



Supplementary Figure 4: Plots of the CDS from the reference organism predicted to have (copper) or not to have (black) a putative orthologue in the comparator organism for A) *Pba1043:Dda3937* and B) *Lactococcus lactis* IL1403:MG1363 a CGH comparisons (the MPP MJ method failed to fit this data). In silico predictions derived from reciprocal best hit comparisons are labelled 'RBH' and indicated in dark copper, for comparison. Analytical methods are 'HMM': hidden Markov model described in the manuscript; GACK_100, GACK_50, GACK_0: GACK binary predictions with the EPP parameter set to 100%, 50% and 0%, respectively; MPP(BPP), MPP(EPP), MPP(MJ): MPP's BPP, EPP and MJ gene entry vector results, respectively. Qualitatively, the HMM method predicts larger contiguous islands with no putative orthologue than the other two methods, and also appears to be more successful at identifying the major genomic islands in the reference organism.