

**Additional file 6:** Dot plot pairwise comparisons (window size of 11) of scaffold\_295 and scaffold\_869 with the reverse complement of BAC 100j9 and PAC 34i7 (allele from an independent animal). BAC 100j9 is most likely represented by scaffold\_869, which is allelic to the corresponding VCBP1/4 region of scaffold\_295, but lacks the VCBP1/4 genes (see Additional file 5). A large (~30 kb) haplotype-specific repetitive region is found at the end of scaffold\_869 (see panel 10) and consists of alternating repetitive units, which are interspersed by exons exhibiting homology to tyrosine kinase-type receptors (possibly pseudogenes). Related repeating units (0.5 kb and larger) are distributed throughout scaffold\_295 (panel 12). Several other repeating units of highly conserved sequence are distributed across these genetic regions, as well as throughout the genome. The allele represented by PAC 34i17 contains a large ~19 kb repeat (panel 7, which is unrelated to the repeat in scaffold\_869) and varies extensively in length across alleles or could reflect an assembly artifact (in PAC 34i17). Segments of nearly identical sequence, ranging from ~0.5 to 3 kb, flank the VCBP1/4 region in both alleles of the reference genome.

Figure S5

