Evaluation of the scoring metric on different samples

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