Evaluating the scoring metric on different samples

Variants are split into those that do not contradict the known positions of DBs, and those that do.

