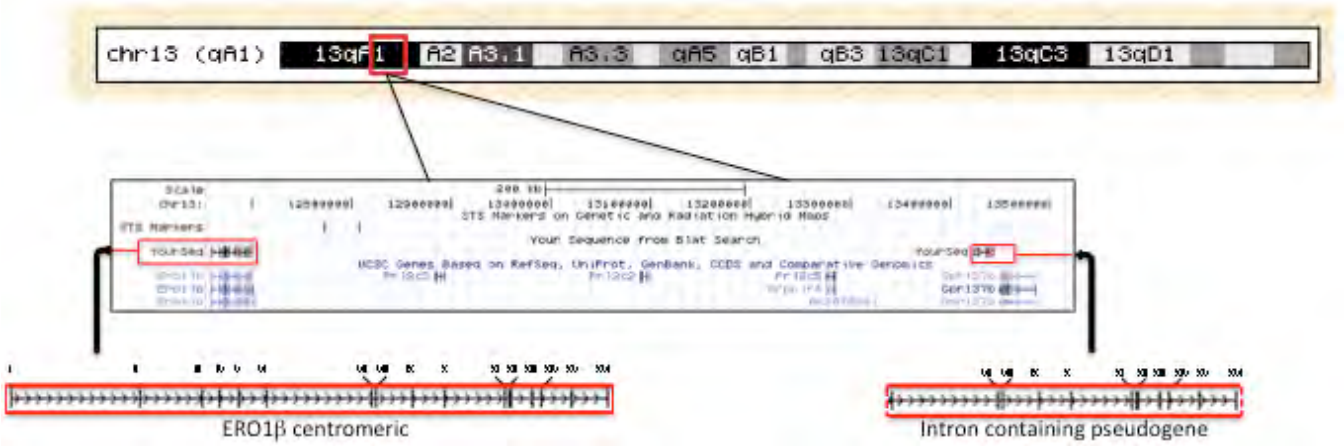
Technical supplement on genotyping the insertion allele of *Ero1β* derived from the P077G11 ES cell line

Sequencing of the genomic fragment 3' of the insertion site in the ES clone bearing the disrupted *Ero1β* (*Ero1lb*) allele P077G11 (using primers located in the insertion's Neo^r gene and in exon 15 of *Ero1β*) revealed significant divergence between the observed sequence of the 3' of intron 14 and that predicted by GenBank. Furthermore, the latest (July 2007) assembly of the mouse genome suggests the existence of large tracks of homologous DNA telomeric of *Ero1β*. This homology extends 5' of exon 6 and 3' of exon 16 (as described in the cartoon below). Experimental confirmation of this feature was provided by the observation of heterogeneity in the sequence of the 3' end of a wildtype-sized genomic DNA fragment recovered by PCR between exon 14 and exon 15 in the genomic DNA of P077G11 ES cells[1]. Further analysis revealed that the 129 mouse genome had two (non-allelic) versions of the 3' end of intron 14 (shown in the alignment below as 129^{lo} and 129^{up}, a reference to the difference in size of the two homologous fragment, figure below). The BL/6 genome also has two non-allelic versions of intron 14, tentatively named here B6$^{C\text{-}type}$ and B6$^{T\text{-}type}$, to denote the fact that one is more closely related to the centromeric sequence and the other to the more telomeric sequence found in the July 2007 build of the mouse genome.

This duplication of the mouse genome imposed a challenge in distinguishing mice heterozygous and homozygous for the P077G11 insertion into *Ero1β* (this is conventionally achieved by noting the loss of a marker specific to the wildtype allele in the homozygous disruption). To circumvent this problem we exploited a sequence polymorphism involving a *PstI* site common to both the centromeric and telomeric sequences in the BL6 and 129 genomes. In F2 progeny of a cross of 129 mutant mice to BL/6 wildtype mice the wildtype chromosome could be followed by the unique BL/6 marker: the absence of a *Pst*I in the fragment of genomic DNA arising from genomic PCR with the primers Ero1b_21S(5'TGGGTGTGTCCACCGAGGCAGTGGA3') and Ero1b_18AS (5'GATCTTCAAGGCAGTTCCTAAACCCTGAG3') (see figure below). Thus F2 mice homozygous for the 129 versions of the *Ero1β* intron 14 were tentatively assigned an *i/i* (homozygous insertion) genotype (see fig. 2A in main text). Rare recombination events between the intron 14 marker and the insertion would be picked up later by immunoblot of mouse pancreas showing the loss of ERO1β protein (no such recombination events were identified).

[1] Note that such variation could not be allelic, as the >3Kb retroviral insertion marks one of the two chromosomes of P077G11 (a diploid ES line that transmitted readily through the mouse germ line).



Cartoon Legend:

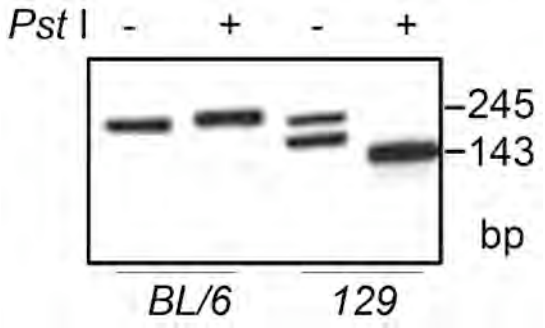
Architecture of mouse chromosome 13:

The upper panel is a low-resolution view of mouse chromosome 13. Beneath it is a close-up view of the region from 12,658,350-13,422,908, with the location of *Ero1lb* (ERO1β) and the homologous telomeric sequence (likely an intron-containing pseudo-gene).

```
129lo         1 TGGGTGTGTCCACCGAGGCAGTGGA----------------------------------  24
B6-C-type     1 TGGGTGTGTCCACCGAGGCAGTGGAGGGTGTCAGATTTCCTGGAGCTGAGGTTACAGGTG  60
129up         1 TGGGTGTGTCCACCGAGGCAGTGGAGGGTGTCAGATTTCCTGGAGCTGGGGTTCCAGGTG  60
B6-T-type     1 TGGGTGTGTCCACCGAGGCAGTGGAGGGTGTCAGATTTCCTGGAGCTGGGGTTCCAGGTG  60
                ************************

129lo        25 ----------------TGAAGAGCTGGGCGTTCCCAGCTGCAGCTGCTGGCTGTCTCTCT  69
B6-C-type    61 GTTCTGAGCCACTGGATGAAGAGCTGGGCATCCCCAGCTGC------TGGCTGTCTCTCT 114
129up        61 GTTCTGAGCCACTGGATGAAGAGCTGGGCGTCCCCAGCTGCAGCTGCTGGCTGTCTCTCT 120
B6-T-type    61 GTTCTGAGCCACTGGATGAAGAGCTGGGCATCCCCAGCTGC------TGGCTGTCTCTCT 114
                              ************* * *********      **************

129lo        70 AGCTCCCTCATATGGGTTTCAGTGTCCTCCCTGAGTAATTTTTCATAGTAGCTATTGTGT 129
B6-C-type   115 AGCTCCCTCATACGGGTTTCAGTGTCCTCCCTAAGTAATTTTTCATAGTAGCTATTGTGT 174
129up       121 AGCTCCCTCATACGGGTTTCAGTGTCCTCCCTGAGTAATTTTTCATAGTAGCTATTGTGT 180
B6-T-type   115 AGCTCCCTCATACGGGTTTCAGTGTCCTCCCTGAGTAATTTTTCATAGTAGCTATTGTGT 174
                ************ ******************* ****************************

129lo       130 AAAGTTAACTGTAAATTTGTGACTGTTTTCTTCAGACTCAGGGTTTAGGAACTGCCTTGA 189
B6-C-type   175 AAAGTTAACTGTAAATTTGTGACTGTTTTCTTCAGACTCAGGGTTTAGGAACTGCCTTGA 234
129up       181 AAAGTTAACTGTAAATTTGTGACTGTTTTCTTCAGACTCAGGGTTTAGGAACTGCCTTGA 240
B6-T-type   175 AAAGTTAACTGTAAATTTGTGACTGTTTTCTTCAGGCTCAGGGTTTAGGAACTGCCTTGA 234
                ********************************** ************************

129lo       190 AGATC 194
B6-C-type   235 AGATC 239
129up       241 AGATC 245
B6-T-type   235 AGATC 239
                *****
```

Alignment legend:

Alignment of the 3' end of intron 14 from 129 and BL/6 genomic DNA. The position of the Ero1b_21S primer is denoted in green highlight and the Ero1b_18AS antisense primer is denoted in teal highlight. Note that the 129 genome contains two clearly distinguishable versions of the intron (denoted here as 129lo and 129up) owing to a 51 base pair size polymorphism, whereas the two versions of the BL/6 gene are indistinguishable in length but differ by three base substitutions (noted here in red). Most importantly, the *Pst*I site present in the 129 versions of the intron (italicized in the alignment) is missing in both BL/6 versions of the gene, giving rise to the different migration on agarose gel of the PCR product derived from 129 and BL/6 mice, as shown in fig. 2A)



Legend

Ethidum Bromide stained agarose gel of genomic DNA fragments recovered by PCR from the *Ero1Lb* (ERO1β) locus of C57BL/6 and 129 strains of mice following digestion with *Pst*I revealing the strain-specific polymorphism at this locus.