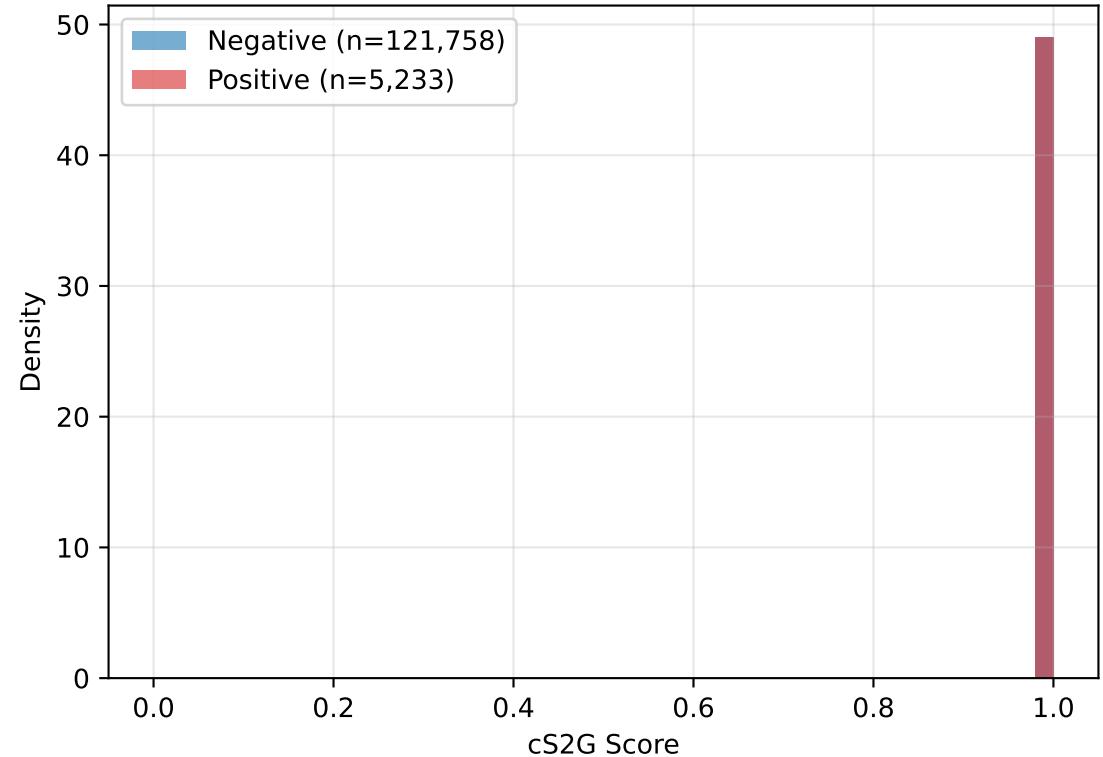
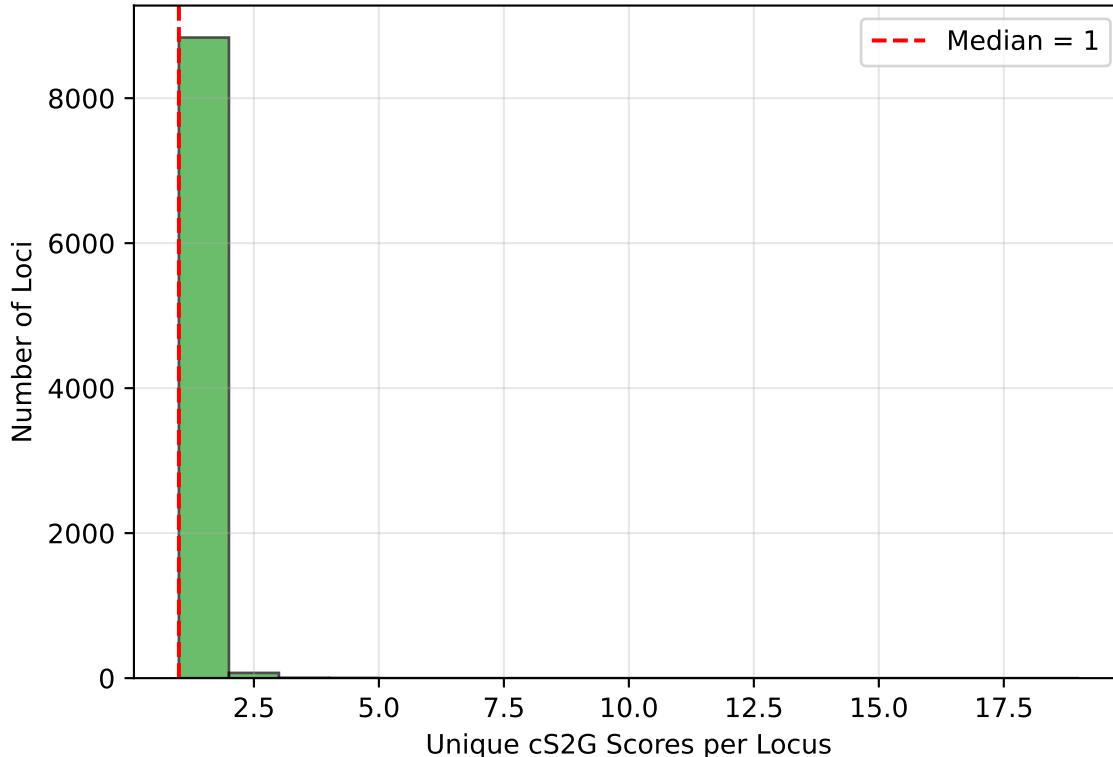


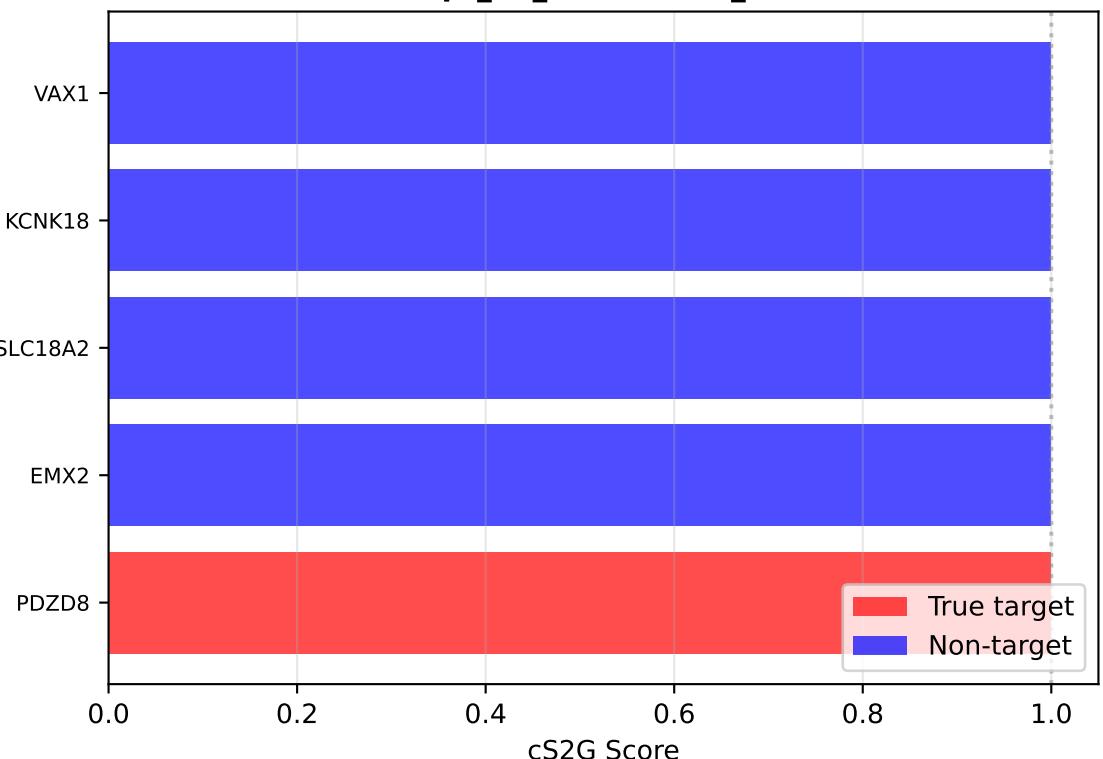
A. cS2G Score Distribution



B. Score Diversity Within Loci



C. Example Locus: All Genes Get Same Score (Locus crispr_10_117411392_117412093)



Why cS2G Shows $AUC \approx 0.50$ on RegulatoryBench

cS2G was designed for GWAS variant-to-gene assignment, providing a maximum score per gene aggregated across all linked SNPs.

The Problem:

- cS2G assigns scores to GENES, not POSITIONS
- All candidate genes at the same locus receive scores based on their own SNP linkages
- The true target gene is not scored relative to the regulatory element position

Result:

- 99.1% of loci have candidates with identical cS2G scores
- Within-locus discrimination is impossible
- $AUC \approx 0.50$ (random performance)

Key Insight:

Methods designed for GWAS-specific positions cannot be directly applied to arbitrary regulatory element-to-gene prediction.

This highlights the need for position-aware methods that score the regulatory-gene link, not just the gene itself.