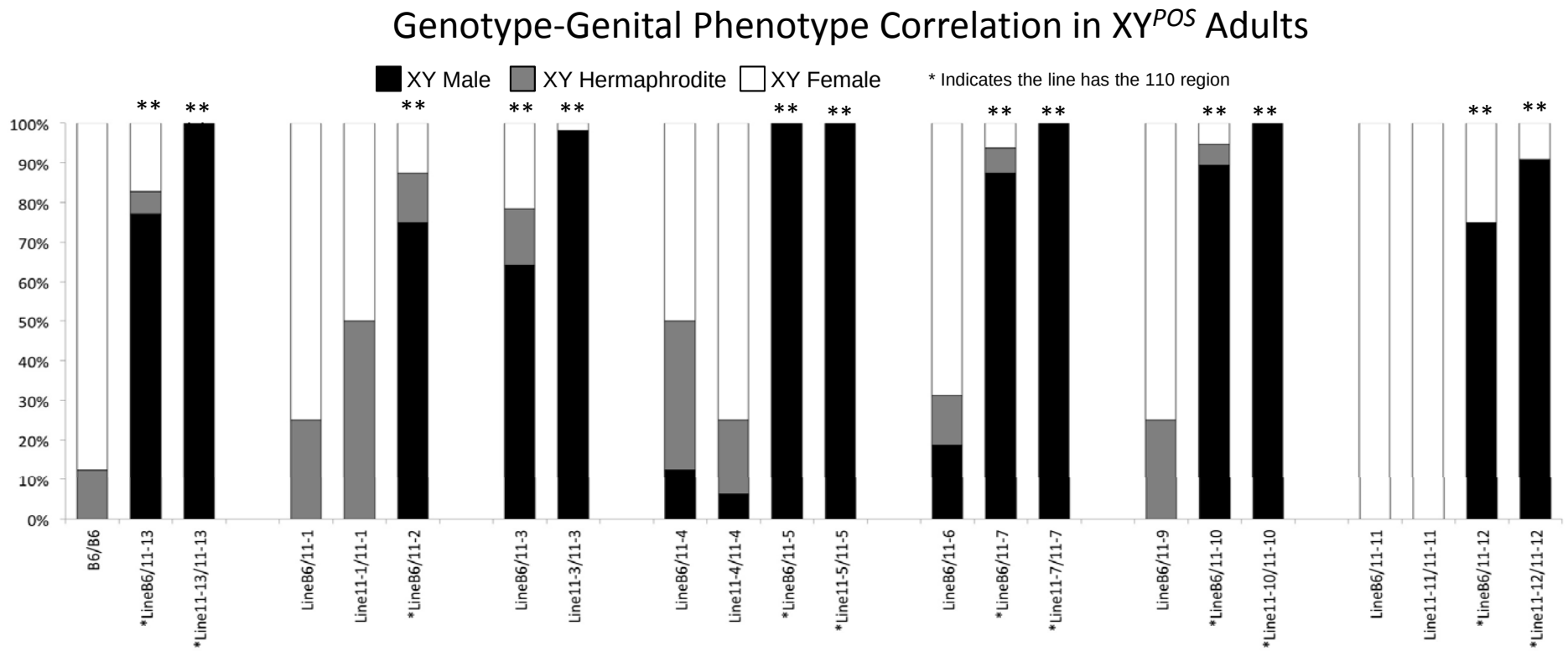
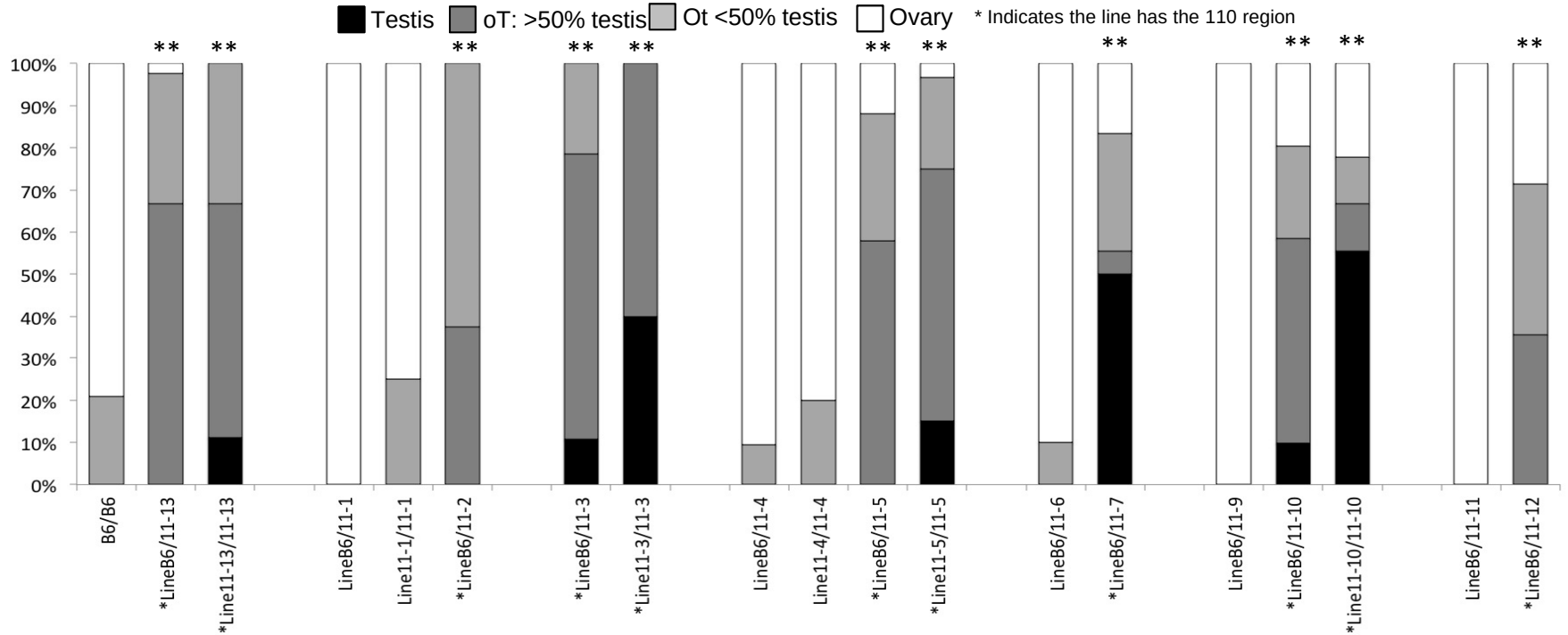


XY FEMALE	21	6	0		12	2	1		6	1		4	12	0	0		11	1	0		3	1	0		5	3	1	1
XY HERM	3	2	0		4	2	1		4	0		3	3	0	0		2	1	0		1	1	0		0	0	0	0
XY MALE	0	27	12		0	0	6		18	53		1	1	17	20		3	14	1		0	17	13		0	0	3	10
TOTAL	24	35	12		16	4	8		28	54		8	16	17	20		16	16	1		4	19	13		5	3	4	11

** Fisher exact test p-value <0.001 compared to B6/B6

Figure S1 Genotype/genital phenotype correlation in XY^{POS} adults shows that the 110 region promotes male sex determination on a B6-Y^{POS} background. Adult XY^{POS} animals that were heterozygous or homozygous for different lines were analyzed for their genital phenotype. All strains were compared to B6/B6-Y^{POS} and animals that were heterozygous for congenic regions, which did not contain the 110 region. An asterisk (*) refers to the presence of the 110 region within the congenic line and a double asterisk (**) refers to Fisher exact test p-value < 0.01. In all strains, only lines that carried the 110 region were significantly protected against sex reversal. Only a total of 5 animals were found to be phenotypically male, even in the absence of the 110 region (see LineB6/11-4, Line11-4/11-4, and LineB6/11-6). The number of animals analyzed in each group is shown below the graph, and lines that are either heterozygous or homozygous for the 110 region are highlighted in grey.

Genotype/Gonad Correlation in XY^{POS} Embryos



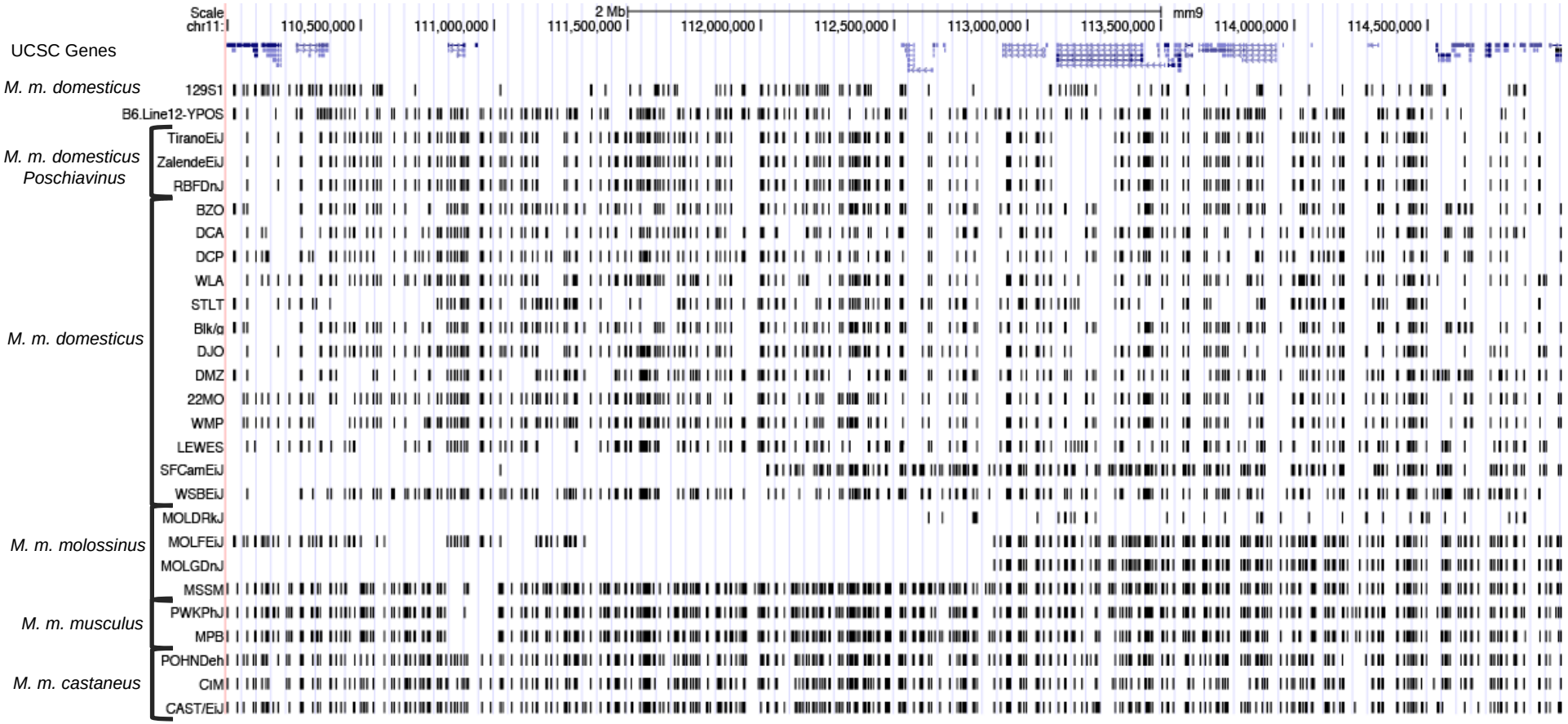
Ovary	19	1	0		6	3	0		0	0		29	8	6	2		9	3		12	8	4		16	4
Ot	5	13	3		0	1	5		6	0		3	2	15	13		1	5		0	9	2		0	5
oT	0	28	5		0	0	3		19	6		0	0	29	36		0	1		0	20	2		0	5
Testis	0	0	1		0	0	0		3	4		0	0	0	9		0	9		0	4	10		0	0
Total	24	42	9		6	4	8		28	10		32	10	50	60		10	18		12	41	18		16	14

** Fisher exact test p-value <0.001 compared to B6/B6

Figure S2 Genotype-genital phenotype correlation in XY^{POS} embryonic gonads shows that the presence of the 110 region results in a shift towards testis formation. The embryonic phenotype is a qualitative assessment closer to the critical timepoint of embryonic sex determination, which occurs between E10.5 and E12. Embryonic XY^{POS} gonads were dissected immediately after sex determination at E14.5-E15.5 and the proportion of testis formation was qualitatively assessed. All lines were compared to B6/B6-Y^{POS} and animals that were heterozygous for congenic regions containing the 110 region. An asterisk (*) refers to the presence of the 110 region within the congenic line and double asterisk (**) refers to a Fisher exact test p-value <0.01. The homozygous presence of the congenic region with 110 resulted in no Y^{POS} ovaries. Heterozygosity for a congenic region with the 110 region resulted in a shift towards increased testis formation with less formation of full ovaries.

A

Mouse Diversity Array Analysis of B6.Line12-Y^{POS} versus wild-derived strains



B

Mouse Diversity Array Analysis of B6.Line12-Y^{POS} vs other *M. m. domesticus poschiavinus strains*

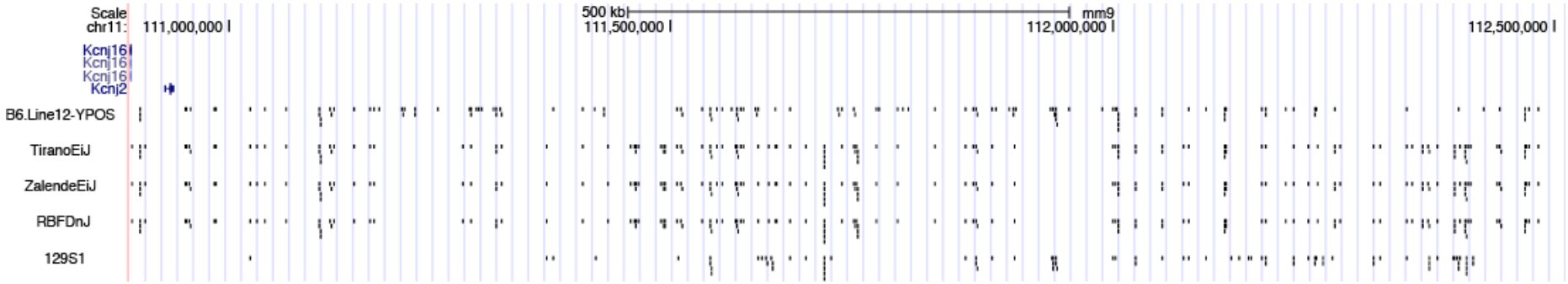


Figure S3 Mouse Diversity array analysis of B6.Line12-Y^{POS} versus wild-derived strains. A) To identify the potential origins of the protective region, we compared Diversity Array data from the data from the 129 strain (top line) to our LineB6/11-13-Y^{POS} protective congenic strain (2nd line), which contains only the 110 region, to and found that the Chromosome 11 110 region (shown here Chr11:110000000-115000000) is not derived from the 129 strain. Further analysis with array data from wild-derived strains of the 4 major subspecies of M. musculus: *domesticus, musculus, molossinus, and castaneus* identified significant similarity between 3 other typed *M. m. domesticus Poschiavinus* wild strains *Tirano*, *Zalende*, and *RBFDnJ* grouped under *M. m. domesticus Poschiavinus*. SNP data for the 129 and wild derived strains is obtained from http://cgd.jax.org/datasets/popgen/diversityarray/yang2011.shtml. B) In-depth analysis of the minimal overlap protective region derived from sub110-1 and sub110-2 (chr11:110,887,739-112,514,446) of the protected strain LineB6/11-13 and 3 genotyped animals from the *M. m. domesticus Poschiavinus* subspecies show significant similarity over the span of this non-coding region.

Predicted Transcription Factor Binding Sites in the mouse syntenic region of a second minimal region identified in cases of human DSD

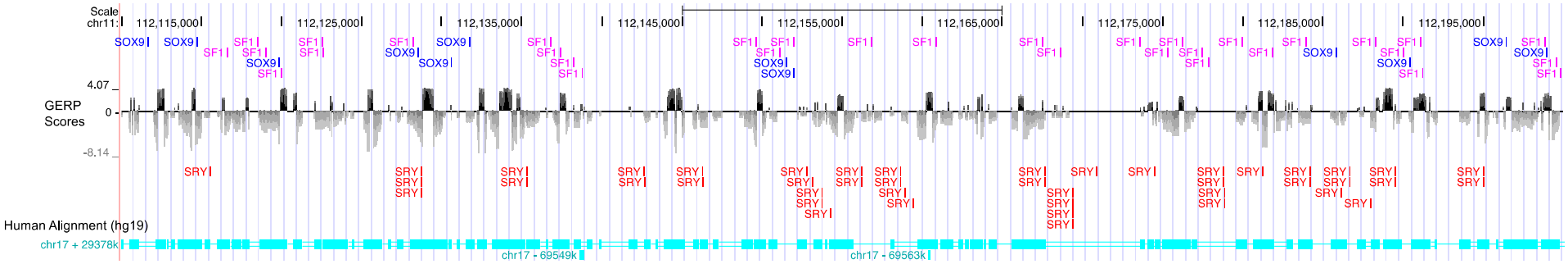


Figure S4 Sry, Sox9 and Sf1 binding sites in mouse region syntenic to the human minimal 90kb protective region identified from human cases of 46,XX and 46,XY DSD. Using MatInspector, we identified predicted transcription factor binding sites for Sry(red), Sox9(blue), and Sf1(pink) in the mouse genomie region spanning 112,110,000 to 112,200,000bp on chromosome 11. Here, 25bp regions representing predicted transcription factor binding sites regions are indicated by the vertical lines and colored to represent one of three factors, as stated above. Additionally, GERP scores were used to identify regions that are undergoing positive selection (GERP score > 0) or neutral selection (GERP score < 0). We show that the majority of predicted binding sites did not fall into highly conserved regions, based on GERP scores. At the bottom in turquoise, we show the alignment to the human genome hg19. The solid boxes indicate regions where there alignment between mouse and human reference sequence. Single lines represent gaps between the mouse and human alignment due to either deletion in the human reference or insertion in the mouse assembly. Double lines represent more complex regions that have gaps in either mouse, humans, or both. Scale bar (at top) = 20kb

Tables S1-S3
Available for download at http://www.genetics.org/lookup/suppl/doi:10.1534/genetics.113.160259/-/DC1

Table S1 PCR primers for SNP genotyping of Congenic Lines.

Table S2 Chromosome 11 results for strains genotyped by the Mouse Diversity Array at the Jackson Laboratories. Each sheet represents an individual mouse from Figure 3. Chromosome 11 position (build 37) is provided in columns B and C. Column D shows dbSNP or JAX lab identifier, if identifier known. The final column is the SNP call as performed at the Jackson laboratories. 0 = B6, 1= heterozygous, 2= homozygous for alternate non-B6 allele, -1= no call.

Table S3 Copy Number Variation (CNV) Analysis performed on a subset of animals did not identify and consistent CNVs in our congenic mice. CNV analysis was performed on the genotyped founders from Animals Line11-4/11-5, Line11-6/Line11-7, Line11-8/11-9, Line11-10/Line11-10 (see Figure 3). Results from the CNV analysis identified between 5 and 9 CNVs in each of these samples. These CNVs were classified as either a gain or a loss, but the number of copies was not quantitatively determined. Four of the CNVs recurred in most of the samples and had the same boundaries. These consisted of CNVs encompassing the variable antibody region on chromosome 12, the histocompatibility locus, erythropoietin-4 immediate early response gene, the intron of Neuroligin-1, and Lrmp. The locations of other CNVs identified in individuals are described in Table 3. Only two copy number variants were identified on chromosome 11, in the within the congenic region of the Line11-6/11-7 animal.