Figures and figure supplements

Panton–Valentine leucocidin is the key determinant of *Staphylococcus aureus* pyomyositis in a bacterial GWAS

Bernadette C Young et al
Figure 1. Phylogeny of S. aureus cultured from children in Cambodia shows strong strain-to-strain variation in pyomyositis prevalence. The phylogeny was estimated by maximum likelihood from SNPs mapping to the USA300 FPR3757 reference genome. Multi-locus sequence type (ST) or clonal complex (CC) groups are shown (outer grey ring). Case/control status is marked in the middle ring: pyomyositis (gold, n = 101) or nasal carriage (green, n = 417). Branches of the phylogeny that correspond to the three principal components (PCs) significantly associated with case/control status (PCs 1, 2 and 20) are marked in red, blue and yellow, respectively. Branch lengths are square root transformed to aid visualization. The presence of the kmers most strongly associated with pyomyositis is indicated by red blocks in the inner ring.

DOI: https://doi.org/10.7554/eLife.42486.003
Figure 1—figure supplement 1. Sampling frequencies of the major strains were stable over time. The year of sampling (2007–2008, blue shaded lines) and MRSA status (orange lines) are illustrated around the phylogeny of the bacteria sampled from pyomyositis cases (gold lines) and asymptomatic carriage controls (green lines). The three PCs most significantly associated with case/control status are also shown (PCs 1, 2 and 20 by red, blue and yellow branches respectively). Branch lengths are square root transformed to aid visualization.

DOI: https://doi.org/10.7554/eLife.42486.004
Figure 2. Kmers associated with pyomyositis. (A) All kmers (n = 10,744,013) were mapped to the genome assembly of one CC121 pyomyositis bacterium (PYO2134). Each point represents a kmer, plotted by the mapped location and the significance of the association with disease (-log_{10} p value). Kmers are coloured by the odds ratio (OR) of kmer presence for disease risk. A Bonferroni-adjusted threshold for significance is plotted in grey. (B) The region between 1.57–1.62 MB in greater detail. Grey arrows depict coding sequences, determined by homology to USA300 FPR3757. (C) Associations for kmers mapping to region 1,571–1,574 kB is plotted. Kmer presence determined from hybrid assembly using short and long-read data. Grey arrows depict coding sequences, determined by homology to USA300 FPR3757.

DOI: https://doi.org/10.7554/eLife.42486.005
Figure 2—figure supplement 1. Alignments of reference genome PYO2134 assembly (R) with 37 de novo assemblies of Illumina short-read sequencing (C) which feature either ambiguities (Ns) or contig boundaries in the region 389 bp upstream of PVL coding sequence. Contig boundaries, when overlapping, are marked with a red diamond. Ns in the assembly are marked in dark grey. Polymorphisms are colour-coded by base.

DOI: https://doi.org/10.7554/eLife.42486.006
Figure 2—figure supplement 2. Alignments of reference genome PYO2134 assembly (R) with 37 de novo assemblies of Illumina short-read sequencing (C) which feature either ambiguities (Ns) or contig boundaries in the region 389 bp upstream of PVL coding sequence. Contig boundaries, when overlapping, are marked with a red diamond. Ns in the assembly are marked in dark grey. Polymorphisms are colour-coded by base.

DOI: https://doi.org/10.7554/eLife.42486.007
Figure 2—figure supplement 3. Presence of PVL and potential PVL carrying phages across the population. Multi-locus sequence type (ST) or clonal complex (CC) groups are shown (outer grey ring). Case/control status is marked in the next outermost ring: pyomyositis (gold, n = 101) or nasal carriage (green, n = 417). Branch lengths are square root transformed to aid visualization. The presence of kmers mapping to the PVL coding sequence (dark red) or of >98% coverage of PVL coding sequence in de novo assembly found by BLAST (red) are marked in the next rings. Phages with >95% coverage on BLAST of the de novo assembly are marked by points (ϕPVL in purple, ϕSaUSA2 in dark blue, ϕSLT in light blue and ϕCC121 in orange), and bars are used to mark the phage with the greatest sequence homology.

DOI: https://doi.org/10.7554/eLife.42486.008
Figure 2—figure supplement 4. Significant association between case/control status and the presence of kmers mapping to PVL was consistent across early and late subsets. Results from repeat kmer GWAS using early and late sub-groups of cases and controls, as well as results from original GWAS of the entire group. We report the presence of kmers mapping across the PVL locus in each subset. Forest plot showing the effect sizes (box) with 95% CI (whiskers) determined by LMM for kmers mapping to PVL kmer in each subset, and in the complete study (diamond centred on effect size extending across 95% CI). The statistical significance of association with case/control status after controlling for population structure is shown for each subset and the complete study.

DOI: https://doi.org/10.7554/eLife.42486.009