**Figure 4 – The distribution and proportion of best BLAST hits is similar for foreign and non-foreign genes across various *H. dujardini* assemblies.**

All protein sequences from one assembly were compared to all protein sequences from another assembly using BLASTp. The graphs show the distribution of each protein (represented by a blue dot) as a function of its query coverage (how much of the protein was covered by the BLAST hit) and percent identity to its best hit in the other assembly. Distributions of foreign (identified using the HGT index - right column) and non-foreign (left column) genes were compared and do not display gross differences in distribution. The number in the upper right-hand corner of each graph shows the percent of genes that align with both a query coverage and percent identity of great than or equal to 99 to their best hit in the other genome assembly. Given that the gross distribution and percent of hits with query coverage and percent identity great than or equal to 99 are approximately equivalent, it appears that foreign and non-foreign genes are equally represented between assemblies.