Fig. S1. The experimental setup performed in this work is shown in a flow chart.

1. Identification of Y-chromosomal markers (red)
   - Generation of Y-chromosomal references
   - Illumina Seq of LRP products in 3 pools to identify Y-chromosomal variants
   - Validation of candidate polymorphisms by capillary sequencing of the candidate region in each horse from the seq-pools
   - Reconstruct the resultant haplotypes and visualize their phylogenetic relationship

2. Distribution of Y-chromosomal haplotypes (green)
   - Screening of Y-chromosomal haplotypes in various domestic horse breeds using the Sequenom MassARRAY iPLEX system. An additional haplotype (HT6) was discovered in this step, due to the impossibility of getting screening results for a particular locus in Shetland ponies (YE3 - Pos 1007-12040). Amplification of this region with walking primers led to the identification of the 966 bp deletion (Y_E3.1.11076-12042del). A single individual carrying HT6 and a Przewalski horse harbouring an alternative Przewalski haplotype (as described in Wallner et al., 2004) were included in the SNP-validation approach (“extended dataset”).
   - Phylogeography
   - Microsatellite analysis was performed in a subset of 100 horses representing all domestic and Przewalski horse haplotypes (including the individuals from the seq-pools)
   - Trace the contribution of domestic horse founders in a subset of horses that have a well documented pedigree

---

**Flow Chart Description**

- **Y-PCR**
  - Illumina sequencing
  - Map reads to reference
  - Identification of polymorphic sites
  - Validation of candidate polymorphisms „Extended dataset“ (n=20)
  - 18 individuals from the three seq pools
    - 1 Shetland pony (962 bp deletion see below)
    - 1 Przewalski horse
  - Reconstruct Y-chromosomal haplotypes and create haplotype network
  - 5 horse-BAC clones
  - 454 sequencing
  - De-novo assembly
  - Contig references from BAC clones
  - Reference sequence from Genbank
  - 22 LRP products total 186 kb
  - Test for male specific PCR amplification pattern

- **Pool LRP products**
  - In three seq pools
  - Pool LRP
  - Pool-breed 9 males from different breeds
  - Pool-prz 1 male Przewalski horse

- **Screening of Y-chromosomal variants**
  - With the Sequenom MassARRAY iPLEX system (n=815)

- **Y-chromosomal microsatellite analysis from a subset (n=100)**

- **Phylogeography**
  - Microsatellite analysis was performed in a subset of 100 horses representing all domestic and Przewalski horse haplotypes (including the individuals from the seq-pools)

- **Trace the contribution of domestic horse founders in a subset of horses that have a well documented pedigree**