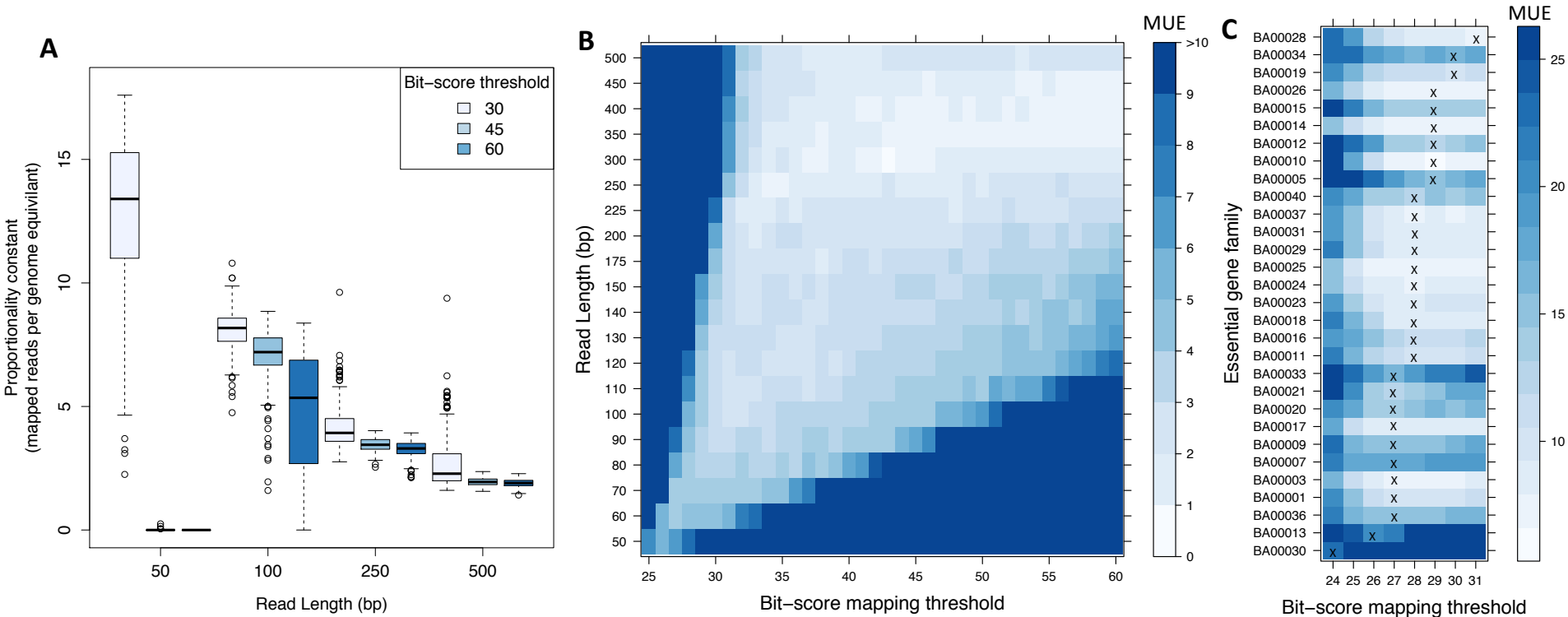




Accurate estimation of AGS depends on read length and gene family specific mapping parameters. a) Distribution of proportionality constants for gene family BA00010 across 329 training libraries. Proportionality constants depend upon read length and mapping parameters. **b)** Median unsigned error (MUE) for gene family BA00010 across 329 training libraries. AGS estimation error is minimized when using read-length specific mapping parameters. **c)** MUE for all 30 gene families across 329 70-bp training libraries. Gene families vary in their usefulness for estimating AGS. Error is minimized when using gene-family specific mapping parameters.