The “bear SAPs” dataset has the following information for each putative coding-region SNP (synonymous or non-synonymous substitution).

1. ref - dog chromosome
2. rPos - position on dog chromosome
3. trns - ENSEMBL transcript name
4. pep - ENSEMBL peptide name
5. AA1 - one amino acid
6. loc - location in the peptide sequence
7. AA2 - variant amino acid
8. %ID - percent amino acid identity of the exon
9. ex - excess number of nucleotide matches in the exon, compared to the second-best alignment
10. KEGG - KEGG pathway identifier for the protein sequence (“U” = undefined)
11. effect = PolyPhen2 prediction of the functional effect of the change (“benign”, “possibly damaging”, “probably damaging”, or “NR” = not relevant because the substitution preserves the amino acid).