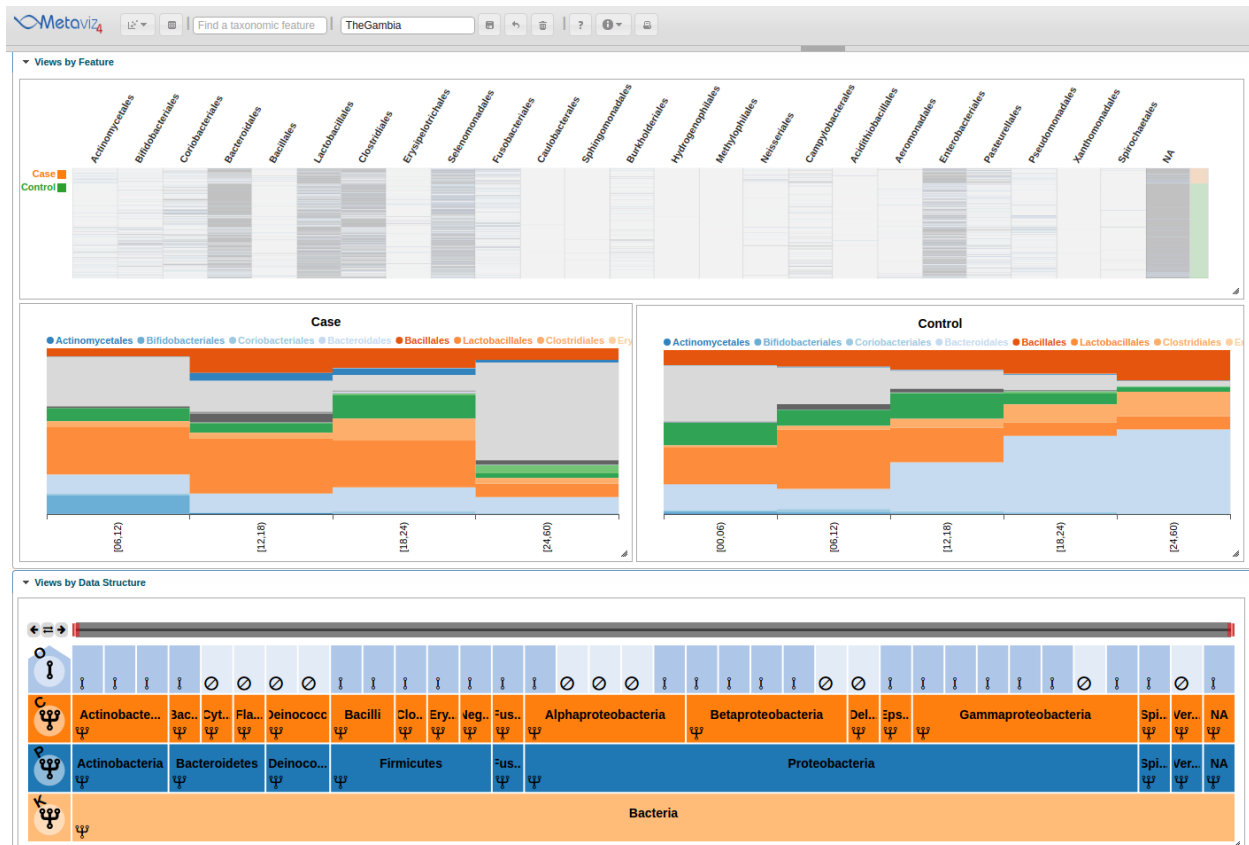
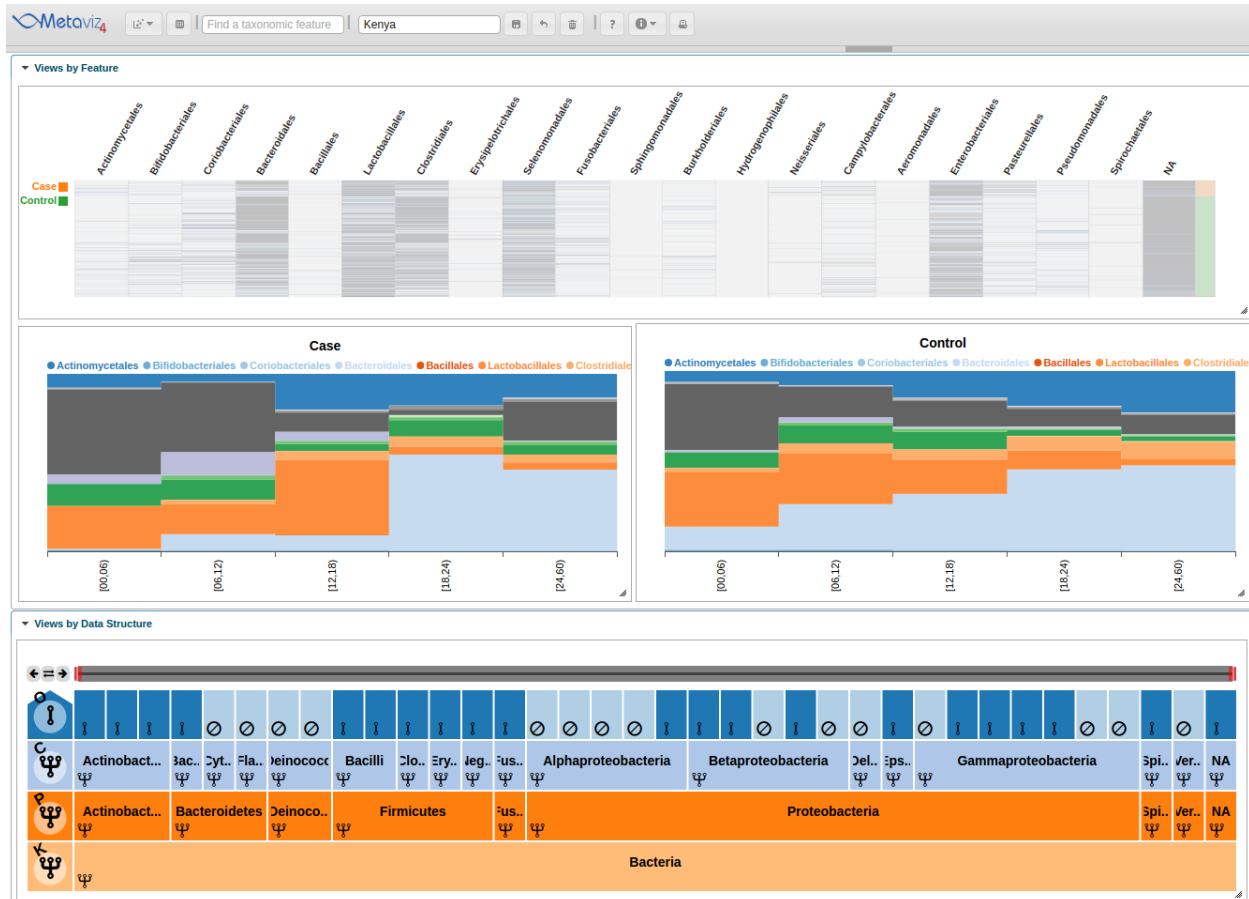




**Supplementary Figure 1: Bangladesh *msd16s* visual analysis** From the heatmap, Actinomycetales, Lactobacillales, Selemononadales, Neisseriales, Campylobacteriales, Enterobacteriales, and Pasteurellales appear more abundant in the case group than the control group. Correspondingly, Bifidobacteriales, Coriobacteriales, Bacteroidales, Clostridiales, and features without an annotation (NA) display higher abundance in the control samples as compared to the case samples. Using *metagenomeSeq*, Campylobacteriales (log fold-change: 2.30, p-value:  $2.46 \times 10^{-4}$ ), Enterobacteriales (1.51,  $6.95 \times 10^{-5}$ ), Neisseriales (1.92,  $3.55 \times 10^{-5}$ ), and Pasteurellales (2.40,  $4.96 \times 10^{-11}$ ), Coriobacteriales (-1.27,  $6.96 \times 10^{-4}$ ), Bacteroidales (-1.23,  $3.77 \times 10^{-4}$ ), and Clostridiales (-1.20,  $6.95 \times 10^{-5}$ ) are differentially abundant while Actinomycetales (.973, .007), Lactobacillales (1.24, .655), Selemonadales (.245, .511), Bifidobacteriales (-.198, .75) and NA (-.596, .752) are not. Looking at the stacked bar plots, Bacteroidales shows a higher proportion in control than case samples at all intervals after 0-6 months. Also, the control samples exhibit a greater proportion of unannotated (NA) in age intervals beyond 0-6 months. Finally, Clostridiales has lower proportion in case than control samples for the intervals of 0-6, 6-12, and 12-18 months then similar proportion for the last two time points. This is available at <http://metaviz.cbcb.umd.edu/?ws=iGPCfth9nQn>.



**Supplementary Figure 2: The Gambia *msd16s* visual analysis** From the heatmap, it appears that Lactobacillales, Campylobacteriales, Enterobacteriales, and Pasteurellales are more abundant in the case samples than control samples. Bacteroidales, Clostridiales, Burkholderiales, and Pseudomonadales are more abundant in the control samples than case. From *metagenomeSeq*, we computed the following log-fold change and adjusted p-values: Lactobacillales (0.62, 0.38), Campylobacteriales (1.10, 0.13), Enterobacteriales (1.46, 0.013), Pasteurellales (1.54,  $4.70 \times 10^{-4}$ ), Bacteroidales (-1.76,  $3.02 \times 10^{-3}$ ), Clostridiales (-.833, .175), Burkholderiales (-.0054, .99), and Pseudomonadales (.320, 865). Examining the stacked bar plots, Bacteroidales shows higher proportion in control samples than case samples for 12-18, 18-24, and 24-60 month age ranges. Lactobacillales decreases in proportion as age increases for both the case and control samples, which a much large decrease from 18-24 to 24-60 months in the case samples. In the case samples, Enterobacteriales has among the highest proportion at 0-6 months, decreases in these samples at 12-18 and 18-24 months, then has the highest proportion in the 24-60 month interval. In the case samples, NA shows lower proportion in 6-12 and 24-60 months than the 12-18 and 18-24 months intervals. The NA proportion increases in the control samples at each age interval. This workspace is available at <http://metaviz.cccb.umd.edu/?ws=Kd8O4u3zOEi>.



**Supplementary Figure 3: Kenya *msd16s* visual analysis** Examining the heatmap, Lactobacillales, Campylobacteriales, and Pasteurellales appear to be more abundant in the case samples than across the control samples. Bifidobacteriales and Clostridiales appears more abundant in control over case. Using *metagenomeSeq*, Pasteurellales has a log fold-change of 1.09 and adjusted p-value of 0.038 while Lactobacillales (.201, .948), Campylobacteriales (1.00, .245), Coriobacteriales (-.993, .073), and Clostridiales (-.898, .152) are not differentially abundant. As for changes across age ranges and case/control status, it appears that Campylobacteriales is more prevalent in 0-6, 6-12, and 12-18 in the case group than the control group. This is available at <http://metaviz.cbcb.umd.edu/?ws=asrAc9DmK2p>.



**Supplementary Figure 4: Mali *msd16s* visual analysis** Bifidobacteriales and Coriobacteriales show greater abundance in control samples as compared to the case samples. Pasteurellales displays increased abundance in the case samples as compared to the distribution in the control samples. With *metagenomeSeq*, we find support for these conclusions with Pasteurellales (2.82,  $5.4 \times 10^{-4}$ ), Neisseriales (1.77,  $1.4 \times 10^{-2}$ ), and Bifidobacteriales (-1.59, .439) but not with Coriobacteriales (-.848, .439). From the stacked plots, the proportion of Enterobacteriales among case samples in age range 6-12 and 12-18 months is much higher than that in the similar age ranges in the control samples. In case samples, Pasteurelles shows higher proportion in the case samples as compared to the controls in the 18-24 age range. For all age ranges, Bacteroidales displays greater proportion in the control compared to case samples. This figure is available at <http://metaviz.cbcb.umd.edu/?ws=EUARocVProf>.