Supplementary Table and Figure Legends

Supplementary Table 1. Genotypes of progeny derived from mating *Sf1+/-* mice. Progeny collected from *Sf1+/-* x *Sf1+/-* intercrosses were genotyped at weaning (PN 21). No *Sf1-/-* mice were recovered ($P < 0.001$). Embryos at E11.5 were collected from females of *Sf1+/-* x *Sf1+/-* intercrosses. Genotyping revealed no *Sf1-/-* embryos at this stage ($P < 0.04$).

Supplementary Fig.1. **Testicular germ cell tumor incidences of 129 and MOLF derived mouse strains**. The boxes represent the 19 autosomal chromosomes, Chr X and Y (smaller box) of the mouse strains. White boxes represent 129 chromosomes. The *Ter* strain has a point mutation on Chr 18, represented by the green mark. MOLF chromosomes are represented in red. 129.MOLF Chr 19 strain carries Chr 19 of MOLF but all other chromosomes are of 129 origin.

Supplementary Fig.2. Inactivation of *Sf1* by gene trap targeting. (**A**) Integration site of the gene trap vector. Clear boxes are first and second exons of *Sf1*. The trap vector was inserted 261 bp downstream of the first exon. Blue boxes (β-geo) indicate tandem insertion of two copies of the trap vector. Red bar indicates the location of the probe used for Southern blotting in C. Arrows indicate the primer pairs used for genotyping in B: gt1F (5'-acgccgctgggtaagct-3'), gt4R (5'-gtgtcctacaacacacactccaacctccg-3'), v1F (5'-ctcctccatgacaaccaggtc-3'), v1R (5'-cgtaatgggataggttacgttgg-3'), i5R (5'-actctcgcgctctatcgg-3'). Triangles indicate the primer pairs geo-F (CGCCGCGGAGGACGAGGAA) and geo-R (GCCAGTTTGAGGGGACGACGACAG) used for genotyping in K. The restriction sites for *Bgl* II (B) and *EcoR* V (E) are shown together with the expected sizes of the fragments. (**B**) Genotyping for *Sf1* mutants. The primer pairs gt1F/gt4R and v1F/v1R identify the mutant allele. gt1F/i5R detects the wild-type allele. (**C**) Southern blotting for *Sf1* gene trap mice. DNA from mutant (+/-) and wild-type (+/+) genotyped mice were digested with *Bgl* II and *EcoR* V prior to hybridization using ^{32}P-labelled probe, as shown in (A). Digestion with *Bgl* II

reveals two bands of 8.6 and 10.4 kb whereas *EcoR* V digestion gives two bands of closely related sizes. No bands are observed in +/+ mice. (**D**) **LacZ expression in PN1 testes of *Sf1+/-* (*$Sf1^{\beta\text{-}geo/+}$*) mice.** X-gal staining of PN1 testes from *Sf1*+/- (*$Sf1^{\beta\text{-}geo/+}$*) mice. PN1 testes were dissected from newborn mice and stained with X-gal. (**E**) PN1 testes immunostained with germ cell specific antibody, anti-SSEA1; (**F**) and Sertoli cell specific antibody, anti-Sox9. The X-gal staining does not overlap with the Sox9-positive Sertoli cells. (**G**) ***In situ* hybridization** of PN1 testes using antisense *Sf1* probe, observed at low magnification (bar indicates 200 um) and (**I**) higher magnification (bar indicates 50 um). (**H)** Control *in situ* hybridization of PN1 testes using sense *Sf1* probe at low and (**J**) higher magnification. (**K**) (top) RT-PCR for *Sf1-geo* using total RNA from PN1 testes of *Sf1+/-* and wild-type (*Sf1*+/+) mice of *Ter/Ter*, *Ter/*+ and +/+ genotypes. PCR was performed on equal amounts of cDNA. RT+ indicates presence of Superscript during cDNA preparation.RT - are control lanes and indicates no Superscript was added. (Bottom) Control RT-PCR of GAPDH on same samples.

Supplementary Fig.3. (**A**) Schematic representation of mouse crosses between the *Ter* and *Sf1+/-* mice. Chr 18 is represented in green and Chr 19 in white. Red bar on Chr 18 represents *Ter* mutation and blue bar on Chr 19 represents the *$Sf1^{\beta\text{-}geo/+}$* gene trap allele. *Ter/+; Sf1+/-* mice were intercrossed with *Ter/*+ mice. Progeny were of genotypes listed in the Table 1. (**B**) Schematic representation of mouse crosses between the M19 and *Sf1+/-* mice. Chr 19 of M19 is represented in red and Chr 19 of *Sf1*+/- mice in white. Blue bar represents the *$Sf1^{b\text{-}geo/+}$* gene trap allele. M19 (or M19/M19) was crossed to *Sf1*+/- mice.

Supplementary Fig.4. (**A**) qRT-PCR performed on total RNA from PN1 testis of *Sf1*+/- and wild-type (*Sf1*+/+) mice of +/+, *Ter/Ter* and *Ter/*+ genotypes. *Sf1* expression is normalized against *Gapdh* expression in the testes. Error bar is the standard deviation derived from two independent experiments. (**B**) **Comparing *Sf1* levels and germ cell tumor incidences of *Ter/Ter* and *Sf1+/-***

***;Ter/Ter* strains.** The boxes represent the 19 autosomal Chrs and Chr X and Y (smaller box) of the mouse strains. White boxes represent 129 chromosomes. The *Ter* allele is represented by the green mark on Chr 18. The $Sf1^{\beta\text{-}geo}$ allele is represented by the blue mark on Chr 19. The relative *Sf1* levels are the median values from the graphs in A.

Supplementary Fig.5. LacZ expression in $Sf1^{\beta\text{-}geo/+}$ mice. Whole-mount X-gal staining of E11.5 mouse embryos: (**A**) $Sf1^{\beta\text{-}geo/+}$ and (**B**) $Sf1^{+/+}$. (**C**) Sagittal sections through the whole embryo reveals tissue-specific expression pattern of LacZ in heart; strong X-gal staining is found in the trabeculated wall of ventricular chamber (vc) and the wall of atrial chamber (ac) of the heart; (**D**) brain; the inner layer of neuroepithelial cells (n) show high LacZ expression; (**E**) eye; the nerve fiber layer (arrow) and a few cells in the neural layer (arrowhead) of the optic cup (oc) are stained positive; and (**F**) gonad; only a subset of cells are positive in the gonadal component of urogenital ridge (g); vc, ventricular chamber of the heart; ac, atrial chamber of the heart; lv, lens vesicle; oc, optic cup; cm, cephalic mesenchyme tissue; n, neuroepithelial cells forming the wall of the hindbrain; v4, fourth ventricle; m, mesonephric component of urogenital ridge; g, gonadal component of urogenital ridge. Scale bars represent 200μm in (C) and (F), 100μm in (D) and (E).