

Figure S1. 3D PCoA plot of unweighted UniFrac distances of oral and stool samples. A 3-dimensional principal coordinates analysis plot of the unweighted Unifrac distances of the 737 oral samples and 551 fecal samples using the package plot3Drgl in R. Oral (blue) and stool (orange) samples are colored by sample site.



Figure S2. 3D PCoA plot of weighted UniFrac distances of oral and stool samples. A 3dimensional principal coordinates analysis plot of the weighted Unifrac distances of the 737 oral samples and 551 fecal samples using the package plot3Drgl in R. Oral (blue) and stool (orange) samples are colored by sample site.



Figure S3. Oral and stool community coalescence is defined by using intra-patient minimum UniFrac distances between oral-stool pairs. Each data point represents the minimum UniFrac distance (from the matrix) between any oral-stool pair from the longitudinal sampling within an individual patient. Plotted is the median and interquartile range. Those that fell below the interquartile range were defined as exhibiting oral-stool microbial community coalescence.



Figure S4. The relative abundance of oral associated bacteria in oral and stool samples of AML patients. The relative abundance of selected oral bacteria (A-B) *Veillonella*, (C-D) *Oribacterium*, (E-F) *Neisseria*, (G-H) *Actinomyces*, (I-J) *Leptotrichia*, and (K-L) *Fusobacterium* was plotted for oral and stool samples. The color scale indicates sampling time point for each of the patient samples, where blue is the first timepoint sampled, and purple is the last sampling time point. P values were calculated between oral and stool samples utilizing a mixed model for repeated measures.



Figure S5. Differential baseline taxonomic abundances among patients who coalesed and did not coalese. Linear discriminant analysis of effect size (LEfSe) of taxonomic abundances for baseline samples of coalesced (1) and not coalesced (0) patients. Horizontal bars represent the effect size for each taxon, the length of each bar represents the log10 transformed LDA score, indicated by the vertical dotted lines. The grouped data was analyzed using the Kruskal-Wallis test with a significant set to 0.05 to determine differentially abundant taxa. Differentially abundant taxa that are present in patients that experienced coalescence are indicated by green, and taxa that are present in patients who did not experience coalescence appear in red. The threshold for the logarithmic LDA score for discriminant features was set to 2.0. A) is stool samples. B) is oral samples.



Figure S6. Summary plot of patient antibiotic administration, sampling period, and coalescence determination. The x-axis shows the duration of sampling and antibiotic administration. Colored dots and lines represent the first administration and continuation of antibiotic treatment. The grey boxes show the sampling period for each patient. The coalescence designation is represented by pink and blue boxes. If coalescence was observed, the time of coalescence is represented within the grey box by a blue band.