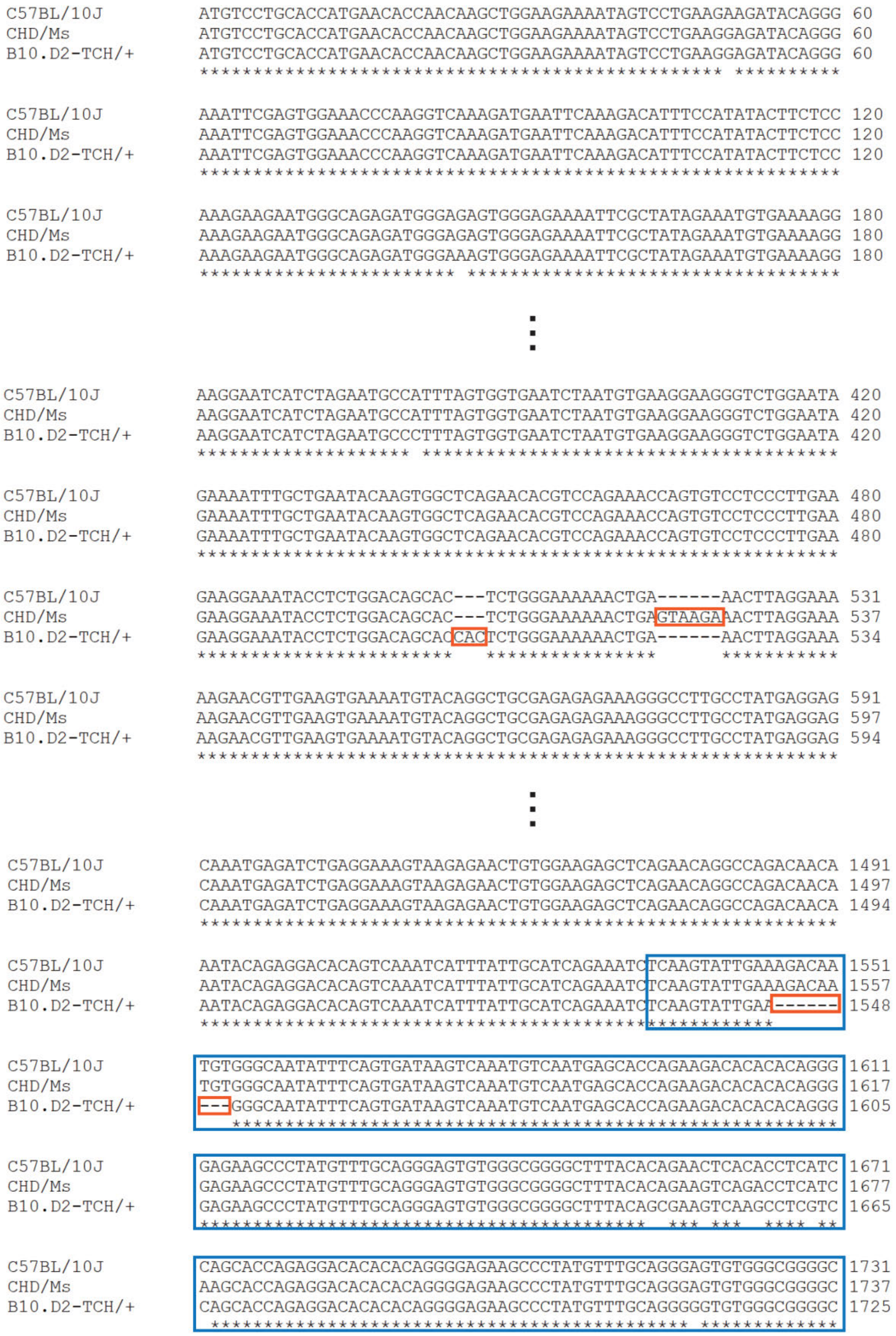
SUPPLEMENTARY INFORMATION

Prdm9 polymorphism unveils mouse evolutionary tracks

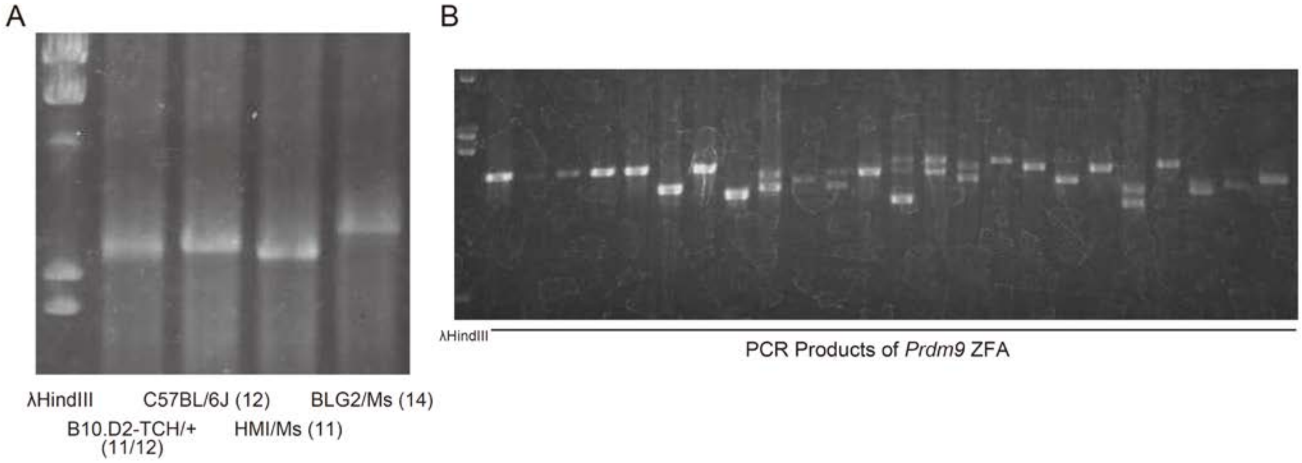
Hiromitsu Kono, Masaru Tamura, Naoki Osada, Hitoshi Suzuki, Kuniya Abe, Kazuo Moriwaki, Kunihiro Ohta and Toshihiko Shiroishi



```
C57BL/10J       ATGTCCTGCACCATGAACACCAACAAGCTGGAAGAAAATAGTCCTGAAGAAGATACAGGG 60
CHD/Ms          ATGTCCTGCACCATGAACACCAACAAGCTGGAAGAAAATAGTCCTGAAGGAGATACAGGG 60
B10.D2-TCH/+    ATGTCCTGCACCATGAACACCAACAAGCTGGAAGAAAATAGTCCTGAAGGAGATACAGGG 60
                ************************************************* **********

C57BL/10J       AAATTCGAGTGGAAACCCAAGGTCAAAGATGAATTCAAAGACATTTCCATATACTTCTCC 120
CHD/Ms          AAATTCGAGTGGAAACCCAAGGTCAAAGATGAATTCAAAGACATTTCCATATACTTCTCC 120
B10.D2-TCH/+    AAATTCGAGTGGAAACCCAAGGTCAAAGATGAATTCAAAGACATTTCCATATACTTCTCC 120
                ************************************************************

C57BL/10J       AAAGAAGAATGGGCAGAGATGGGAGAGTGGGAGAAAATTCGCTATAGAAATGTGAAAAGG 180
CHD/Ms          AAAGAAGAATGGGCAGAGATGGGAGAGTGGGAGAAAATTCGCTATAGAAATGTGAAAAGG 180
B10.D2-TCH/+    AAAGAAGAATGGGCAGAGATGGGAAAGTGGGAGAAAATTCGCTATAGAAATGTGAAAAGG 180
                *********************** ************************************

                                              ⋮

C57BL/10J       AAGGAATCATCTAGAATGCCATTTAGTGGTGAATCTAATGTGAAGGAAGGGTCTGGAATA 420
CHD/Ms          AAGGAATCATCTAGAATGCCATTTAGTGGTGAATCTAATGTGAAGGAAGGGTCTGGAATA 420
B10.D2-TCH/+    AAGGAATCATCTAGAATGCCCTTTAGTGGTGAATCTAATGTGAAGGAAGGGTCTGGAATA 420
                ******************** ***************************************

C57BL/10J       GAAAATTTGCTGAATACAAGTGGCTCAGAACACGTCCAGAAACCAGTGTCCTCCCTTGAA 480
CHD/Ms          GAAAATTTGCTGAATACAAGTGGCTCAGAACACGTCCAGAAACCAGTGTCCTCCCTTGAA 480
B10.D2-TCH/+    GAAAATTTGCTGAATACAAGTGGCTCAGAACACGTCCAGAAACCAGTGTCCTCCCTTGAA 480
                ************************************************************

C57BL/10J       GAAGGAAATACCTCTGGACAGCAC---TCTGGAAAAAAACTGA------AACTTAGGAAA 531
CHD/Ms          GAAGGAAATACCTCTGGACAGCAC---TCTGGAAAAAAACTGAGTAAGAAACTTAGGAAA 537
B10.D2-TCH/+    GAAGGAAATACCTCTGGACAGCACCACTCTGGAAAAAAACTGA------AACTTAGGAAA 534
                ************************   ****************      ***********

C57BL/10J       AAGAACGTTGAAGTGAAAATGTACAGGCTGCGAGAGAGAAAGGGCCTTGCCTATGAGGAG 591
CHD/Ms          AAGAACGTTGAAGTGAAAATGTACAGGCTGCGAGAGAGAAAGGGCCTTGCCTATGAGGAG 597
B10.D2-TCH/+    AAGAACGTTGAAGTGAAAATGTACAGGCTGCGAGAGAGAAAGGGCCTTGCCTATGAGGAG 594
                ************************************************************

                                              ⋮

C57BL/10J       CAAATGAGATCTGAGGAAAGTAAGAGAACTGTGGAAGAGCTCAGAACAGGCCAGACAACA 1491
CHD/Ms          CAAATGAGATCTGAGGAAAGTAAGAGAACTGTGGAAGAGCTCAGAACAGGCCAGACAACA 1497
B10.D2-TCH/+    CAAATGAGATCTGAGGAAAGTAAGAGAACTGTGGAAGAGCTCAGAACAGGCCAGACAACA 1494
                ************************************************************

C57BL/10J       AATACAGAGGACACAGTCAAATCATTTATTGCATCAGAAATCTCAAGTATTGAAAGACAA 1551
CHD/Ms          AATACAGAGGACACAGTCAAATCATTTATTGCATCAGAAATCTCAAGTATTGAAAGACAA 1557
B10.D2-TCH/+    AATACAGAGGACACAGTCAAATCATTTATTGCATCAGAAATCTCAAGTATTGAA------ 1548
                ******************************************************

C57BL/10J       TGTGGGCAATATTTCAGTGATAAGTCAAATGTCAATGAGCACCAGAAGACACACACAGGG 1611
CHD/Ms          TGTGGGCAATATTTCAGTGATAAGTCAAATGTCAATGAGCACCAGAAGACACACACAGGG 1617
B10.D2-TCH/+    ---GGGCAATATTTCAGTGATAAGTCAAATGTCAATGAGCACCAGAAGACACACACAGGG 1605
                   *********************************************************

C57BL/10J       GAGAAGCCCTATGTTTGCAGGGAGTGTGGGCGGGGCTTTACACAGAACTCACACCTCATC 1671
CHD/Ms          GAGAAGCCCTATGTTTGCAGGGAGTGTGGGCGGGGCTTTACACAGAAGTCAGACCTCATC 1677
B10.D2-TCH/+    GAGAAGCCCTATGTTTGCAGGGAGTGTGGGCGGGGCTTTACAGCGAAGTCAAGCCTCGTC 1665
                ******************************************  *** ***  **** **

C57BL/10J       CAGCACCAGAGGACACACACAGGGGAGAAGCCCTATGTTTGCAGGGAGTGTGGGCGGGGC 1731
CHD/Ms          AAGCACCAGAGGACACACACAGGGGAGAAGCCCTATGTTTGCAGGGAGTGTGGGCGGGGC 1737
B10.D2-TCH/+    CAGCACCAGAGGACACACACAGGGGAGAAGCCCTATGTTTGCAGGGGGTGTGGGCGGGGC 1725
                 ********************************************* *************
```

Supplementary Figure S1. ***Prdm9*** **cDNA sequences from three inbred mouse strains.**

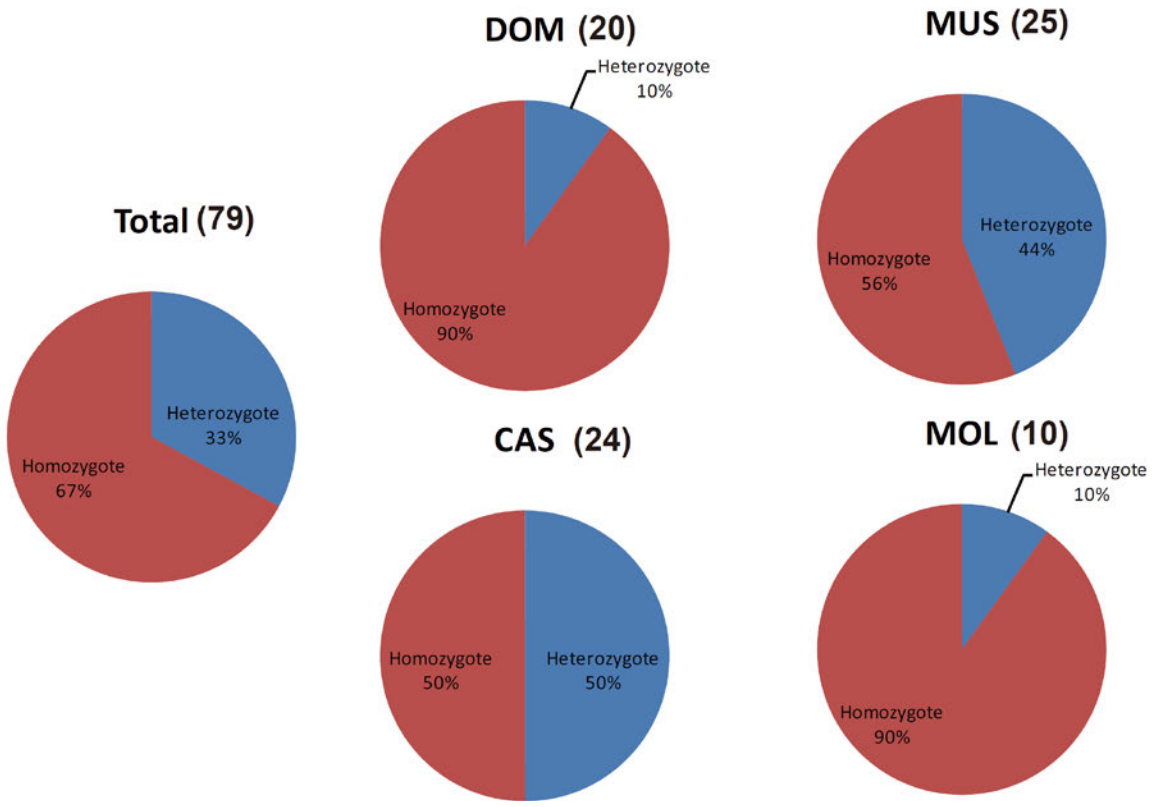
Prdm9 cDNA sequences of three inbred strains, C57BL/10J, B10.D2-TCH/+ and CHD/Ms, are aligned. Insertion and deletion of B10.D2-TCH/+ and CHD/Ms for the C57BL/6J reference sequence and C57BL/10J sequence are indicated by an orange box. The ZFA region is indicated by a blue box. B10.D2-TCH/+ has a single insertion and deletion. CHD/Ms has a single deletion.



Supplementary Figure S2. Agarose gel electrophoresis separation of PCR-amplified products.

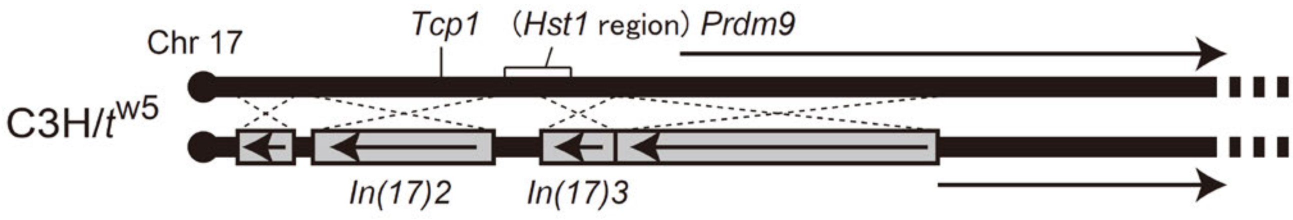
(A) PCR products of *Prdm9* cDNAs from representative strains. The strain names are shown under each lane. The number of ZF repeats in each *Prdm9* allele is shown in brackets. The B10.D2-TCH/+ stock is heterozygous for *Prdm9*.

(B) PCR products of *Prdm9* ZFA from representative wild-derived strains.



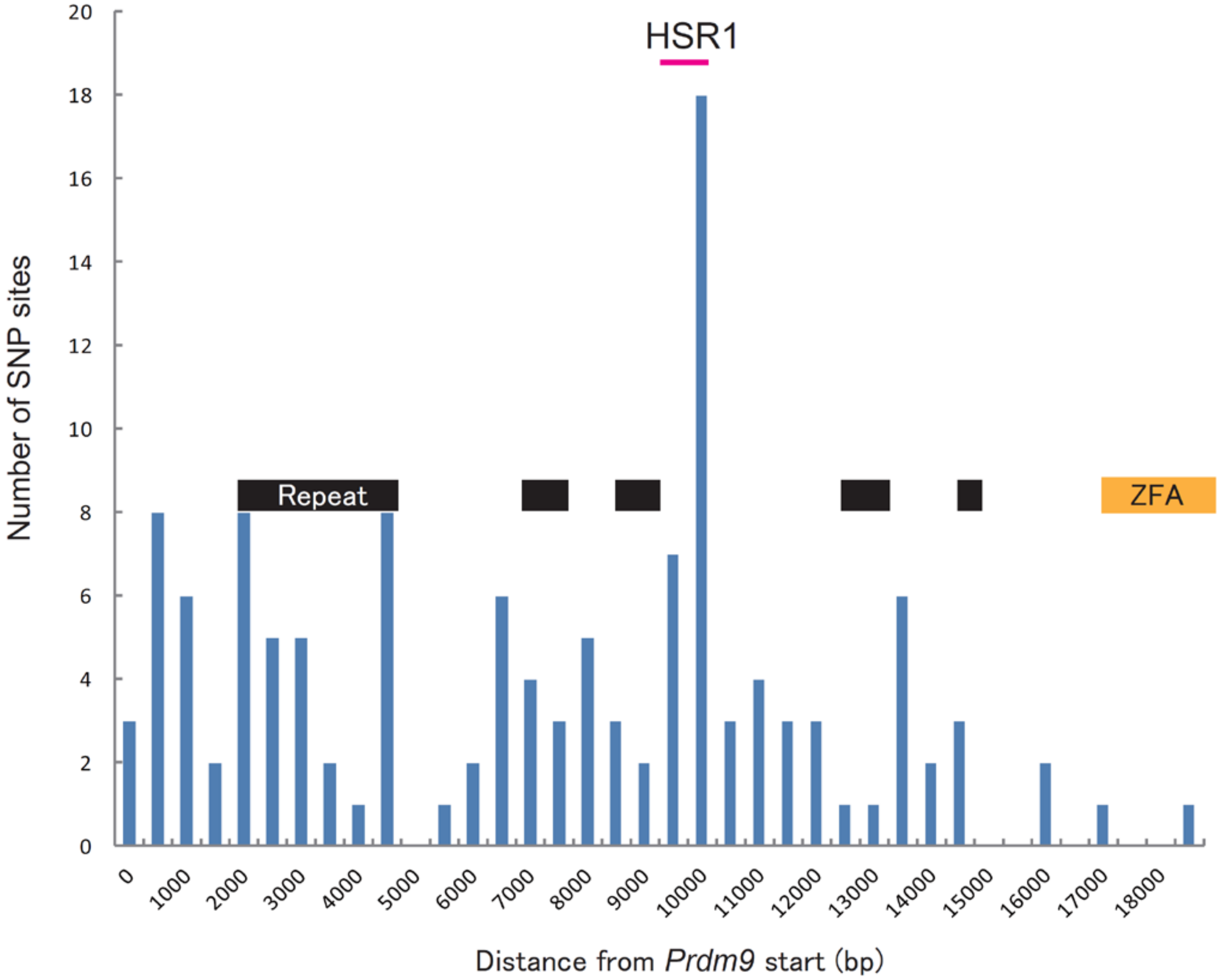
Supplementary Figure S3. Frequency of mice with heterozygous *Prdm9* alleles.

We calculated the frequency of mice with heterozygous *Prdm9* alleles using wild-captured samples. The number in parenthesis at the right side of the subspecies abbreviation is the sample number used in the calculations.



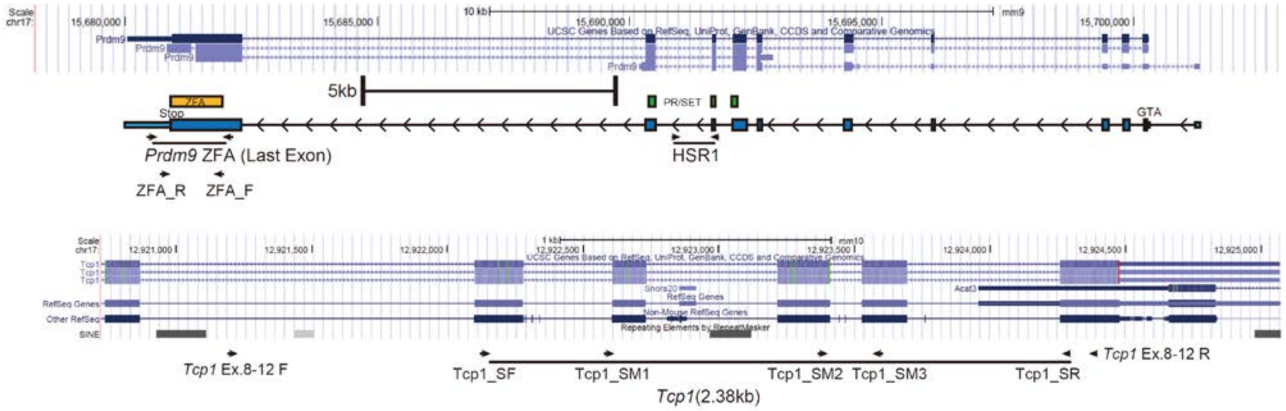
Supplementary Figure S4. Diagram of mouse *t*-haplotype chromosome.

The mouse *t*-haplotype chromosome has four inversions in chromosome 17 relative to a wild-type chromosome. The *Prdm9* (*Hst1*) locus is linked to the third inversion, *In(17)3*. *Tcp1* is located in the second inversion, *In(17)2*.



Supplementary Figure S5. Highly variable regions in the *Prdm9* genomic sequence.

Prdm9 genomic sequences of three inbred strains, C57BL/6J, CAST/EiJ and PWK/PhJ, were obtained from the Jackson Laboratory polymorphism database (http://phenome.jax.org/). The number of SNPs among these three strains were counted for every 500 bp bin, starting from the transcription start site of *Prdm9*. The most variable region resides in the intron near exon 10. A 750 bp stretch (after removing repeat sequences), referred to as High SNP Region 1 (HSR1), was used for phylogenetic analysis.



Supplementary Figure S6. Location on the genome of PCR and sequencing primers used in this study. Genomic locations of the primers used in this experiment are shown. The genomic map was obtained from the UCSC Genome Browser (http://genome.ucsc.edu/cgi-bin/hgGateway).

Supplementary Table S1. Inbred strains used for sequencing of *Prdm9* cDNA

Strain	**Origin**	**Locality of collection**		**ZFA variant type**	**Remarks**
MSM/Ms	*M. m. molossinus*	Mishima	Japan	Ma5	
JF1/Ms	*M. m. molossinus*		(Denmark)	Ma5	Japanese fancy mouse
R209-wm7	*M. m. castaneus*	Sasaguri	Japan	Ca2	
C57BL/10J	*M. m. domesticus*			Da2	
BFM/2Ms	*M. m. domesticus*	Montpellier	France	Dc1	
PGN2/Ms	*M. m. domesticus*	Pigion	Canada	Dc2	
HMI/Ms	*M. m. castaneus*	Hemei	Taiwan	Ca2	
B10.D2-TCH/+	*M. m. castaneus*	Taichung	Taiwan	Da2 / *t*	Heterozygous for *t*-haplotype
NJL/Ms	*M. m. musculus*	Northern Jutland	Denmark	Ma7	
BLG2/Ms	*M. m. musculus*	Toshevo	Bulgaria	Ma12	
CHD/Ms	*M. m. musculus*	Chengdu	China	Ma12	
SPR2	*M. spretus*			SPR2	Neighboring species

Supplementary Table S2. Samples used for sequencing of the ZFA

Sample ID (Strain)	Territory (origin)	Locality of collection		ZFA variant Type	Remarks
MG201	*M. m. molossinus*	Niigata	Japan	Ma5	Wild-captured
MG270	*M. m. molossinus*	Morioka	Japan	Ma5	Wild-captured
MG296	*M. m. molossinus*	Teine, Sapporo	Japan	Ma5	Wild-captured
MG495	*M. m. molossinus*	Hachijo Is.	Japan	Ma5	Wild-captured
MG5013	*M. m. molossinus*	Kagoshima	Japan	Ca2	Wild-captured
MG252	*M. m. molossinus*	Takatsuki	Japan	Ma5	Wild-captured
MG307	*M. m. molossinus*	Wakayama	Japan	Ma5	Wild-captured
MG254	*M. m. molossinus*	Kanazawa	Japan	Ma5 / Ca2	Wild-captured
MG442	*M. m. molossinus*	Tanegashima Is.	Japan	Ma5	Wild-captured
MG440	*M. m. molossinus*	Tsushima Is.	Japan	Ca2	Wild-captured
RFM/J	*M. m. domesticus*			Da1	Laboratory inbred strain
C57BR/J	*M. m. domesticus*			Da2	Laboratory inbred strain
SJL/J	*M. m. domesticus*			Da1	Laboratory inbred strain
P/J	*M. m. domesticus*			Da1	Laboratory inbred strain
BDP/J	*M. m. domesticus*			Da1	Laboratory inbred strain
A/J	*M. m. domesticus*			Da2	Laboratory inbred strain
129s1/SvImJ	*M. m. domesticus*			Da2	Laboratory inbred strain
LP/J	*M. m. domesticus*			Da2	Laboratory inbred strain
CBA/J	*M. m. domesticus*			Da1	Laboratory inbred strain
DBA2/J	*M. m. domesticus*			Da2	Laboratory inbred strain
SK/CamEiJ	*M. m. domesticus*	Skokholm Is.	UK	Da1	Laboratory inbred strain
PERA/EiJ	*M. m. domesticus*	Rimac Valley	Peru	Da1	Laboratory inbred strain
WSB/EiJ	*M. m. domesticus*	Centreville, Queen Anne City	USA	Da1	Laboratory inbred strain
AVZ/Ms		Ahwaz	Iran	Db5	Wild-derived inbred strain
MG305	*M. m. domesticus*	Montpellier	France	Dc1	Wild-captured
MG5188	*M. m. domesticus*	Seychelles	Seychelles	Da1	Wild-captured
MG301	*M. m. domesticus*	Mauritius Is.	Mauritius	Da3	Wild-captured
MG392	*M. m. domesticus*	Dakar	Senegal	Db2 /Dc2	Wild-captured
MG393	*M. m. domesticus*	Monastir	Tunisia	Dd1	Wild-captured
MG383	*M. m. domesticus*	Oran	Algeria	Db4	Wild-captured
MG378	*M. m. domesticus*	Haifa	Israel	Db1	Wild-captured
MG375	*M. m. domesticus*	Afula	Israel	Dc2	Wild-captured
MG394	*M. m. domesticus*	Cislago	Italy	Da1	Wild-captured
MG380	*M. m. domesticus*	Minorca	Spain	Da1	Wild-captured

MG382	*M. m. domesticus*	Cairo	Egypt	Db3	Wild-captured
MG385	*M. m. domesticus*	Corse	France	Da4	Wild-captured
MG384	*M. m. domesticus*	Annemasse	France	Da1	Wild-captured
MG430	*M. m. domesticus*	Coppock	Peru	Da3	Wild-captured
MG364	*M. m. domesticus*	Amlin	Canada	Da3	Wild-captured
MG390	*M. m. domesticus*	Azrou	Morocco	Da2	Wild-captured
HS4378	*M. m. domesticus*	Altindere	Turkey	Da1	Wild-captured
HS4379	*M. m. domesticus*	Altindere	Turkey	Da5	Wild-captured
HS4414	*M. m. domesticus*	Burra, New South Wales	Australia	Da1	Wild-captured
HS4415	*M. m. domesticus*	Burra, New South Wales	Australia	Da1 /Dc2	Wild-captured
B10.CAS3/Kfl	*M. m. castaneus*		Thailand	Ca2	H2 congenic strain
B10.CAS4(R28)/Kfl	*M. m. castaneus*		Thailand	Cd1	H2 congenic strain
CAST/EiJ	*M. m. castaneus*		Thailand	Ca2	Wild-derived inbred strain
MG421	*M. m. castaneus*	Quezon City	Philippines	Cb2	Wild-captured
MG456	*M. m. castaneus*	Bandung	Indonesia	Ce3	Wild-captured
MG5059	*M. m. castaneus*		Malaysia	Ca2 /Da1	Wild-captured
MG501	*M. m. castaneus*	Guilin	China	Ce1	Wild-captured
MG503	*M. m. castaneus*	Guangzhou	China	Ca2 /Cb4	Wild-captured
HI303	*M. m. castaneus*	Bhubaneswar	India	Cc1 /t	Wild-captured
HI306	*M. m. castaneus*	Bhubaneswar	India	Cd2/ Ca2	Wild-captured
MG5123	*M. m. castaneus*	BandarLanpung	Indonesia	Ce2	Wild-captured
MG2111	*M. m. castaneus*	Sanya	China	Ca2 /Cb2	Wild-captured
HI115	*M. m. castaneus*	Bali Denpasar	Indonesia	Ca2	Wild-captured
HI153	*M. m. castaneus*	Ujung Pandang	Indonesia	Cb2	Wild-captured
HI350	*M. m. castaneus*	Mymensingh	Bangladesh	Ce4	Wild-captured
HI366	*M. m. castaneus*	Dhaka	Bangladesh	Cb5 /t	Wild-captured
HI185	*M. m. castaneus*	Delhi	India	Ca2 /Cb1	Wild-captured
HI327	*M. m. castaneus*	Hyderabad	India	Cb1	Wild-captured
MG678	*M. m. castaneus*	Kunming	China	Da2	Wild-captured
HI276	*M. m. castaneus*	Mysore	India	Ca2 /Cd3	Wild-captured
MG5134	*M. m. castaneus*	NowShahr	Iran	Ca1 /Mb3	Wild-captured
MG5046	*M. m. castaneus*	Mashhad	Iran	Ma13	Wild-captured
HI172	*M. m. castaneus*	Islamabad	Pakistan	Cc3	Wild-captured
HI175	*M. m. castaneus*	Islamabad	Pakistan	Ca2 /Mb4	Wild-captured
HI159	*M. m. castaneus*	Leh	India	Ce5 /Cb3	Wild-captured
HI260	*M. m. castaneus*	Sahiwal	Pakistan	Ca2 /Cb5	Wild-captured

BRC3566	*M. m. castaneus*	Kabul	Afghanistan	Da5	Wild-captured
PWD/PhJ	*M. m. musculus*	Kunratice	Czech	Ma7	Wild-derived inbred strain
KJR/Ms	*M. m. musculus*	Kojuri	Korea	Ma16	Wild-derived inbred strain
MG80	*M. m. musculus*	Beijing	China	Ma2 /Ma14	Wild-captured
MG686	*M. m. musculus*	Kashgar	China	Mb1/Ma2	Wild-captured
MG716	*M. m. musculus*	Lhasa	China	Ma15 /Ma2	Wild-captured
MG3025	*M. m. musculus*	Vladivostok	Russia	Da5 /Ma9	Wild-captured
MG3030	*M. m. musculus*	Chita Region	Russia	Ma15	Wild-captured
MG3045	*M. m. musculus*	Novosibirsk	Russia	Da5	Wild-captured
MG3069	*M. m. musculus*	Magadan	Russia	Cb1	Wild-captured
MG5130	*M. m. musculus*	Bialowieza	Poland	Ma17	Wild-captured
MG740	*M. m. musculus*	Khotan	China	Mb1	Wild-captured
MG867	*M. m. musculus*	Turpan	China	Ma2	Wild-captured
MG2086	*M. m. musculus*	Jiangyin	China	Ma6	Wild-captured
MG961	*M. m. musculus*	Laiyang	China	Ma10	Wild-captured
MG3010	*M. m. musculus*	Grozny	Russia	Cc4 /Ma8	Wild-captured
MG3044	*M. m. musculus*	Tallin	Estonia	Ma9	Wild-captured
MG3057	*M. m. musculus*	Moscow	Russia	Da5	Wild-captured
MG3088	*M. m. musculus*	Samarkand	Uzbekistan	Cb6 /Cc2	Wild-captured
MG3089	*M. m. musculus*	Samarkand	Uzbekistan	Ma18 /Ma19	Wild-captured
MG3090	*M. m. musculus*	Tashkent	Uzbekistan	Ma1 /Ma11	Wild-captured
MG3104	*M. m. musculus*?	Balkhash Lake	Kazakhstan	Ma2 /Ma4	Wild-captured
MG3065	*M. m. musculus*?	Donetsk	Ukraine	Ca1 /Mb2	Wild-captured
MG2127	*M. m. musculus*?	Chongquin	China	Dc2 /Cb2	Wild-captured
MG810	*M. m. musculus*?	Tongliao	China	Ma3	Wild-captured
MG3091	*M. m. musculus*?	Termiz	Uzbekistan	Da2	Wild-captured
MG399	*M. m. musculus*?	Ljubljana	Slovenija	Ma7	Wild-captured
MG3054	*M. m. musculus*?	Poronaysk	Russia	Ma9 /*t*	Wild-captured
C3H/tw71	*M. m. domesticus*	Jutland	Denmark	Da1 /*t*	*t*-haplotype
C3H/tw75	*M. m. domesticus*	Jena	Germany	Da1 /*t*	*t*-haplotype
C3H/tw5	*M. m. domesticus*	New York	USA	Da1 /*t*	*t*-haplotype
C3H.Ttf/t12+	*M. m. domesticus*			Da1 /*t*	*t*-haplotype
Ttf/tw12+	*M. m. domesticus*			Da1 /*t*	*t*-haplotype
C3H.Ttf/t0+	*M. m. domesticus*			Da1 /*t*	*t*-haplotype
tw2/tw2	*M. m. domesticus*			*t / t*	*t*-haplotype
SEG/Pas	*M. spretus*			SEG	Neighboring species
SPRET/EiJ	*M. spretus*	Puerto Real	Spain	SPRET	Neighboring species
ZBN	*M. spicilegus*		Bulgaria	ZBN	Inbred strain

				(Neighboring species)
HS537	*M. macedonicus*	Israel	macedonicus1 / macedonicus2	Neighboring species
CAROLI/EiJ	*M. caroli*	Thailand	caroli	Other species
PAHARI/EiJ	*M. pahari*	Thailand	pahari	Other species

Supplementary Table S3. PCR primer sequence and PCR condition

Primer name	**Sequence**	**Annealing TEMP (°C)**	**Elongation time**
Prdm9 cDNA_F	TTCTAGGCTCAGGCCCGACCATAGG	56	2min
Prdm9 cDNA_R	GCTGTTGGCTTTCTCATTCTTTTCG		
Prdm9 Exon12_F	ACAACCTTTCCCAGCACACTCAAAGAAC	68	3min (2 step)
Prdm9 Exon12_R	ATGAGAATTTGCAATGGGGCTTTACTCC		
Prdm9 ZFA-L_F	GAAAGTAAGAGAACTGTGGAAGAGCTCAGAA	68	3min (2 step)
Prdm9 ZFA-L_R	GAGATGTGGTTTTATTGCTGTTGGCTTTCTC		
Prdm9 HSR1_F	CTGACACTCAGATTACCAAGGGAAGAAACTG	68	1min30sec (2 step)
Prdm9 HSR1_R	TGCCCAGGCTACCATAAAAGCTTAAGAGATC		
Tcp1 Exon8-12_F	GACATCACCCCAGCAGACAGCTTTCAC	65	2min
Tcp1 Exon8-12_R	GGCCTCTGTTGCGAACTTCAGGCTCTTC		

Supplementary Table S4. Sequencing primer sequence

Primer name	**Sequence**
M13 (-20)	GTAAAACGACGGCCAG
M13_R (-47)	CGCCAGGGTTTTCCCAGTCACGAC
M13_R	CAGGAAACAGCTATGAC
cDNA-Seq_338	GGGTGAAGCACAGTAAACAGC
cDNA-Seq_836	CCTATGAGGGTCAGATCACAGAG
cDNA-Seq_1143	CACAGCAGGAAGAGAACTAAGGA
cDNA-Seq_1450	CAGGCCAGACAACAAATACAGAG
ZFA_R	GCTGTTGGCTTTCTCATTCTTTTCG
Tcp1_SF	CAAGATAGAAGGTGCAAGGGTTGTGT
Tcp1_SM1	TTCTGGAGGTACGTCCTCTCCTGG
Tcp1_SM2	GTGGTCCCAGGTGGAGGTGCTGTAGAAG
Tcp1_SM3	GTTCCCGAGATCCCTAAGGAAGAAAC
Tcp1_SR	CAGCCTCTGTTGCGAACTTCAGGCTCTT