

## A Mouse Speciation Gene Encodes a Meiotic Histone H3 Methyltransferase

Ondrej Mihola,<sup>1\*</sup> Zdenek Trachtulec,<sup>1\*</sup> Cestmir Vlcek,<sup>1</sup> John C. Schimenti,<sup>2</sup> Jiri Forejt<sup>1†</sup>

<sup>1</sup>Institute of Molecular Genetics, Academy of Sciences of the Czech Republic, Videnska 1083, 142 20 Prague, Czech Republic.

<sup>2</sup>Center for Vertebrate Genomics, Department of Biomedical Sciences, College of Veterinary Medicine, Cornell University, T9014A Vet Research Tower, Ithaca, NY 14853, USA.

\*These authors contributed equally to this work.

†To whom correspondence should be addressed. E-mail: jforejt@img.cas.cz

**Speciation genes restrict gene flow between the incipient species and related taxa. Three decades ago, we mapped a mammalian speciation gene, hybrid sterility 1 (*Hst1*), in the intersubspecific hybrids of house mouse. Here we identify this gene as *Prdm9*, encoding a histone H3 lysine-4 trimethyltransferase. We rescued infertility in male hybrids with bacterial artificial chromosomes carrying *Prdm9* from a strain with the “fertility” *Hst1<sup>f</sup>* allele. Sterile hybrids display down-regulated microorchidia 2B (*Morc2b*) and fail to compartmentalize  $\gamma$ H2AX into the pachynema sex (XY) body. These defects, seen also in *Prdm9* null mutants, are rescued by the *Prdm9* transgene. Identification of a vertebrate hybrid sterility gene reveals a role for epigenetics in speciation, and opens a window to a hybrid sterility gene network.**

Hybrid sterility is one of the postzygotic reproduction isolating mechanisms that play an important role in speciation. Hybrid sterility is defined as a situation where parental forms, each fertile inter se, produce infertile offspring (1, 2). Hybrid sterility follows Haldane’s rule by affecting predominantly the heterogametic sex (XY or ZW) in crosses where one sex of the progeny is sterile or missing (3). Identification of speciation genes has not been particularly successful. In spite of many decades of effort, only two hybrid sterility genes have been isolated, both from *Drosophila* species (4, 5).

Here we report identification of a hybrid sterility gene in a vertebrate species. Hybrid sterility 1 (*Hst1*) is one of several genes responsible for spermatogenic failure in *Mus m. musculus* - *Mus m. domesticus* (*Mmm-Mmd*) hybrids (6, 7). It was genetically mapped to mouse Chromosome 17 (Chr17) in hybrids between the *Mmm*-derived PWD/Ph inbred strain (8) and several classical laboratory strains, mostly of *Mmd* origin (9). Whereas most laboratory inbred strains including C57BL/6J (B6) carry the *Hst1<sup>s</sup>* (sterility) allele, a few strains, such as C3H/DiSnPh (C3H) or P/J, carry the *Hst1<sup>f</sup>* (fertility) allele (see Table S1) (10). In sterile male hybrids the *Hst1*

interacts, among other genes, with *Hst1<sup>ws</sup>* locus on Chr17 of *Mmm* subspecies. However, it remains to be determined whether *Hst1* and *Hst1<sup>ws</sup>* are identical genes.

A series of high-resolution genetic mapping experiments (11–13) and haplotype analyses (14, 15) localized *Hst1* to a 255-kb single copy candidate region on Chr17, harbouring six protein-coding genes (*Dll1*, *Pgcc1*, *Psmbl1*, *Tbp*, *Pdcd2*, and *Prdm9*) and six pseudogenes (Fig. 1A). To narrow the *Hst1* critical region, we attempted rescue of the hybrid sterility phenotype by transgenesis with bacterial artificial chromosomes (BACs) derived from the C3H/HeJ strain carrying the „fertile“ *Hst1<sup>f</sup>* allele. Four overlapping BAC clones (CHORI-34-45F17 – hereafter BAC5, CHORI-34-255E14 -BAC19, CHORI-34-289M8 -BAC21, and CHORI-34-331G23 -BAC24) (16, 17) were transfected into embryonic stem (ES) cells of (129 x B6)F<sub>1</sub>, predominantly of *Mmd* origin. The mice with BAC19 did not transmit the BAC to progeny and were not pursued further. The other three BACs were transmitted, and, as expected, none of them interfered with fertility after outcrossing to the B6 laboratory strain. However, in crosses to PWD (*Mmm*) females, BAC5 and BAC24 fully restored male fertility of the F<sub>1</sub> hybrid males, whereas BAC21 transgenic F<sub>1</sub> hybrid males remained sterile (Table 1). *Psmbl1*, *Tbp*, and *Pdcd2* were excluded as *Hst1* candidates based on their expression in BAC5 (rescuing) as well as in BAC21 (non rescuing) transgenic lines (see below). *Pgcc1* was excluded because of its absence in rescuing BAC24. Previous sequencing of *Dll1*, *Pgcc1*, *Psmbl1*, *Tbp*, and *Pdcd2* alleles suggested they are unlikely candidates for *Hst1* (13, 14). Thus, the newly defined *Hst1* critical region was restricted to 15.9-kb interval (Chr17 - 15689705-15705634, NCBI Build 37.1) shared by rescuing BAC24 and BAC5 but absent in BAC21. This region is occupied by the 5’ end of the PR-domain 9 (*Prdm9*) gene and the *Mrps21-rs* pseudogene (Fig. 1B).

To exclude the possibility that BAC21 did not rescue because its genes were silenced in the BAC integration site, we analysed the BAC transgenics for the expression of the

genes within the *Hst1* candidate region. The C3H allele-specific transcript of *Tbp* was found in adult testis, proving its activity in the BAC transgene (fig. S1). *Psmbl1* and *Pdcd2* could not be tested due to the lack of suitable polymorphism, but the dosage-dependent increase of their testicular expression (Fig. 2A) suggested that the genes within the BAC21 were active but unable to rescue the meiotic arrest. A dosage-dependent increase in *Prdm9* expression was seen in the BAC5 and BAC24 hybrids but was absent in BAC21 hybrids, confirming that the *Prdm9* transcript from the *Hst1<sup>f</sup>* allele was not present in the latter (Fig. 2A). The analysis of prepubertal fertile and sterile hybrids revealed no significant differences in mRNA expression in any of the six *Hst1* candidate genes (fig. S2). These results further corroborated *Prdm9* as the only candidate gene. Novel testicular mRNA isoforms of *Prdm9* were found (fig. S3), however, none of them exhibited reproducible differential expression between prospectively fertile and sterile prepubertal hybrid testis (fig. S4).

Next, we sequenced the 25-kb region containing the C3H allele of *Prdm9* including the 5'-flanking region (GeneBank EU719625) and found 57 differences between the B6 and C3H strains: 35 microsatellite length polymorphisms, 21 single-nucleotide polymorphisms (SNPs) and one zinc-finger encoding repeat variant. All except the zinc-finger variant were in non-coding regions. Whereas PRDM9<sup>B6</sup> contained 13 C<sub>2</sub>H<sub>2</sub> zinc fingers, we found that PRDM9<sup>C3H</sup> contains 14 of them (fig. S5). The pseudogene *Mrps21-rs* is not polymorphic between C3H and B6. Thus *Prdm9* remains the only candidate for *Hst1*.

The *Prdm9* gene, also known as Meisetz, is expressed in testis and ovaries (18). It encodes a histone H3K4 trimethyltransferase. Trimethylation of histone H3K4 at promoters leads to the transcriptional activation of genes. The *Prdm9*-null mice show arrest of spermatogenesis and oogenesis at pachynema, impairment of double-strand break repair, chromosome asynapsis and disrupted sex-body formation (18). To further verify *Prdm9* as *Hst1*, the phenotypes of the sterile *Hst<sup>ws</sup>/Hst1<sup>s</sup>* (PWD x B6)F<sub>1</sub> males were compared with the published phenotypes of *Prdm9<sup>-/-</sup>* mutants. Sterile hybrids have small testes with spermatogenic arrest predominantly during pachytene and no sperm in the seminiferous tubules (6, 19, 20) (Table 1). Rare surviving primary spermatocytes at diakinesis-metaphase I manifest two to six univalents and frequent X-Y dissociation (6), resembling the impairment of synapsis between homologous chromosomes in the *Prdm9<sup>-/-</sup>* mutants. Both, sterile hybrids and *Prdm9<sup>-/-</sup>* mice display abnormal sex body formation in pachytene spermatocytes (Fig. 2B). *Prdm9<sup>-/-</sup>* pachytene spermatocytes lack a sex body, and exhibit patches of  $\gamma$ H2AX over the synaptonemal complexes (18). We observed a comparable failure of sex-body formation with scattered

$\gamma$ H2AX in 60% of pachytene spermatocytes in sterile (PWD x B6)F<sub>1</sub> hybrids versus 7% in fertile hybrid controls (Fig. 2B).

The microrchidia 2b gene, *Morc2b* or 4932411A10Rik, encoding a gonad-specific protein, is directly induced by *Prdm9*. Similar to *Prdm9<sup>-/-</sup>* testis (18) we found that *Morc2b* mRNA is barely detectable in sterile hybrids. The *Morc2b* expression in hybrid males was restored by the *Prdm9*-containing BACs (Fig. 2C). Transcription from *Morc2b* corresponds to the levels of histone H3K4 trimethylation controlled by the enzymatic activity of PRDM9 (18). Chromatin immunoprecipitation revealed decreased H3K4 trimethylation of *Morc2b* in sterile hybrid testis (Fig. 3).

Two differences were observed between sterile hybrids and *Prdm9*-null mutants. Following Haldane's rule (3) hybrid sterility is male-limited, yet meiotic arrest of *Prdm9<sup>-/-</sup>* mice affects both sexes. This discordance could be explained by incompatible *domesticus* – *musculus* epistatic interaction(s) of the *Prdm9* gene in sterile hybrids in contrast to the complete silencing of *Prdm9* in the knockout. Similar meiotic effects, sterility of both sexes or dominant male-limited sterility, have been described for the null and missense mutations of the *Dmcl1* gene (21), respectively. Secondly, the *Prdm9* null mutation acts independently, whereas meiotic arrest in F<sub>1</sub> hybrids results from the epistatic interaction of the *Hst1* gene (*Hst<sup>ws</sup>/Hst1<sup>s</sup>*) with several independently segregating genes (10). Hybrids between the consomic strain B6.PWD-Chr17 and B6 (22) as well as hybrids (B6 female x PWD male, Table 1) carry the "sterile" *Hst<sup>ws</sup>/Hst1<sup>s</sup>* genotype but are fertile due to their lack of an interaction of *Hst1* with other hybrid sterility genes (10, 23).

The parallel between the role of *Hst1* in mouse hybrid sterility and the role of *Lhr* in hybrid male inviability of *Drosophila* is quite striking. In both cases, a variant form able to rescue hybrid incompatibility was found within a species. It behaved as an autosomal locus, *Hst1<sup>f</sup>* within *Mmd* and *Lhr* in the case of *D. simulans*, and interacted with an X-linked genetic factor (7, 23, 24). Finally, both, in the mouse and *Drosophila*, the two loci were necessary but not sufficient to reconstitute the hybrid incompatibility phenotype.

Our data demonstrate that *Prdm9*, known to activate genes essential for meiosis by methylation of histone H3 at lysine 4, is the only candidate for *Hst1*. It is the only known gene located within the newly defined 15.9-kb hybrid sterility 1 critical region, expressed at the right tissue and at the right time of germ cell differentiation (primary spermatocytes). The perturbation of *Prdm9* function observed in sterile hybrid males corresponds to the phenotype of the *Prdm9<sup>-/-</sup>* mutants.

The genes that reduce hybrid fitness because of their divergent evolution can be the cause or the consequence of speciation, depending whether they evolved before or after the complete reproductive isolation of the studied taxa (25, 26). It is the advantage of our mouse model that reproductive

isolation of *Mmm* and *Mmd* is still incomplete. Thus *Prdm9* may be an essential component of a Dobzhansky-Muller incompatibility that is part of an incipient speciation event. It can lead us to the Dobzhansky-Muller incompatibility gene(s) that interfere with the normal meiotic function of histone H3K4 methyltransferase. The meiosis-specific function of *Prdm9* can explain the breakdown of meiotic cells with no effect on somatic tissues in intersubspecific hybrids. Uncovering *Prdm9* as a hybrid sterility gene will permit us to search for the epigenetically regulated downstream genes and their role in the hybrid sterility gene network.

## References and Notes

1. T. Dobzhansky, in *Genetics and the origin of Species*. (Columbia University, New York, 1951).
2. H. A. Orr, *Proc Natl Acad Sci U S A* **102 Suppl 1**, 6522 (2005).
3. J. Haldane, *Journal of Genetics* **12**, 101 (1922).
4. C. Ting, S. Tsaur, M. Wu, C. Wu, *Science* **282**, 1501 (1998).
5. J. P. Masly, C. D. Jones, M. A. Noor, J. Locke, H. A. Orr, *Science* **313**, 1448 (2006).
6. J. Forejt, P. Ivanyi, *Genet Res* **24**, 189 (1974).
7. J. M. Good, M. A. Handel, M. W. Nachman, *Evolution Int J Org Evolution* **62**, 50 (2008).
8. S. Gregorova, J. Forejt, *Folia Biol (Praha)* **46**, 31 (2000).
9. H. Yang, T. A. Bell, G. A. Churchill, F. Pardo-Manuel de Villena, *Nat Genet* **39**, 1100 (2007).
10. J. Forejt, *Trends Genet* **12**, 412 (1996).
11. S. Gregorova *et al.*, *Mamm Genome* **7**, 107 (1996).
12. Z. Trachtulec *et al.*, *Mamm Genome* **8**, 312 (1997).
13. Z. Trachtulec *et al.*, *Biological Journal of the Linnean Society* **84**, 637 (2005).
14. O. Mihola, J. Forejt, Z. Trachtulec, *BMC Genomics* **8**, 20 (2007).
15. Z. Trachtulec *et al.*, *Genetics* (2008).
16. K. Osoegawa *et al.*, *Genome Res* **10**, 116 (2000).
17. Materials and methods are available as supporting material on Science Online.
18. K. Hayashi, K. Yoshida, Y. Matsui, *Nature* **438**, 374 (2005).
19. J. Forejt, in *Current Trends in histocompatibility R*. Reisfeld, S. Ferrone, Eds. (Plenum, New York, 1981), vol. 1, pp. 103-131.
20. A. Yoshiki, K. Moriwaki, T. Sakakura, M. Kusakabe, *Development Growth & Differentiation* **35**, 271 (1993).
21. L. A. Bannister *et al.*, *PLoS Biol* **5**, e105 (2007).
22. S. Gregorova *et al.*, *Genome Res* **18**, 509 (2008).
23. R. Storchova *et al.*, *Mamm Genome* **15**, 515 (2004).
24. N. J. Brideau *et al.*, *Science* **314**, 1292 (2006).
25. R. Lewontin, *The genetic basis of evolutionary change* (Columbia University, New York, 1974).
26. H. A. Orr, D. C. Presgraves, *Bioessays* **22**, 1085 (2000).
27. We thank M. Nachman, University of Arizona, K. Paigen, The Jackson Laboratory, and D. Barbash, Cornell University for valuable comments, to P. Jansa for advice with the immunofluorescence, and P. Flachs for assistance. This work was funded by the Czech Science Foundation, the Grant Agency of the Academy of Sciences of the Czech Republic and the Ministry of Education, Youth and Sports of the Czech Republic (J.F., Z.T.), and the NIH grant (J.C.S.). J.F. was supported as International Fellow of Howard Hughes Medical Institute.

## Supporting Online Material

www.sciencemag.org/cgi/content/full/1163601/DC1

Materials and Methods

Figs. S1 to S5

Tables S1 and S2

References

22 July 2008; accepted 21 October 2008

Published online 11 December 2008;

10.1126/science.1163601

Include this information when citing this paper.

**Fig. 1.** The *Prdm9* gene encodes *Hst1*. **(A)** The co-segregating *Hst1* region is defined by the markers *CR212* and *M33* (see Table S2). The arrows point in the direction of gene transcription; the boxes stand for pseudogenes. The C3H BAC clones used for transgenesis are shown as horizontal lines with their sizes on the left. The BAC19 chimeras did not transmit the transgene. The blue lines show the BACs rescuing hybrid sterility, BAC21 did not rescue; the C region is necessary for the rescue. **(B)** The *Hst1* critical region. Dark blue boxes stand for coding exons; light blue box for an untranslated region; red boxes for alternative exons (marked 5a, 5u, B2, A, S1, and S2); grey boxes for putative pseudogenes; empty boxes or vertical black lines for repetitive sequences; asterisks for polyadenylation sites; the vertical arrow points to the site of insertion of a zinc-finger in the last exon of *Prdm9* in the C3H mouse strain. The numbers at the top indicate the positions on Chr17 (in kb, NCBI m37 assembly).

**Fig. 2.** The effect of BAC transgenes on gene expression in testes of hybrid males. **(A)** fold change of gene expression in the *Hst1* critical region in BAC carriers versus their wildtype littermates. Stars label the presence of an intact gene within the BAC. *Psmbl* and *Pdcd2* predominantly expressed in spermatogonia and pachynema show dosage-dependent increase in expression (BAC21 – 2 copies, BAC5 – 2 copies, BAC24 – 6 copies). **(B)** Indirect immunofluorescence of  $\gamma$ H2AX (red) and SCP1 (synaptonemal complexes - green) in pachytene spermatocytes. The majority of pachytene spermatocytes of fertile hybrids (*Hst<sup>wt</sup>/Hst<sup>f</sup>*, right panel)

display  $\gamma$ H2AX within the sex body. In the sterile hybrids (*Hst<sup>wt</sup>/Hst<sup>f</sup>*, left panel), the patches of  $\gamma$ H2AX are scattered over the autosomes in most of the examined cells. Three-hundred pachytene spermatocytes of each genotype were analysed. Congenic strain B10.P carries the *Hst<sup>f</sup>* allele on C57BL/10 background. (C) Transcription of *Morc2b* is induced in prepubertal fertile (PWDxB6.C3H-*Hst<sup>f</sup>*) testes but is missing in (PWDxB6)F<sub>1</sub> sterile hybrids. The expression at day 17.5 is rescued in a dosage-dependent manner in (PWDxBAC) hybrids by BAC5 and BAC24, but not by BAC21. The fertility status of F<sub>1</sub> hybrids of various combinations of inbred mouse strains is shown in Table S1.

**Fig. 3.** Histone 3 K4 trimethylation in prepubertal hybrid testis by chromatin immunoprecipitation. Gray and white columns show values for 17-day-old sterile and fertile hybrids, respectively. At the bottom, the quantified regions are indicated. The values of immunoprecipitated DNA on the y-axis are normalized by input DNA (non-immunoprecipitated positive control). The *Psmbl* promoter and the *Tbp-Pdcd2* 3'-intergenic region served as positive and negative control for H3K4 trimethylation, respectively. The *Psmbl* gene was selected as a control because it is expressed on the same levels in prepubertal hybrid testis (fig. S3). The probability (P) of the difference was determined by the Welsch's t-test; \*, P=0.02.

**Table 1.** The effect of BAC transgenes on male fertility phenotypes.

Tg line <sup>¶</sup>	Tg	Genetic background*	TW (mg)	N	Sperm count x10 <sup>-6</sup>	N	OFM	N	Fertility
BAC5	-	B6x(B6x129)	199 ± 41	5	4.5 ± 1.1	5	ND		F
BAC5	+	B6x(B6x129)	212 ± 22	7	4.6 ± 1.1	7	6.3 ± 1.2	3	F
BAC5	-	PWDx(B6x129)	63 ± 7	15	0	15	0	5	S
BAC5	+	PWDx(B6x129)	145 ± 25	22	2.4 ± 0.9	22	6.5 ± 1	7	F
BAC21	-	PWDx(B6x129)	54 ± 5	12	0	12	ND		S
BAC21	+	PWDx(B6x129)	55 ± 6	12	0	12	0	2	S
BAC24	-	PWDx(B6x129)	63 ± 8	12	0	6	ND		S
BAC24	+	PWDx(B6x129)	211 ± 16	18	6.9 ± 1.2	14	7.3	2	F
None	-	B6xPWD	152 ± 13	9	1.0	2	3.6	2	F
None	-	PWDxB6	60 ± 4	12	0	6	ND		S
None	-	PWDxC3H	128 ± 2	2	1.0	2	4.2 ± 0.7	5	F

<sup>¶</sup>Transgenic lines carried two BAC copies in BAC5 and BAC21 lines, and six in BAC24. \*B6 (*Mmd*) or PWD (*Mmm*) females were crossed with male BAC carriers on a mixture of B6 and 129 genetic background (*Hst1<sup>s</sup>*). The C3H strain and BAC clones carry the *Hst1<sup>f</sup>* allele. The presence of BAC (Tg +) was tested with an SSLP polymorphic marker. B6xPWD, reciprocal hybrid (PWD male). TW – wet weight of paired testes; OFM –offspring per female per month; ND, not determined; F, fertile, S, sterile.





