

Figure S1. Graphical output of Roary-derived conserved genes and pan-genome. A) Number of conserved genes as a function of the number of genomes analysed. The number of conserved genes appears to begin to plateau around 5,400. B) Number of genes in the pan-genome as a function of the number of genomes analysed. The total number of pan-genome genes continues to increase for each genome that is added and shows no plateau, suggesting an open pan-genome.

В



Figure S2. Concentration of proteins in OMVs does not correlate with number of OMV proteins. A) Concentration of proteins in OMVs from each isolate. Columns are ordered by the total number of OMV proteins. Protein concentration was measured using a Qubit Protein Assay. B) Scatterplot of protein concentration and number of identified proteins in vesicles.



Figure S3. Characterization and distribution of core and accessory pan-proteome protein sequences. Red bars are predicted signal peptide: Twin Arginine translocation; Lipoprotein; Sec and Other or Non-Classical (Other). Yellow bars are predicted subcellular location: Cytoplasmic; Inner Membrane; Periplasmic; Outer Membrane; Extracellular and Unknown. Green bars are predicted KEGG pathway. Blue bars are predicted COG orthology.