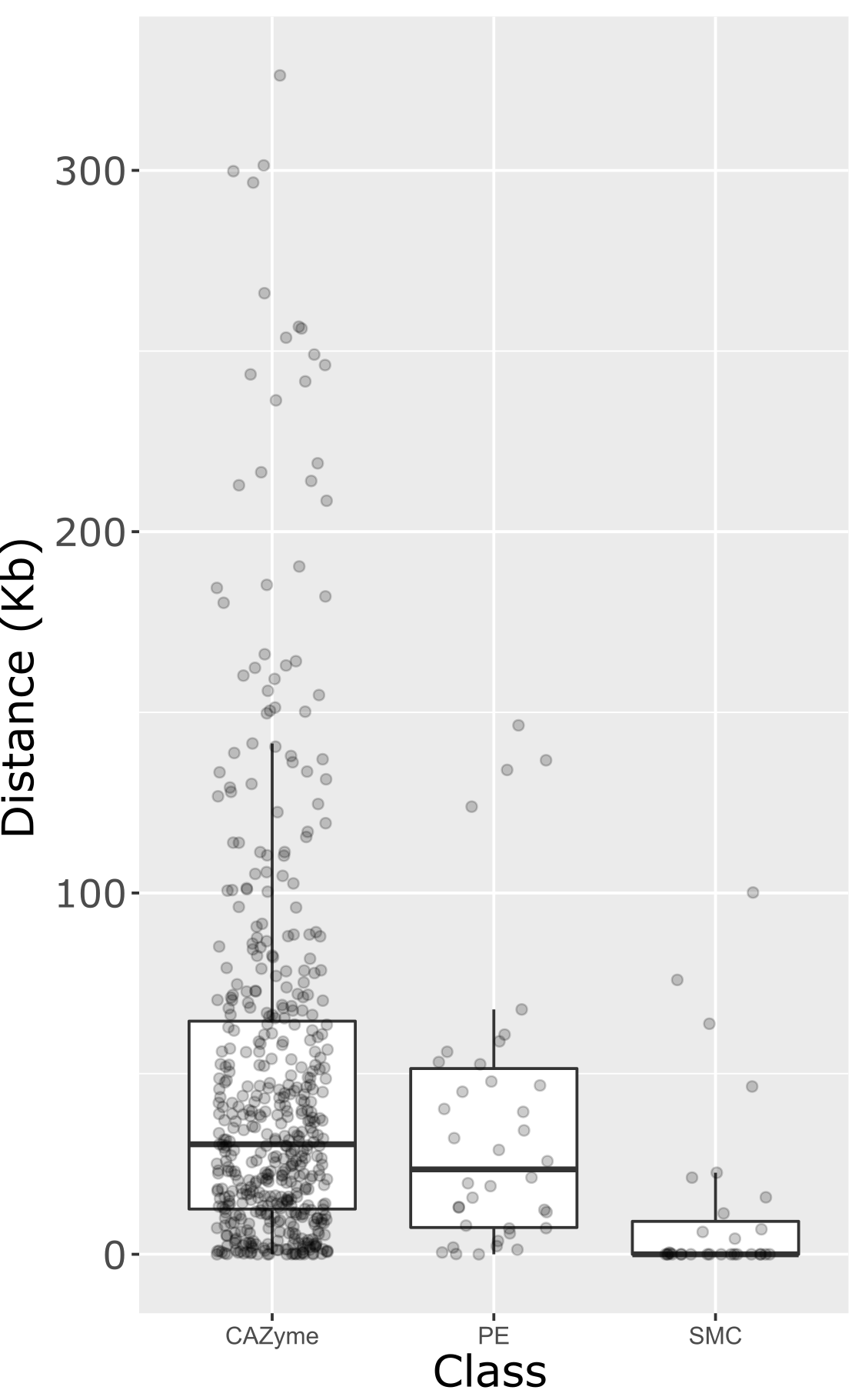
**Supplementary Figure Figure\_S4**

Boxplot of distance in Kb from the nearest contig end or AT-rich DNA region in *A. lentis* Al4, to three classes of gene: Carbohydrate-active enzymes (CAZymes), putative effector genes (PE) as predicted in the current study, and secondary metabolite gene clusters (SMC). CAZymes and putative effectors were distributed more broadly across the genome than SMCs. SMCs had a clear bias for location close to AT-rich DNA, with 18 of 31 SMCs being located directly adjacent to an AT-rich DNA region and a further five being within 10 Kb. Eleven PEs were within 10 Kb of an AT-rich region and a further 17 were within 50 Kb. Kolmogorov-Smirnov (KS) tests showed that the distribution of distances from SMCs to AT-rich DNA was significantly different to CAZymes and PEs (p-value < 0.05). However, mean and distribution of distances for CAZymes and PEs were not significantly different to each other (KS statistic D = 0.14; p-value = 0.49).