeric mitochondrial transcript that is modified by *Rf* loci (CASE and WILLIS 2008) identified as PPRs (this study) is consistent with a history of selfish mitochondrial evolution and compensatory nuclear coevolution within *M. guttatus*. Further work will be necessary to precisely identify the *Rf* and *rf* alleles and investigate their evolutionary histories; however, this work provides preliminary support for the idea that intraspecific epistasis driven by selfish evolution, rather than chance interactions of independently fixed alleles, may often cause hybrid incompatibilities.

Evidence that PPRs restore CMS in Mimulus hybrids: We began fine mapping with a strong hypothesis that *Mimulus Rf* was a PPR gene, given known molecular mechanisms of restoration in other systems and the nature of the *Mimulus* mitochondrial CMS gene (CASE and WILLIS 2008). Four of the five *Rf* genes discovered to date have been PPRs and, in three taxa (petunia *Rf*, radish/Brassica *Rfo*, and rice *Rf1a* and *Rf1b*), the *Rf*/PPRs are homologous despite spanning the plant kingdom and the monocot/dicot split (HANSON and BENTOLILA 2004; LURIN *et al.* 2004; KLEIN *et al.* 2005). Very preliminary evidence suggests that the cotton *Rf2* restorer may also be a PPR, adding gameto-

phytic restoration to the above sporophytic systems (WANG *et al.* 2007). Only a few members of this recently discovered and very large family of genes have been functionally characterized (LURIN *et al.* 2004), but, in general, PPR genes appear to function in the modification of organellar mRNA transcripts (ANDRES *et al.* 2007). Restoration of CMS, which occurs by modification of chimeric mitochondrial transcripts in all known cases (BENTOLILA *et al.* 2002), including *Mimulus* (CASE and WILLIS 2008), falls within this broad functional role.

As in other genomes (GEDDY and BROWN 2007; O'TOOLE *et al.* 2008), restorer-like PPRs appear to cluster in *Mimulus* (Figure 1; Table 1). Most notably, the adjacent genome scaffolds (s14 and s97) containing *Rf1* and *Rf2* have very high densities of restorer-like PPRs (Table 1), with s14 containing the highest number of PPRs of any scaffold in the *Mimulus* genome (Figure 1, Table S2). Although it is remotely possible that some other gene(s) in this region encodes *Rf1* and *Rf2*, this coincidence strongly suggests that *Mimulus Rf* loci are also restorer-like PPRs. The *Mimulus Rf* candidates on s14 and s97 have the P-type PPR motif shared by all known *Rf* PPRs (LURIN *et al.* 2004; GEDDY and BROWN 2007), and most also have 13–16 motifs (Figure 2). While PPR motif number in Arabidopsis ranges from 2 to 26 (LURIN *et al.* 2004), the known *Rf* loci all range from 13 to 16 motifs: petunia (14 motifs), rice (16 motifs), sorghum (14 motifs), and radish/Brassica (16 motifs). The significance of motif number for fertility restora-

tion is unknown; it is likely that similar motif numbers reflect homology and not direct effects on restoration. Sequence polymorphisms, and not motif number, differentiate fertility restoring *Rf* and nonrestoring *rf* in sorghum and radish (KOIZUKA *et al.* 2003; KLEIN *et al.* 2005), and deletions on the promoter region differentiate *Rf* and *rf* in petunia and rice (BENTOLILA *et al.* 2002; KAZAMA and TORIYAMA 2003; KOMORI *et al.* 2004). If we restrict our candidate genes to mitochondrially targeted PPR genes with 13–16 repeat motifs in the relevant regions of s14 and s97, there is one remaining strong candidate gene for *Rf2* and six for *Rf1* (Figure 2). Given their sequence similarity (see below) and difficulties with the genome assembly in this chromosomal region, positionally cloning and functional confirmation of the individual *Rf1* and *Rf2* genes will likely be difficult. However, we are confident that these few candidate PPRs accurately represent the functional category of *Mimulus Rf* genes.

Commonalities with CMS restoration in other systems: In addition to the features used in their identification, the candidate PPRs for *Mimulus Rf* loci share several features with molecularly characterized *Rf* loci in other systems. First, all of the PPRs in these intervals (both mitochondrially and nonmitochondrially targeted), cluster together in a single distinct clade including the rice *Rf* locus (KAZAMA and TORIYAMA 2003; KOMORI *et al.* 2004), with petunia and sorghum *Rf* genes as the closest sisters (BENTOLILA *et al.* 2002; KLEIN *et al.* 2005). Because PPRs appear to evolve very rapidly and sequence alignments are complicated by variation in motif number (RIVALS *et al.* 2006), we do not want to overinterpret this finding. However, it is clear that our candidates share close amino acid sequence homology with restorers in several other systems.

Second, it appears that the PPRs in the *Rf* region expanded via tandem duplication and/or transposition, as they are more closely related to each other than to PPRs elsewhere in the *Mimulus* genome. This appears to be a common pattern for PPR genes at multiple evolutionary scales. PPR proteins are present in low copy numbers in other eukaryotes (five in yeast, six in humans, and two in *Drosophila*), but have undergone a massive expansion in land plants, with >450 in *Arabidopsis* and rice (LURIN *et al.* 2004; ANDRES *et al.* 2007). This primary expansion is thought to have occurred early in land plant evolution, through whole-genome duplications and retrotransposition (GEDDY and BROWN 2007; KATO *et al.* 2007; O'TOOLE *et al.* 2008). The *Rf*-like PPRs appear to have expanded more recently than the rest of the gene family and show particularly strong evidence for duplication, transposition, and diversifying selection (GEDDY and BROWN 2007; O'TOOLE *et al.* 2008). Unlike other PPRs, the *Rf* PPRs in rice, radish, and petunia, as well as the *Rf* homolog PPRs in *Arabidopsis*, exist in clusters suggesting relatively recent local duplication (HANSON and BENTOLILA 2004; LURIN *et al.*

2004; KATO *et al.* 2007). This pattern, which is also evident with the *Rf*-associated PPRs in *Mimulus*, is consistent with ongoing selection for new (or more) PPRs by selfish evolution of mitochondrial CMS (GEDDY and BROWN 2007).

Third, the presence of more than one *Rf* allele in *M. guttatus* parallels other taxa. In the CMS-BT strain of rice, fertility restoration is achieved by at least two tightly linked *Rf* alleles, *Rf1a* and *Rf1b*, either of which is sufficient to restore male fertility (WANG *et al.* 2006). A similar situation exists in the Ogura CMS of radish, in which *Rfo* and *Rfo2* are tightly linked and both sufficient to restore male fertility (WANG *et al.* 2008), along with at least several more loci with complicated effects (NIEUWHOF 1990; BETT and LYDIATE 2004). Recent work with the CMS-C system in maize suggests two tightly linked *Rf* loci with dominant effects, *Rf4* and *Rf5*, only one of which is sufficient to restore fertility over all CMS-C types (HU *et al.* 2006). In cotton, the sporophytic *Rf1* and gametophytic *Rf2* are also tightly linked (ZHANG and STEWART 2001; WANG *et al.* 2007). It is not yet clear what evolutionary processes contribute to this pattern, but it is intriguing, particularly given population genetic theory about the cospread of CMS and *Rf* alleles (see below).

Evolutionary implications: The genetic and molecular characterization of *Mimulus Rf* sets the stage for understanding the evolutionary dynamics of a major source of plant hybrid incompatibility. CMS genes are largely generated through mitochondrial genome rearrangements, which are thought to be common in plants (PALMER and HERBON 1988). This mutational input, along with strong selection favoring any mitochondrial variant that can achieve increased female fertility through decreases in male function (FRANK 1989; ATLAN *et al.* 1992; ASHMAN 1994), generates the conditions for repeated selfish mitochondrial evolution. This favors nuclear *Rf* alleles, which under some conditions can be polymorphic and maintain gynodioecy, but are more commonly expected to spread to fixation along with the CMS gene (CHARLESWORTH and GANDERS 1979; HURST *et al.* 1996; BURT and TRIVERS 2006). This would leave populations with no phenotypic evidence of past conflict, but with genotypic traces of rapid divergence in the cytoplasmic and nuclear loci involved, and potential hybrid incompatibilities with non-*Rf* populations (HURST *et al.* 1996). Although high frequencies of anther sterility have never been reported in numerous field studies of this genus (WU *et al.* 2008), it is likely that Iron Mountain *M. guttatus* (like hermaphroditic crop plants with hybrid CMS) has experienced the history of selfish mitochondrial evolution and nuclear coevolution predicted by these models. However, unlike crop plants with a selective and demographic history complicated by domestication, *M. guttatus* provides the opportunity to explicitly test for the molecular population genetic signature of the

coevolutionary process underlying this extremely common phenomenon (KAUL 1988). As in the few animal systems in which D-M incompatibility genes have been characterized (WITTBRODT *et al.* 1989; TING *et al.* 1998; BARBASH *et al.* 2003; PRESGRAVES *et al.* 2003; BRIDEAU *et al.* 2006; MASLY *et al.* 2006; PHADNIS and ORR 2009; TANG and PRESGRAVES 2009), it may be that selfish evolution and compensatory coevolution are major drivers of hybrid breakdown in plants.

Although we cannot yet say whether the *Mimulus* cytonuclear incompatibility evolved via selfish mitochondrial evolution and nuclear coevolution, the finding that the *Mimulus Rf* loci map to a cluster of PPR genes provides circumstantial evidence for such a scenario. PPRs function by modifying organellar transcripts, and precise transcript modification is necessary for removal of the chimeric ORF associated with NAD6 in Iron Mountain *M. guttatus* (CASE and WILLIS 2008). The finding that cytonuclear hybrid male sterility in plants commonly involves interactions between rearranged respiratory proteins and PPRs (HANSON and BENTOLILA 2004) argues for a common, and predictable, evolutionary process. The theory of selfish CMS evolution provides such a process (HURST *et al.* 1996; BURT and TRIVERS 2006; GEDDY and BROWN 2007) and, unlike other scenarios for the evolution of hybrid sterility, predicts that male sterility should have no negative effect on female fitness. Of course, in the absence of strong evidence for selection on *Mimulus* CMS and/or restorer loci, it remains possible that one or both restorers predates the CMS, such that male sterility was never expressed. However, the chance fixation of a novel neutral mitochondrial mutation in the very large (census size >10,000 annually), predominantly outcrossing Iron Mountain *M. guttatus* population is difficult to envision.

The particular genetics of nuclear restoration in *Mimulus* raises additional questions about the origins of cytonuclear hybrid incompatibilities and mating system variation in plants. Traditional discussions of CMS assume a matching-alleles model with the introduction and increase of a novel CMS-inducing mitochondrial gene driving the selection of a unique and matching nuclear restorer (FRANK 1989). Loss of the particular CMS is predicted to lead to loss of the matching restorer, likely through a negative fitness cost of the restorer (BAILEY 2002). The presence of two linked loci in *Mimulus*, each capable of effecting fertility restoration, as well as more complicated genetic details known from other systems, challenges the simplicity of this model. These details underscore the importance of incorporating the increasingly readily available molecular and genetic information about hybrid incompatibilities into future theoretical models.

We are very grateful to Arpiar Saunders for skilled assistance with the bioinformatics. We also thank Scott Miller for help with the phylogenetic trees. Dan Rokhsar and Jeremy Schmutz of Department of

Energy (DOE) Joint Genome Institute provided the *Mimulus guttatus* draft whole-genome sequence. We thank Ben Sundy, Samantha Campbell, Ben Ewen-Campen, Chelsea Luce, and John Tuthill for assistance with DNA extractions and greenhouse care, and John Willis and Andrea Case for helpful discussions. Funding was provided by National Science Foundation grants DEB-0316786 (L.F.) and BIO-0328326 (J. H. Willis, L. F., T. J. Vision, H. D. Bradshaw, Jr., D. W. Schemske and J. Tomkins.)

LITERATURE CITED

- ABASCAL, F., R. ZARDOYA and D. POSADA, 2005 ProtTest: selection of best-fit models of protein evolution. *Bioinformatics* **21**: 2104–2105.
- ALTSCHUL, S. F., W. GISH, W. MILLER, E. W. MYERS and D. J. LIPMAN, 1990 Basic local alignment search tool. *J. Mol. Biol.* **215**: 403–410.
- ANDRES, C., C. LURIN and I. D. SMALL, 2007 The multifarious roles of PPR proteins in plant mitochondrial gene expression. *Physiol. Plant.* **129**: 14–22.
- ASHMAN, T. L., 1994 Reproductive allocation in hermaphrodite and female plants of *Sidalcea oregana* ssp. *spicata* (Malvaceae) using 4 currencies. *Am. J. Bot.* **81**: 433–438.
- ATLAN, A., P. H. GOUYON, T. FOURNIAL, D. PONENTE and D. COUVET, 1992 Sex allocation in an hermaphroditic plant: the case of gynodioecy in *Thymus vulgaris* L. *J. Evol. Biol.* **5**: 189–203.
- BAILEY, M. F., 2002 A cost of restoration of male fertility in a gynodioecious species, *Lobelia siphilitica*. *Evolution* **56**: 2178–2186.
- BARBASH, D. A., D. F. SHINO, A. M. TARONE and J. ROOTE, 2003 A rapidly evolving MYB-related protein causes species isolation in *Drosophila*. *Proc. Natl. Acad. Sci. USA* **100**: 5302–5307.
- BARBASH, D. A., P. AWADALLA and A. M. TARONE, 2004 Functional divergence caused by ancient positive selection of a *Drosophila* hybrid incompatibility locus. *PLoS Biol.* **2**: 839–848.
- BATESON, W., 1909 Heredity and variation in modern lights, pp. 85–101 in *Darwin and Modern Science*, edited by A. C. STEWARD. Cambridge University Press, Cambridge, UK.
- BENTOLILA, S., A. A. ALFONSO and M. R. HANSON, 2002 A pentatricopeptide repeat-containing gene restores fertility to cytoplasmic male-sterile plants. *Proc. Natl. Acad. Sci. USA* **99**: 10887–10892.
- BETT, K. E., and D. J. LYDIATE, 2004 Mapping and genetic characterization of loci controlling the restoration of male fertility in Ogura CMS radish. *Mol. Breed.* **13**: 125–133.
- BIKARD, D., D. PATEL, C. LE METTE, V. GIORGI, C. CAMILLERI *et al.*, 2009 Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science* **323**: 623–626.
- BIRKY, C. W., 2001 The inheritance of genes in mitochondria and chloroplasts: laws, mechanisms, and models. *Annu. Rev. Genet.* **35**: 125–148.
- BORDENSTEIN, S. R., F. P. O'HARA and J. H. WERREN, 2001 Wolbachia-induced incompatibility precedes other hybrid incompatibilities in *Nasonia*. *Nature* **409**: 707–710.
- BORDENSTEIN, S. R., and J. H. WERREN, 2007 Bidirectional incompatibility among divergent Wolbachia and incompatibility level differences among closely related Wolbachia in *Nasonia*. *Heredity* **99**: 278–287.
- BOWMAN, J. L., H. SAKAI, T. JACK, D. WEIGEL, U. MAYER *et al.*, 1992 SUPERMAN, a regulator of floral homeotic genes in *Arabidopsis*. *Development* **114**: 599–615.
- BRIDEAU, N. J., H. A. FLORES, J. WANG, S. MAHESHWARI, X. WANG *et al.*, 2006 Two Dobzhansky-Muller genes interact to cause hybrid lethality in *Drosophila*. *Science* **314**: 1292–1295.
- BROWN, G. G., 1999 Unique aspects of cytoplasmic male sterility and fertility restoration in *Brassica napus*. *J. Hered.* **90**: 351–356.
- BROWN, G. G., N. FORMANOVA, H. JIN, R. WARGACHUK, C. DENDY *et al.*, 2003 The radish Rfo restorer gene of *Ogura* cytoplasmic male sterility encodes a protein with multiple pentatricopeptide repeats. *Plant J.* **35**: 262–272.
- BURGE, C., and S. KARLIN, 1997 Prediction of complete gene structures in human genomic DNA. *J. Mol. Biol.* **268**: 78–94.
- BURT, A., and R. TRIVERS, 2006 *Genes in Conflict: The Biology of Selfish Genetic Elements*. Harvard University Press, Cambridge, MA.
- BURTON, R. S., 1990 Hybrid breakdown in developmental time in the copepod *Tigriopus californicus*. *Evolution* **44**: 1814–1822.

- CASE, A. L., and J. H. WILLIS, 2008 Hybrid male sterility in *Mimulus* (Phrymaceae) is associated with a geographically restricted mitochondrial rearrangement. *Evolution* **62**: 1026–1039.
- CHARLESWORTH, D., and F. R. GANDERS, 1979 The population genetics of gynodioecy with cytoplasmic-genic male-sterility. *Heredity* **43**: 213–218.
- COSMIDES, L. M., and J. TOOBY, 1981 Cytoplasmic inheritance and intragenomic conflict. *J. Theor. Biol.* **89**: 83–129.
- COYNE, J. A., and H. A. ORR, 1989 Patterns of speciation in *Drosophila*. *Evolution* **43**: 362–381.
- COYNE, J. A., and H. A. ORR, 2004 *Speciation*. Sinauer, Sunderland, MA.
- CUI, X. Q., R. P. WISE and P. S. SCHNABLE, 1996 The rf2 nuclear restorer gene of male-sterile T-cytoplasm maize. *Science* **272**: 1334–1336.
- DESLOIRE, S., H. GHERBI, W. LALOU, S. MARHADOUR, V. CLOUET *et al.*, 2003 Identification of the fertility restoration locus, Rfo, in radish, as a member of the pentatricopeptide-repeat protein family. *EMBO Rep.* **4**: 588–594.
- DIAZ, A., and M. R. MACNAIR, 1999 Pollen tube competition as a mechanism of prezygotic reproductive isolation between *Mimulus nasutus* and its presumed progenitor *M. guttatus*. *New Phytol.* **144**: 471–478.
- DOBZHANSKY, T., 1937 *Genetics and the Origin of Species*. Columbia University Press, New York, NY.
- ELLISON, C. K., and R. S. BURTON, 2008 Interpopulation hybrid breakdown maps to the mitochondrial genome. *Evolution* **62**: 631–638.
- ELLISON, C. K., O. NIEHUIS and J. GADAU, 2008 Hybrid breakdown and mitochondrial dysfunction in hybrids of *Nasonia* parasitoid wasps. *J. Evol. Biol.* **21**: 1844–1851.
- FENSTER, C. B., and D. E. CARR, 1997 Genetics of sex allocation in *Mimulus* (Scrophulariaceae). *J. Evol. Biol.* **10**: 641–661.
- FISHMAN, L., and J. H. WILLIS, 2001 Evidence for Dobzhansky-Muller incompatibilities contributing to the sterility of hybrids between *Mimulus guttatus* and *M. nasutus*. *Evolution* **55**: 1932–1942.
- FISHMAN, L., and J. H. WILLIS, 2005 A novel meiotic drive locus almost completely distorts segregation in *Mimulus* (monkey-flower) hybrids. *Genetics* **169**: 347–353.
- FISHMAN, L., and J. H. WILLIS, 2006 A cytonuclear incompatibility causes anther sterility in *Mimulus* hybrids. *Evolution* **60**: 1372–1381.
- FRANK, S. A., 1989 The evolutionary dynamics of cytoplasmic male sterility. *Am. Nat.* **133**: 345–376.
- FRANK, S. A., 1991 Divergence of meiotic drive-suppression systems as an explanation for sex-biased hybrid sterility and inviability. *Evolution* **45**: 262–267.
- GEDDY, R., and G. G. BROWN, 2007 Genes encoding pentatricopeptide repeat (PPR) proteins are not conserved in location in plant genomes and may be subject to diversifying selection. *BMC Genomics* **8**: 130–142.
- GOUYON, P. H., and D. COUVET, 1987 A conflict between two sexes, females and hermaphrodites. *Experientia Suppl.* **55**: 245–261.
- HANSON, M. R., and S. BENTOLILA, 2004 Interactions of mitochondrial and nuclear genes that affect male gametophyte development. *Plant Cell* **16**: S154–S169.
- HAVEY, M. J., 2004 The use of cytoplasmic male sterility for hybrid seed production, pp. 623–634 in *Molecular Biology and Biotechnology of Plant Organelles*. Springer, Berlin.
- HU, Y. M., J. H. TANG, H. YANG, H. L. XIE, X. M. LU *et al.*, 2006 Identification and mapping of Rf1 an inhibitor of the Rf5 restorer gene for Cms-C in maize (*Zea mays* L.). *Theor. Appl. Genet.* **113**: 357–360.
- HURST, L. D., and A. POMIANKOWSKI, 1991 Causes of sex-ratio bias may account for unisexual sterility in hybrids: a new explanation of Haldane's rule and related phenomena. *Genetics* **128**: 841–858.
- HURST, L. D., A. ATLAN and B. O. BENGTSOON, 1996 Genetic conflicts. *Quart. Rev. Biol.* **71**: 317–364.
- IWABUCHI, M., J. KYOZUKA and K. SHIMAMOTO, 1993 Processing followed by complete editing of an altered mitochondrial atp6 RNA restores fertility of cytoplasmic male sterile rice. *EMBO J.* **12**: 1437–1446.
- KATO, H., K. TEZUKA, Y. Y. FENG, T. KAWAMOTO, H. TAKAHASHI *et al.*, 2007 Structural diversity and evolution of the Rf1 locus in the genus *Oryza*. *Heredity* **99**: 516–524.
- KAUL, M. L. H., 1988 *Male Sterility in Higher Plants*. Springer, Berlin.
- KAZAMA, T., and K. TORIYAMA, 2003 A pentatricopeptide repeat-containing gene that promotes the processing of aberrant atp6 RNA of cytoplasmic male-sterile rice. *FEBS Lett.* **544**: 99–102.
- KLEIN, R. R., P. E. KLEIN, J. E. MULLET, P. MINX, W. L. ROONEY *et al.*, 2005 Fertility restorer locus *Rf1* of sorghum (*Sorghum bicolor* L.) encodes a pentatricopeptide repeat protein not present in the colinear region of rice chromosome 12. *Theor. Appl. Genet.* **111**: 994–1012.
- KOIZUKA, N., R. IMAI, H. FUJIMOTO, T. HAYAKAWA, Y. KIMURA *et al.*, 2003 Genetic characterization of a pentatricopeptide repeat protein gene, orf687, that restores fertility in the cytoplasmic male-sterile Kosenia radish. *Plant J.* **34**: 407–415.
- KOMORI, T., S. OHTA, N. MURAI, Y. TAKAKURA, Y. KURAYA *et al.*, 2004 Map-based cloning of a fertility restorer gene, Rf-1, in rice (*Oryza sativa* L.). *Plant J.* **37**: 315–325.
- LASER, K. D., and N. R. LERSTEN, 1972 Anatomy and cytology of microsporogenesis in cytoplasmic male sterile angiosperms. *Bot. Rev.* **38**: 425.
- LEE, H.-Y., J.-Y. CHOU, L. CHEONG, N.-H. CHANG, S.-Y. YANG *et al.*, 2008 Incompatibility of nuclear and mitochondrial genomes causes hybrid sterility between two yeast species. *Cell* **135**: 1065–1073.
- LEWIS, D., 1941 Male sterility in natural populations of hermaphrodite plants. *New Phytol.* **40**: 56–63.
- LOWRY, D. B., J. L. MODLISZEWSKI, K. M. WRIGHT, C. A. WU and J. H. WILLIS, 2008 The strength and genetic basis of reproductive isolating barriers in flowering plants. *Philos. Trans. R. Soc. B Biol. Sci.* **363**: 3009–3021.
- LURIN, C., C. ANDRES, S. AUBOURG, M. BELLAOUI, F. BITTON *et al.*, 2004 Genome-wide analysis of Arabidopsis pentatricopeptide repeat proteins reveals their essential role in organelle biogenesis. *Plant Cell* **16**: 2089–2103.
- MARTIN, N. H., and J. H. WILLIS, 2007 Ecological divergence associated with mating system causes nearly complete reproductive isolation between sympatric *Mimulus* species. *Evolution* **61**: 68–82.
- MASLY, J. P., C. D. JONES, M. A. F. NOOR, J. LOCKE and H. A. ORR, 2006 Gene transposition as a cause of hybrid sterility in *Drosophila*. *Science* **313**: 1448–1450.
- MEIERHOFER, K., S. FELDER, T. NAKAMURA, N. BECHTOLD and G. SCHUSTER, 2003 HCF152, an Arabidopsis RNA binding pentatricopeptide repeat protein involved in the processing of chloroplast psbB-psbT-psbH-petB-petD RNAs. *Plant Cell* **15**: 1480–1495.
- MULLER, H. J., 1942 Isolating mechanisms, evolution and temperature. *Biol. Symp.* **6**: 71–125.
- NIEHUIS, O., A. K. JUDSON and J. GADAU, 2008 Cytonuclear genetic incompatibilities cause increased mortality in male F2 hybrids of *Nasonia giraulti* and *N. vitripennis*. *Genetics* **178**: 413–426.
- NIEUWHOF, M., 1990 Cytoplasmic-genetic male-sterility in radish (*Raphanus sativus* L.). Identification of maintainers, inheritance of male-sterility and effect of environmental factors. *Euphytica* **47**: 171–177.
- NOOR, M. A. F., and J. L. FEDER, 2006 Speciation genetics: evolving approaches. *Nat. Rev. Genet.* **7**: 851–861.
- O'TOOLE, N., M. HATTORI, C. ANDRES, K. IIDA, C. LURIN *et al.*, 2008 On the expansion of the pentatricopeptide repeat gene family in plants. *Mol. Biol. Evol.* **25**: 1120–1128.
- ORR, H. A., 2005 The genetic basis of reproductive isolation: insights from *Drosophila*. *Proc. Natl. Acad. Sci. USA* **102**: 6522–6526.
- PALMER, J. D., and L. A. HERBON, 1988 Plant mitochondrial DNA evolves rapidly in structure, but slowly in sequence. *J. Mol. Evol.* **28**: 87–97.
- PHADNIS, N., and H. A. ORR, 2009 A single gene causes both male sterility and segregation distortion in *Drosophila* hybrids. *Science* **323**: 376–379.
- PRESGRAVES, D. C., L. BALAGOPALAN, S. M. ABMAYR and H. A. ORR, 2003 Adaptive evolution drives divergence of a hybrid inviability gene between two species of *Drosophila*. *Nature* **423**: 715–719.
- PRESGRAVES, D. C., and W. STEPHAN, 2007 Pervasive adaptive evolution among interactors of the *Drosophila* hybrid inviability gene, Nup96. *Mol. Biol. Evol.* **24**: 306–314.
- RAND, D. M., R. A. HANEY and A. J. FRY, 2004 Cytonuclear coevolution: the genomics of cooperation. *Trends Ecol. Evol.* **19**: 645–653.
- RIESEBERG, L. H., and J. H. WILLIS, 2007 Plant speciation. *Science* **317**: 910–914.

- RITLAND, C., and K. RITLAND, 1989 Variation of sex allocation among eight taxa of the *Mimulus guttatus* species complex (Scrophulariaceae). *Am. J. Bot.* **76**: 1731–1739.
- RIVALS, E., C. BRUYERE, C. TOFFANO-NIOCHE and A. LECHARNY, 2006 Formation of the Arabidopsis pentatricopeptide repeat family. *Plant Physiol.* **141**: 825–839.
- SAMBATTI, J. B. M., D. ORTIZ-BARRIENTOS, E. J. BAACK and L. H. RIESEBERG, 2008 Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecol. Lett.* **11**: 1082–1091.
- SCHNABLE, P. S., and R. P. WISE, 1998 The molecular basis of cytoplasmic male sterility and fertility restoration. *Trends Plant Sci.* **3**: 175–180.
- SMALL, I., N. PEETERS, F. LEGEAI and C. LURIN, 2004 Predotar: a tool for rapidly screening proteomes for N-terminal targeting sequences. *Proteomics* **4**: 1581–1590.
- SWEIGART, A. L., L. FISHMAN and J. H. WILLIS, 2006 A simple genetic incompatibility causes hybrid male sterility in *Mimulus*. *Genetics* **172**: 2465–2479.
- SWEIGART, A. L., and J. H. WILLIS, 2003 Patterns of nucleotide diversity in two species of *Mimulus* are affected by mating system and asymmetric introgression. *Evolution* **57**: 2490–2506.
- TANG, S. W., and D. C. PRESGRAVES, 2009 Evolution of the *Drosophila* nuclear pore complex results in multiple hybrid incompatibilities. *Science* **323**: 779–782.
- TIFFIN, P., M. S. OLSON and L. C. MOYLE, 2001 Asymmetrical crossing barriers in angiosperms. *Proc. R. Soc. Lond. Ser. B Biol. Sci.* **268**: 861–867.
- TING, C. T., S. C. TSAUR, M. L. WU and C. I. WU, 1998 A rapidly evolving homeobox at the site of a hybrid sterility gene. *Science* **282**: 1501–1504.
- TURELLI, M., and L. C. MOYLE, 2007 Asymmetric postmating isolation: Darwin's corollary to Haldane's rule. *Genetics* **176**: 1059–1088.
- USUKA, J., W. ZHU and V. BRENDEL, 2000 Optimal spliced alignment of homologous cDNA to a genomic DNA template. *Bioinformatics* **16**: 203–211.
- VICKERY, R. K. JR., 1978 Case studies in the evolution of species complexes in *Mimulus*. *Evol. Biol.* **11**: 405–507.
- WANG, Z. H., Y. J. ZOU, X. Y. LI, Q. Y. ZHANG, L. CHEN *et al.*, 2006 Cytoplasmic male sterility of rice with boro II cytoplasm is caused by a cytotoxic peptide and is restored by two related PPR motif genes via distinct modes of mRNA silencing. *Plant Cell* **18**: 676–687.
- WANG, F., J. M. STEWART and J. ZHANG, 2007 Molecular markers linked to the Rf(2) fertility restorer gene in cotton. *Genome* **50**: 818–824.
- WANG, Z. W., Y. J. ZHANG, C. P. XIANG, S. Y. MEI, Y. ZHOU *et al.*, 2008 A new fertility restorer locus linked closely to the Rfo locus for cytoplasmic male sterility in radish. *Theor. Appl. Genet.* **117**: 313–320.
- WILLETT, C. S., 2008 Significant variation for fitness impacts of ETS loci in hybrids between populations of *Tigriopus californicus*. *J. Hered.* **99**: 56–65.
- WILLETT, C. S., and J. N. BERKOWITZ, 2007 Viability effects and not meiotic drive cause dramatic departures from Mendelian inheritance for malic enzyme in hybrids of *Tigriopus californicus* populations. *J. Evol. Biol.* **20**: 1196–1205.
- WILLETT, C. S., and R. S. BURTON, 2004 Evolution of interacting proteins in the mitochondrial electron transport system in a marine copepod. *Mol. Biol. Evol.* **21**: 443–453.
- WILLIS, J. H., 1993 Partial self-fertilization and inbreeding depression in two populations of *Mimulus guttatus*. *Heredity* **71**: 145–154.
- WITTBRODT, J., D. ADAM, B. MALITSCHKE, W. MAUELER, F. RAULF *et al.*, 1989 Novel putative receptor tyrosine kinase encoded by the melanoma-inducing *Tu* locus in *Xiphophorus*. *Nature* **341**: 415–421.
- WOLFE, K. H., W. H. LI and P. M. SHARP, 1987 Rates of nucleotide substitution vary greatly among plant mitochondrial, chloroplast, and nuclear DNAs. *Proc. Natl. Acad. Sci. USA* **84**: 9054–9058.
- WU, C. A., D. B. LOWRY, A. M. COOLEY, K. M. WRIGHT, Y. W. LEE *et al.*, 2008 *Mimulus* is an emerging model system for the integration of ecological and genomic studies. *Heredity* **100**: 220–230.
- WU, C. I., and C. T. TING, 2004 Genes and speciation. *Nat. Rev. Genet.* **5**: 114–122.
- ZHANG, J. F., and J. M. STEWART, 2001 Inheritance and genetic relationships of the D8 and D2–2 restorer genes for cotton cytoplasmic male sterility. *Crop Sci.* **41**: 289–294.

Communicating editor: D. CHARLESWORTH

GENETICS

Supporting Information

<http://www.genetics.org/cgi/content/full/genetics.109.108175/DC1>

The Nuclear Component of a Cytonuclear Hybrid Incompatibility in *Mimulus* Maps to a Cluster of Pentatricopeptide Repeat Genes

Camille M. Barr and Lila Fishman

Copyright © 2009 by the Genetics Society of America
DOI: 10.1534/genetics.109.108175

TABLE S1**Primers for markers designed on *Rf*-associated scaffolds**

Marker	Forward Primer	Reverse Primer
97_329	TCCCGAAAGGAACATTCTTG	TTTTATCCACCCGCAATTC
191_27	CGAAGAAACACAGCAGCAAG	TCGATTTTCGATCGTTTCCTT
191_45	AGCCATGTTGCAGTGTCAAG	CCGGGAGTTATAAGCCCTTT
14_172	TTTGACACCATAAGGATGGACA	CGGAAACAGAAATGATAAGGTAGGA
271_89	GGATTGAGCCGAAGAGGCTA	CGGCAATTTAAGCAACCTCA
1_36	AAGTGAATGGTCCTCCGATG	ACAACGATCCATTGTGGTGA

TABLE S2***Mimulus guttatus* genome scaffolds with >10 PPR genes, in rank order.****Linkage group anchoring was done as described in Table 1.**

Scaffold	LG	#PPRs	Scaffold size (Mb)	PPR Density (#PPR/Mb)
14	7	46	2.654	17.33
8	6	32	3.479	9.20
11	8	27	2.975	9.08
16	6	22	2.572	8.55
4	13	18	4.107	4.38
18	2	16	2.290	6.99
13	10	16	2.638	6.07
37	12	16	1.767	9.05
2	14	16	4.558	3.51
87	10	15	1.052	14.26
1	4	14	4.922	2.84
69	1	13	1.241	10.47
7	9	11	3.840	2.86
104	6	11	0.957	11.49
22	11	11	2.205	4.99
27	2	10	2.067	4.84