



Figure 4. Charting biosynthetic potential of the EPIC library. (A) Relative abundance of different biosynthetic gene cluster (BGC) types found across eight body sites. Y-axis depicts relative abundance, while x-axis indicates distinct body sites. Each color within the bars corresponds to a unique BGC type. (B) A phylogeny of the 182 isolates from our study with the inner track corresponding to the genus of the isolate and the outer barplot depicting the number of BGCs identified in the genome of each isolate by antiSMASH. (C) A network of 305 GCF (nodes) identified in EPIC isolates using BiG-SCAPE. Edges indicate common antiSMASH-based BGC types shared between GCFs. The size of the nodes corresponds to the number of genomes a GCF was found in and the color composition indicates proportion of BGCs belonging to a GCF originating from EPIC, SMGC, or MIBiG. (D) Cumulative GCF discovery by BiG-MAP is plotted as a function of aggregate sequencing depth at each body site across multiple individuals using skin metagenomes from Swaney et al. 2022. Metagenomes are ordered from lowest to highest sequencing depth.