

1 **Supplementary Information**

2 **Spatiotemporal dynamics of multidrug resistant bacteria on intensive care unit surfaces**

3 Alaric W. D'Souza^{1*}, Robert F. Potter^{1*}, Meghan Wallace², Angela Shupe², Sanket Patel^{1,2}, Xiaoqing Sun^{1,2},

4 Danish Gul³, Jennie H. Kwon⁴, Saadia Andleeb^{3**}, Carey-Ann D. Burnham^{2,4,5,6**}, Gautam Dantas^{1,2,5,7**}

5

6 *These authors contributed equally to this work.

7

8 **Corresponding Authors

9 (Saadia Andleeb: saadiamarwat@yahoo.com, Carey-Ann D. Burnham: cburnham@wustl.edu, Gautam Dantas:

10 dantas@wustl.edu)

11

12 ¹The Edison Family Center for Genome Sciences and Systems Biology, Washington University School of
13 Medicine, St. Louis, MO, USA

14 ²Department of Pathology and Immunology, Washington University School of Medicine, St. Louis, MO, USA

15 ³Atta ur Rahman School of Applied Biosciences, National University of Sciences and Technology Islamabad,
16 Pakistan

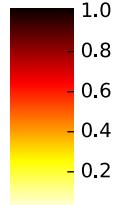
17 ⁴Department of Medicine, Washington University School of Medicine, St. Louis, MO, USA

18 ⁵Department of Molecular Microbiology, Washington University School of Medicine, St. Louis, MO, USA

19 ⁶Departments of Pediatrics, Washington University School of Medicine, St. Louis, MO, USA

20 ⁷Department of Biomedical Engineering, Washington University in St. Louis, St. Louis, MO, USA

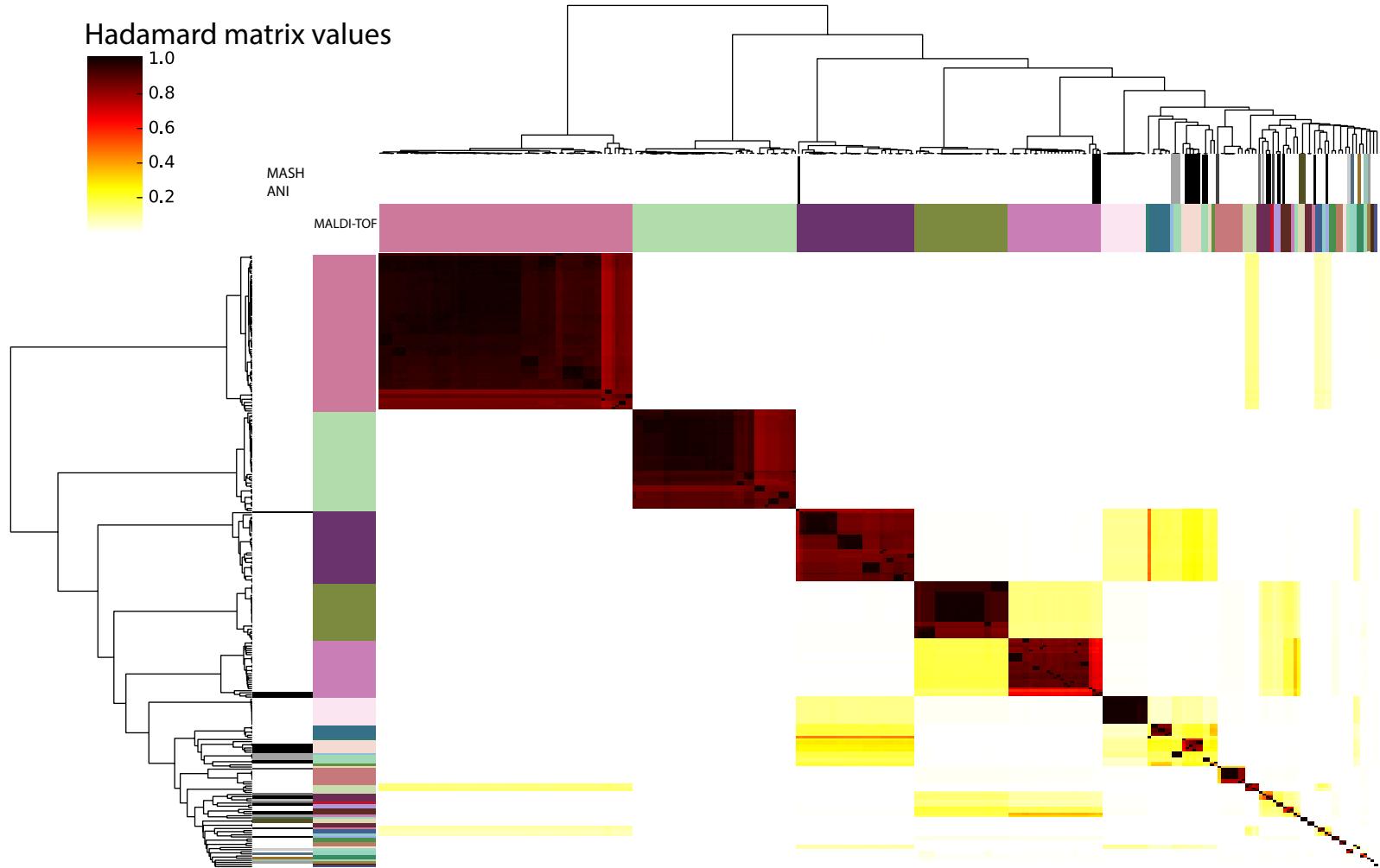
Hadamard matrix values



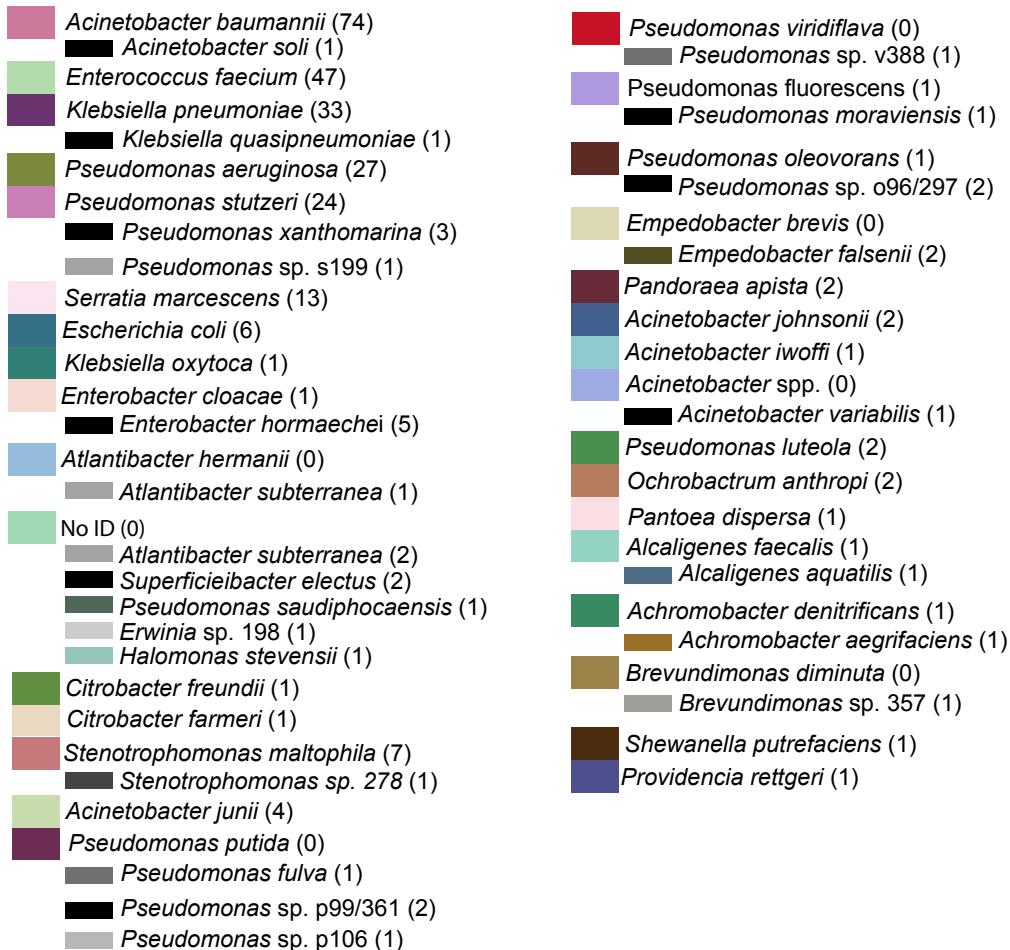
MASH

ANI

MALDI-TOF

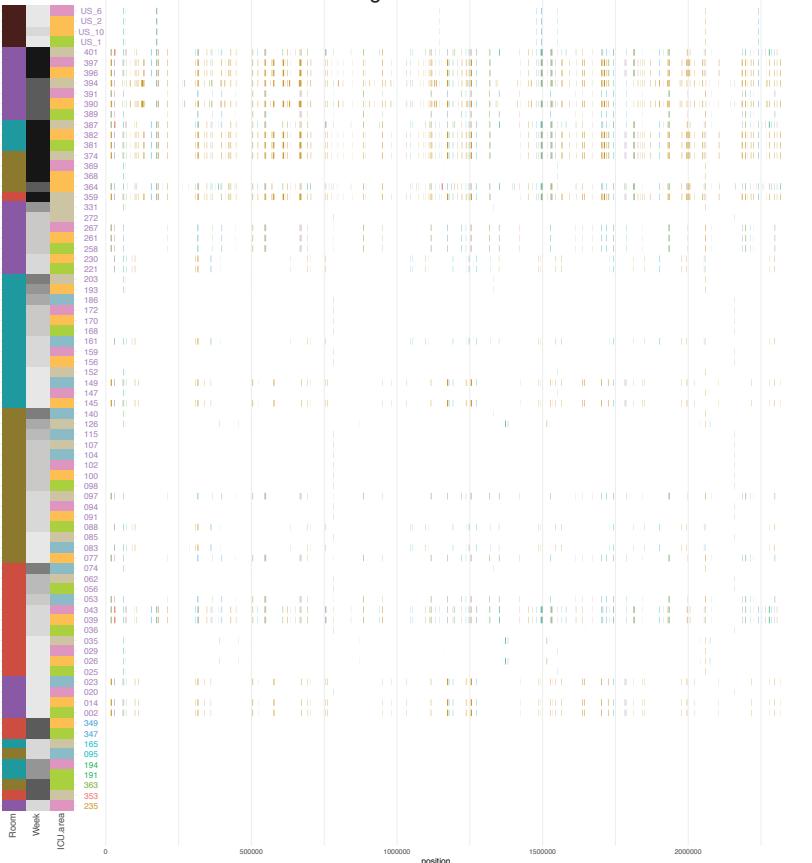
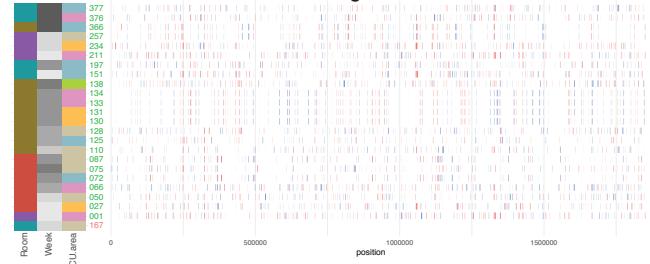
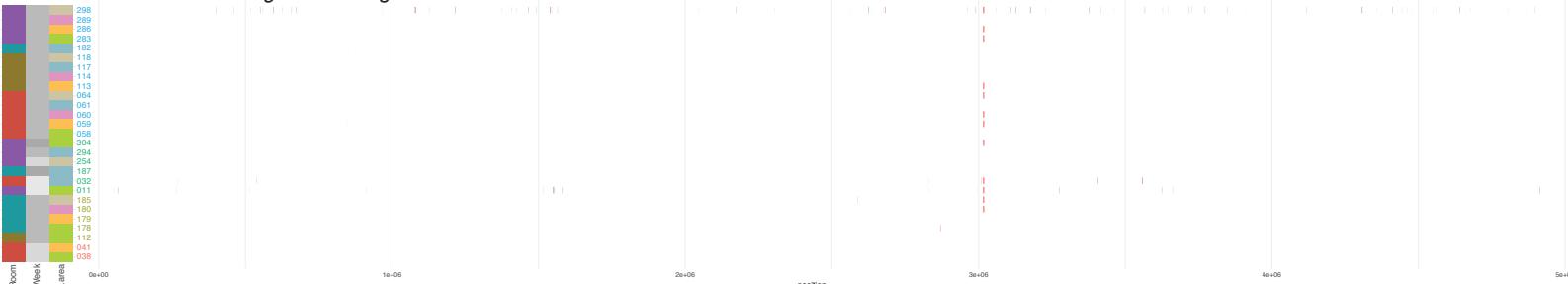


Species

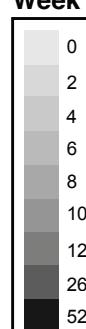


21 **Supplementary Figure 1: Taxonomic identification by WGS gives higher resolution than MALDI-TOF MS**

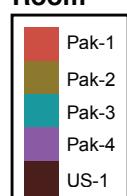
22 Pairwise average nucleotide identity (ANI) between all PAK-H isolates. Axes of this graph are hierarchically
23 clustered and the MALDI-TOF MS identifications are shown as colored annotations next to the heatmap. MASH
24 ANI species identifications to type strains are shown as colored annotations where they differed from MALDI-
25 TOF MS annotations. Source data is provided in the source data file.

a *Acinetobacter baumannii* core genome recombination**b** *Enterococcus faecium* core genome recombination**c** *Pseudomonas stutzeri* core genome recombination**d** *Pseudomonas aeruginosa* core genome recombination**e** *Klebsiella pneumoniae* core genome recombination

Week



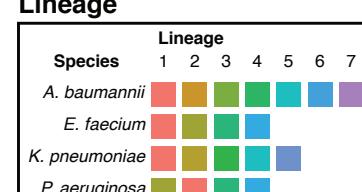
Room



ICU.area



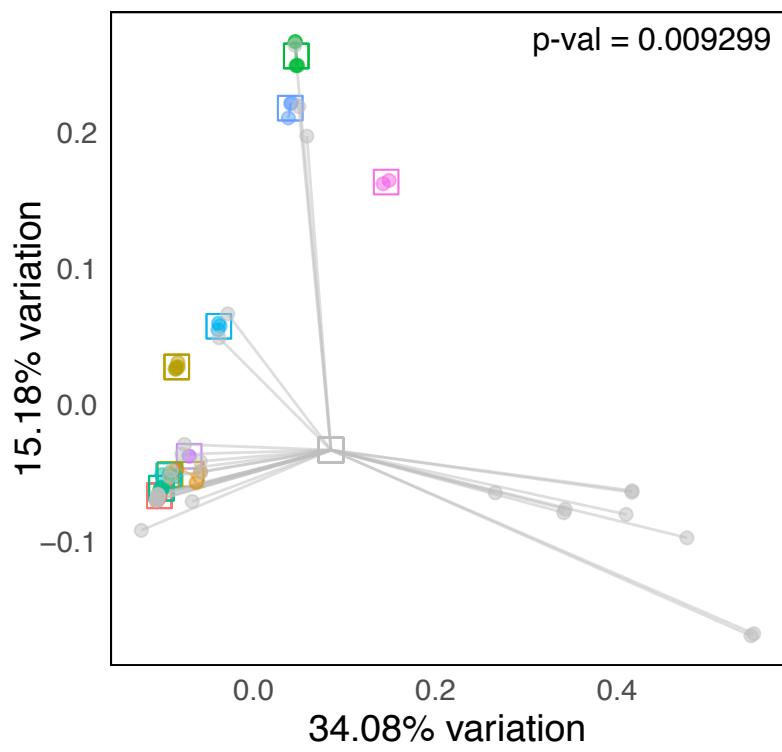
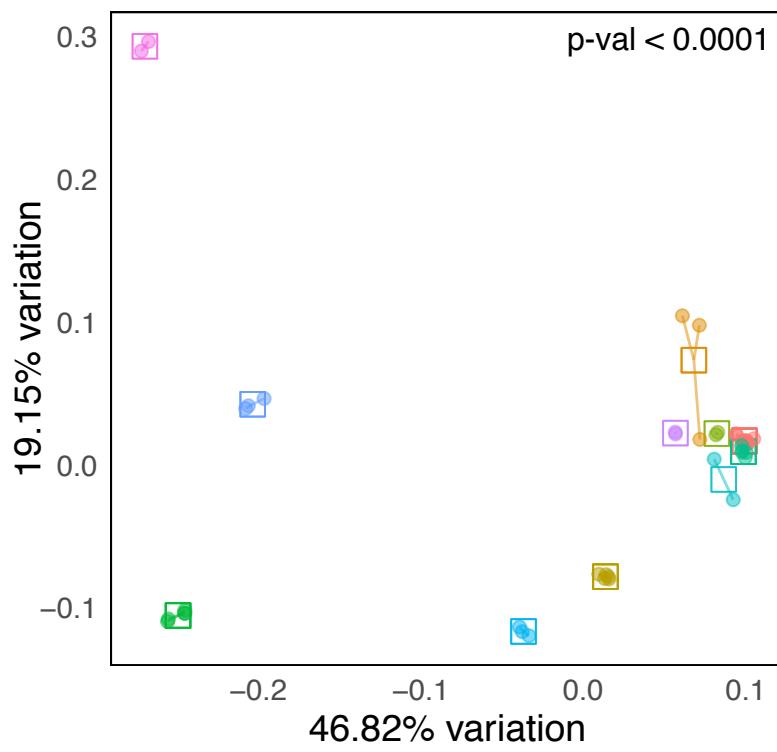
Lineage



26 **Supplementary Figure 2: Recombinant core genome positions for major species**
27 Core-genome alignments with recombinant positions for **(a)** *A. baumannii*, **(b)** *E. faecium*, **(c)** *P. stutzeri*, **(d)** *P.*
28 *aeruginosa*, **(e)** *K. pneumoniae*. The x-axis for each plot is the position in the core genome and each row is an
29 isolate. Vertical bars in the plot show where recombination is predicted, and the colors correspond to the lineage
30 of origin. Isolate numbers are colored by hierBAPS lineage and each isolate is annotated on the left with room,
31 week, and surface information. Source data for all panels is provided in the source data file.

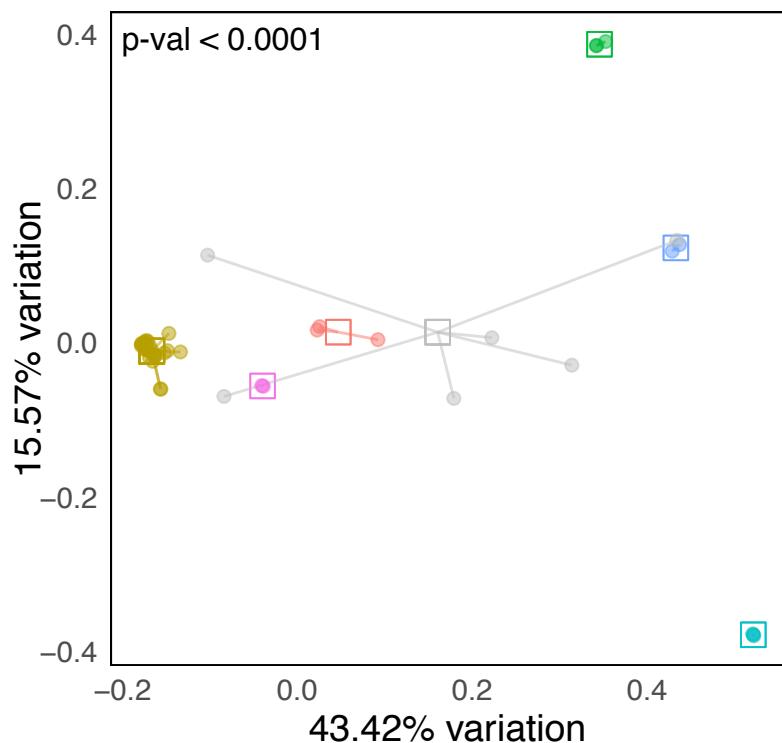
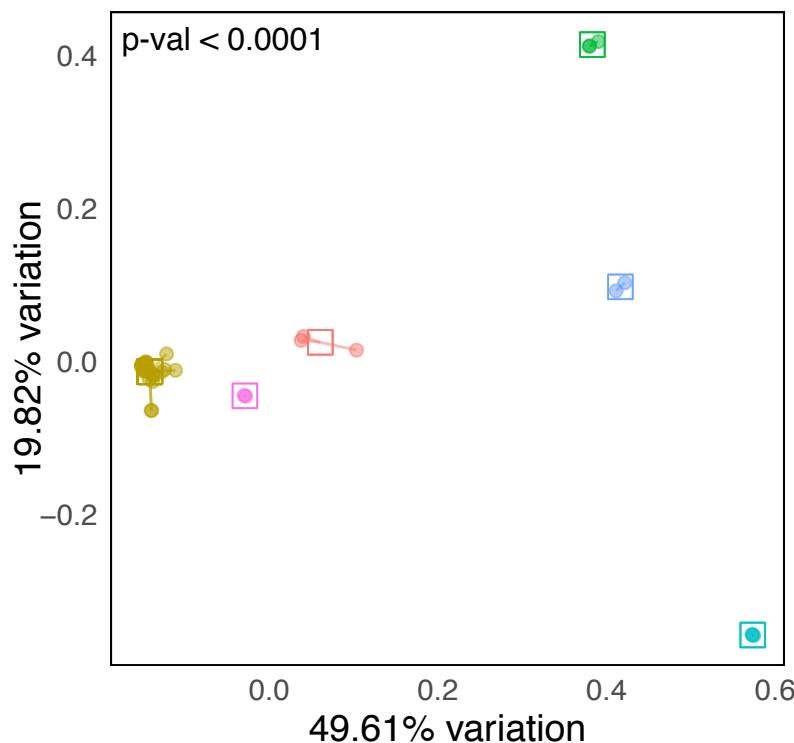


32 **Supplementary Figure 3: Histograms of pairwise core genome SNP distances for major species**
33 Distribution of pairwise SNP distances between isolates for **(a)** *A. baumannii*, **(b)** *E. faecium*, **(c)** *K. pneumoniae*,
34 **(d)** *P. aeruginosa*, **(e)** *P. stutzeri*. The x-axis is the ratio of number of pairwise SNPs to core-genome length and
35 the y-axis is counts for the histograms and density for the density plots. Source data for all panels is provided in
36 the source data file.

a *A. baumannii* accessory genes**b** *A. baumannii* accessory genes

significance from Permanova

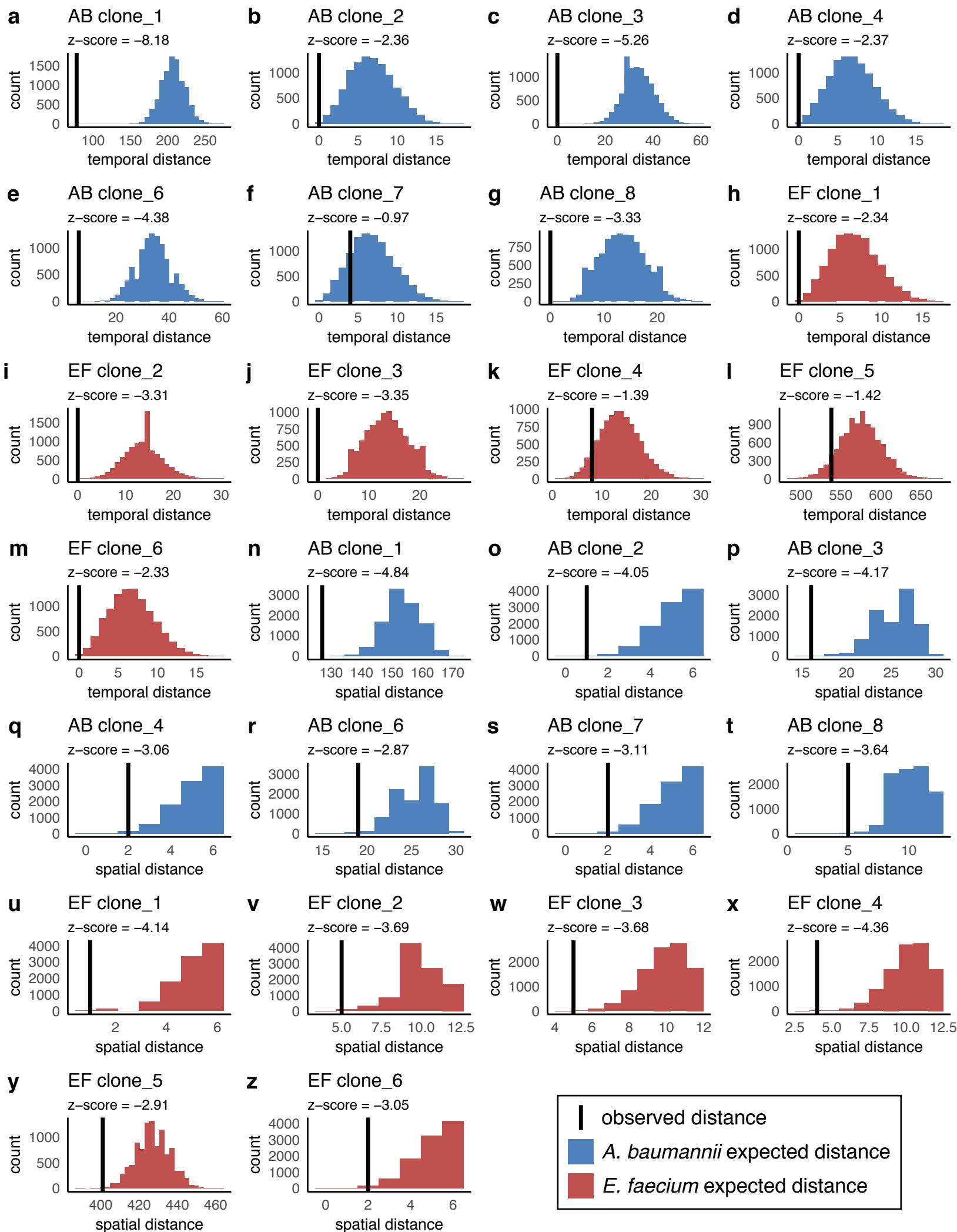
10000 permutations

c *E. faecium* accessory genes**d** *E. faecium* accessory genes

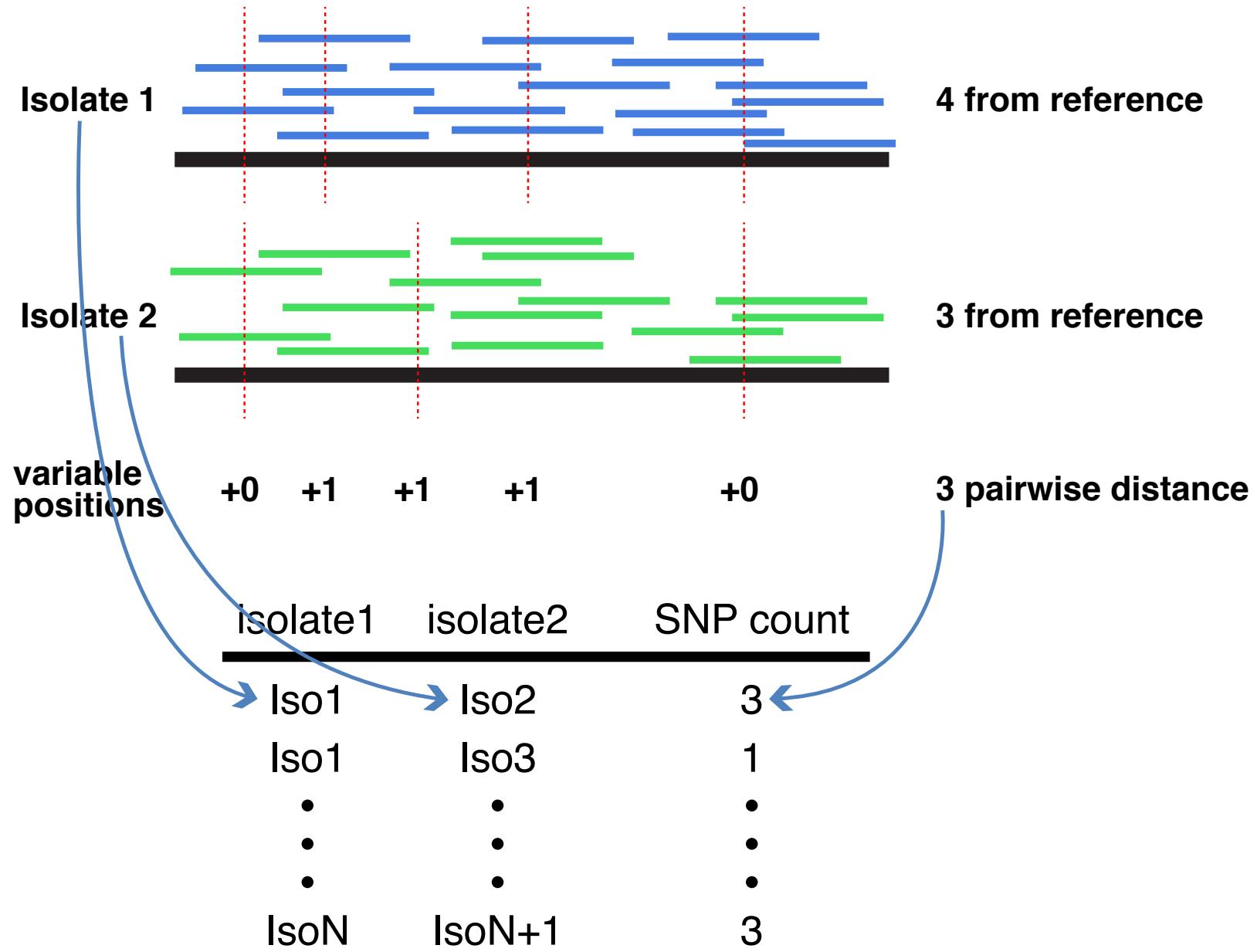
significance from Permanova

10000 permutations

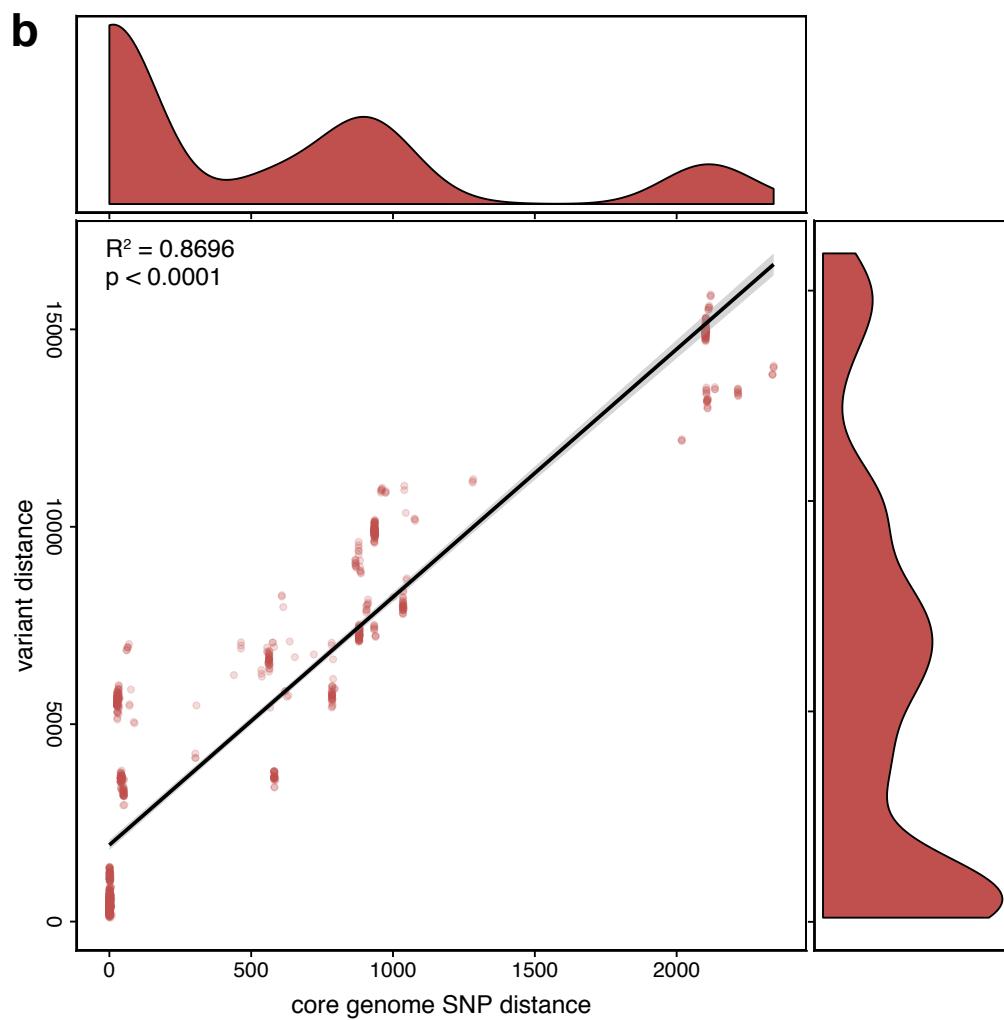
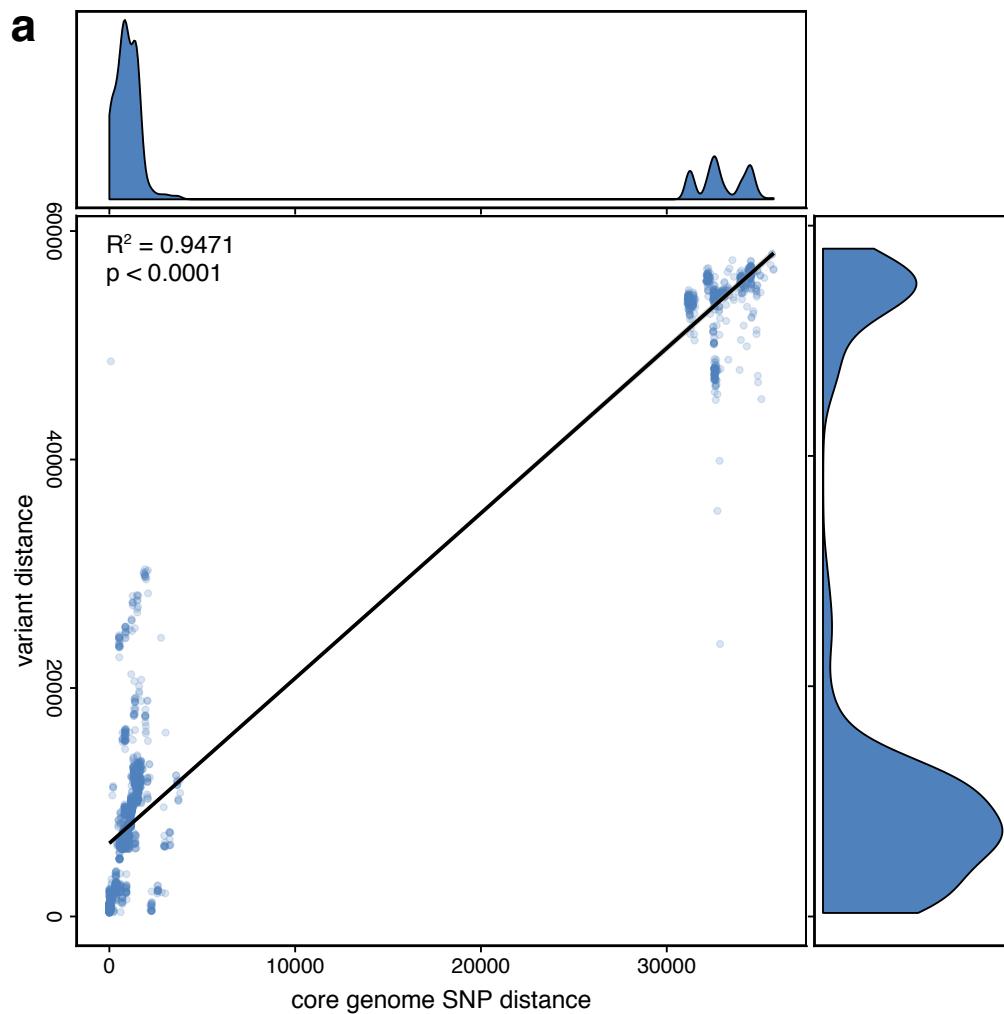
37 **Supplementary Figure 4: Accessory genome ordination of *A. baumannii* and *E. faecium***
38 Principal coordinates analysis of accessory genome for **(a)** *A. baumannii* with non-clonal isolates included, **(b)** *A.*
39 *baumannii* without non-clonal isolates, **(c)** *E. faecium* with non-clonal isolates included, **(d)** *E. faecium* without
40 non-clonal isolates. Samples for each specific isolate are filled circles and centers of gravity for each clonal group
41 are open squares. Lines connect samples to their centers of gravity. PERMANOVA was used to determine was
42 used compare clonal groups' centroids and dispersion. Source data for all panels is provided in the source data
43 file.



44 **Supplementary Figure 5: Spatial and temporal linkage of core genome SNP identified clonal groups**
45 Temporal and spatial linkage for *A. baumannii* and *E. faecium* clones during the first 3 months of collection.
46 Temporal distances are calculated as +1 for every 2 week span separating isolate collections. Spatial distances
47 are given as +0 if isolates were collected from the same surface and room, +1 if they were collected from the
48 same room, but different surfaces, and +2 if they were collected from different rooms. Expected distance
49 distributions for *A. baumannii* are shown in blue and expected distance distributions for *E. faecium* are shown in
50 red. Observed distances are shown as vertical black lines. Z-scores for the observed values compared to the
51 expected distribution are given for each graph. **(a-g)** Temporal distance distributions for *A. baumannii* clones. **(h-**
52 **m)** Temporal distance distributions for *E. faecium* clones. **(n-t)** Spatial distance distributions for *A. baumannii*
53 clones. **(u-z)** Spatial distance distributions for *E. faecium* clones. Source data for all panels is provided in the
54 source data file.



55 **Supplementary Figure 6: Identification of pairwise variant distances from reference read mapping**
56 Dark black horizontal lines represent reference genomes and colored short horizontal lines represent mapped
57 short reads. Vertical dashed red lines represent variant positions from the reference. Every position is checked
58 for differences from the reference and from all other isolates.



59 **Supplementary Figure 7: Correlation between core genome SNP and variant distances**

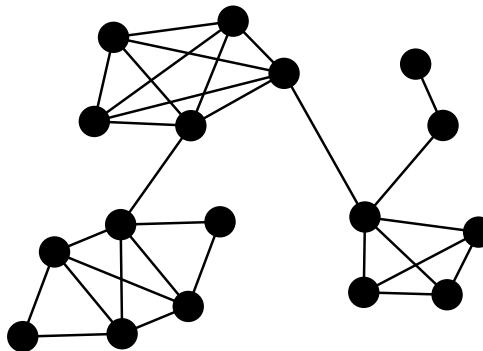
60 Comparison of pairwise core genome SNP distances to pairwise variant distances for **(a)** *A. baumannii* in blue
61 and **(b)** *E. faecium* in red. Each point shows the distance a pairwise comparison of two isolates with the x-axis
62 showing the distance by core genome SNPs and the y-axis showing the distance by variants called by mapping
63 quality filtered reads to the type strain. The density of the points for the core genome SNP and variant axis are
64 shown above and to the right of the scatterplots. The black line is the prediction given by the linear model (*variant*
65 *distance* ~ *SNP distance*) and the shaded grey area around the line is the 95% confidence interval for the
66 correlation. The adjusted R² coefficient of determination and the p-value are given on the respective plots. Source
67 data for all panels is provided in the source data file.

For each cutoff:

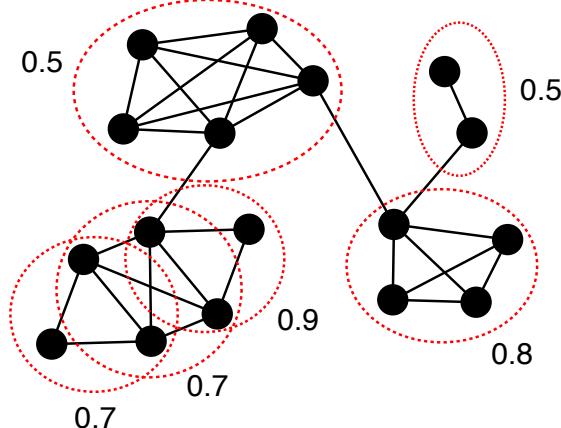
1. filter pairwise distance table by cutoff

isolate1	isolate2	variant dist
Iso1	Iso2	3
Iso1	Iso3	1
•	•	•
•	•	•
•	•	•
IsoN	IsoN+1	3

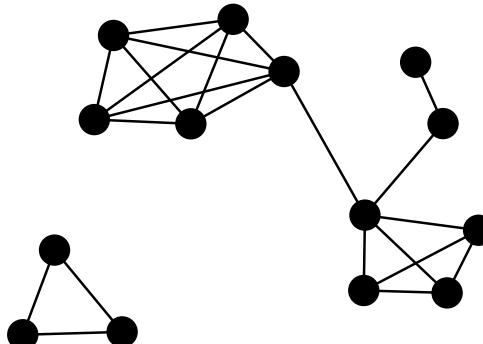
2. find all maximally connected cliques



3. sort cliques by highest minimum edgeweighting



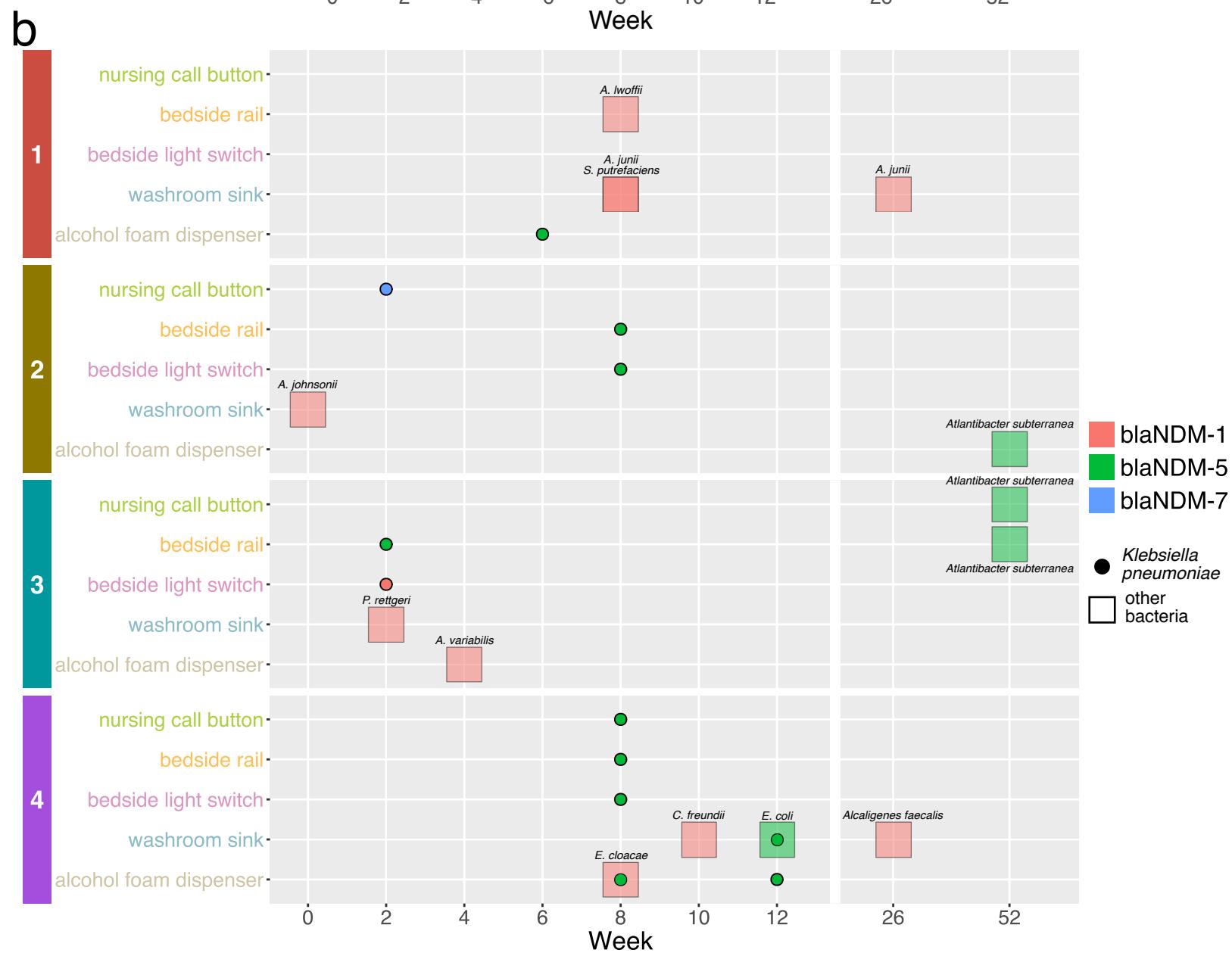
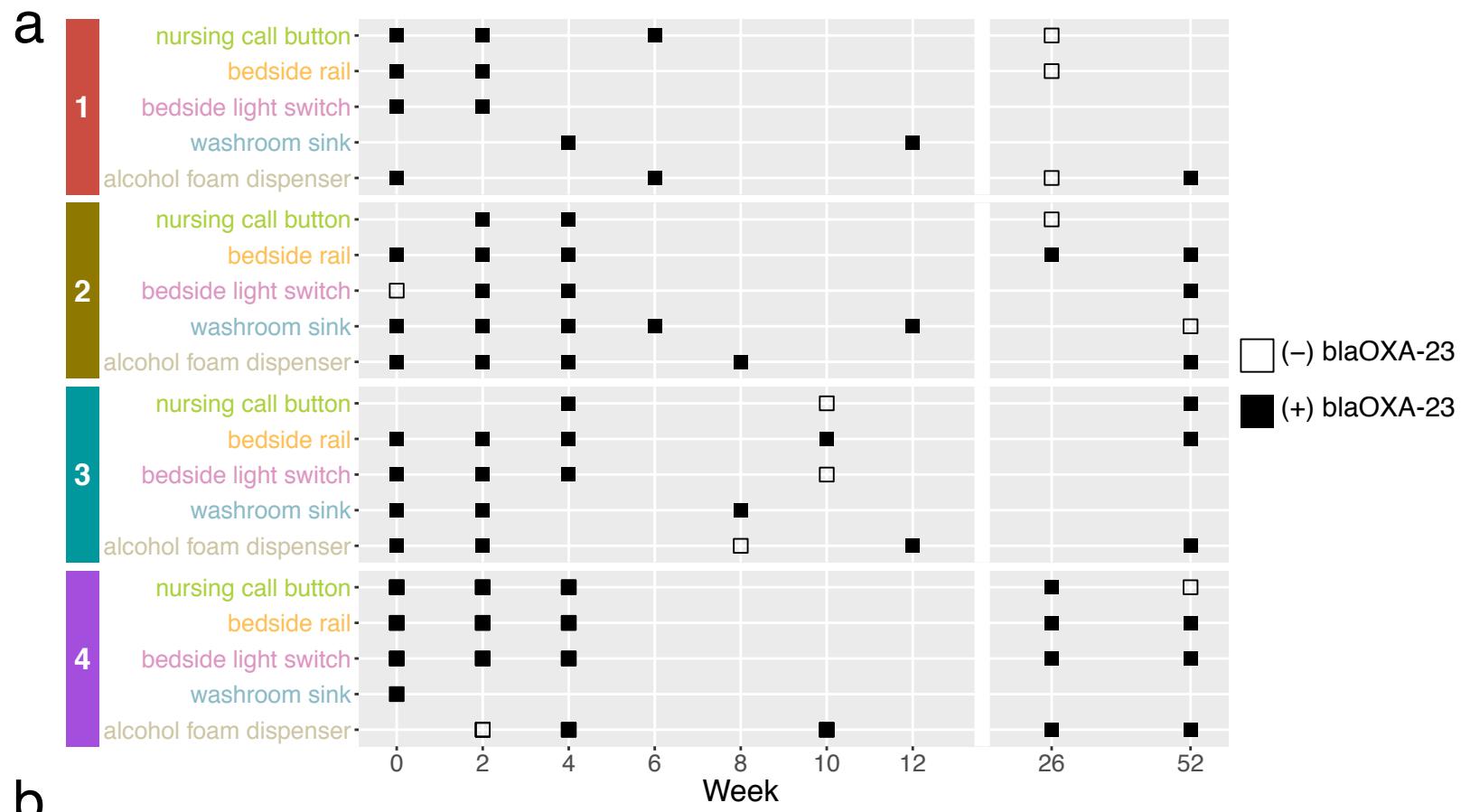
4. remake graph without strongest clique



5. recurse on steps 3 and 4

68 **Supplementary Figure 8: Identification of cliques at given cutoff value**

69 In step 1, the pairwise distance table is filtered by the cutoff value. In step 2, this filtered table is then used to
70 generate a network. In step 3, the network is searched for fully connected cliques. The weakest connection in each
71 fully connected clique is identified and cliques are sorted in descending order based on their weakest connection.
72 In step 4, the clique with the highest value weakest connection is stored and its associated nodes are removed
73 from the filtered table. The network is then regenerated without the nodes. For step 5, steps 3 and 4 are repeated
74 until all maximally connected cliques are identified.



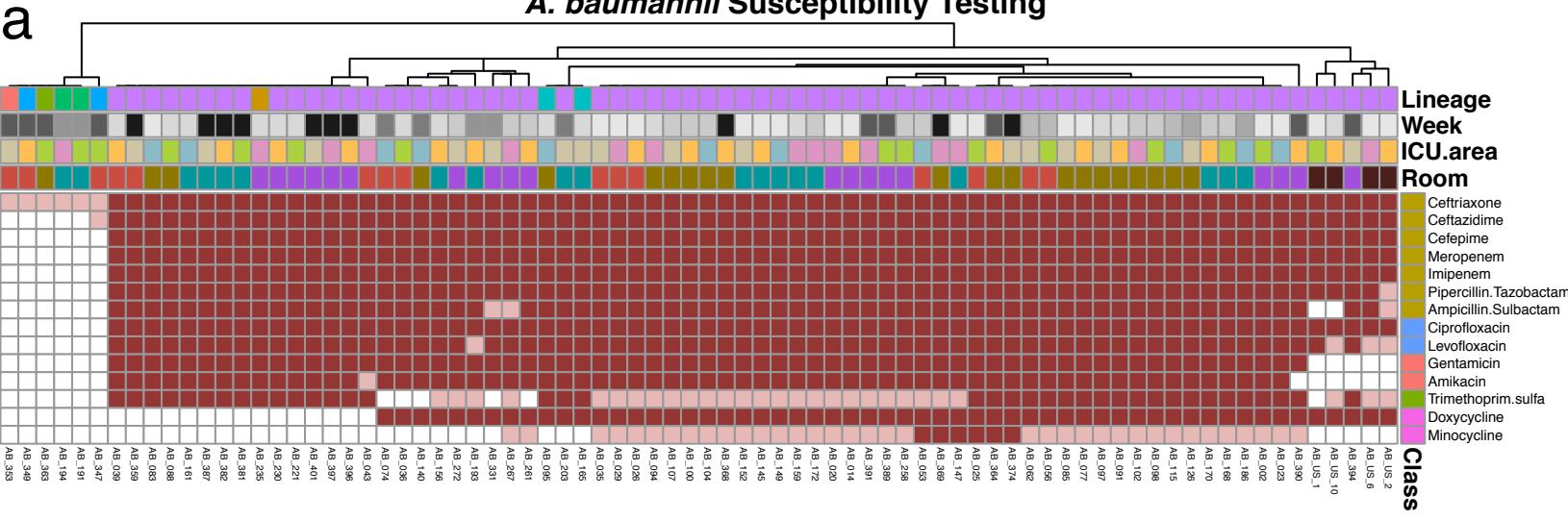
75 **Supplementary Figure 9: Spatial and temporal organization of high-risk resistance genes**

76 **(a)** Overview of *bla*_{OXA-23} on PAK-H surfaces. Each horizontal grey panel represents a PAK-H room. Within each
77 room, the horizontal gridded white lines are the five sampled surfaces. Each vertical white line is one of the
78 collection weeks. Places where the horizontal and vertical white lines intersect represent a sampling. Filled black
79 squares are *A. baumannii* isolates positive for *bla*_{OXA-23} and unfilled black squares are *A. baumannii* isolates
80 negative for *bla*_{OXA-23}.

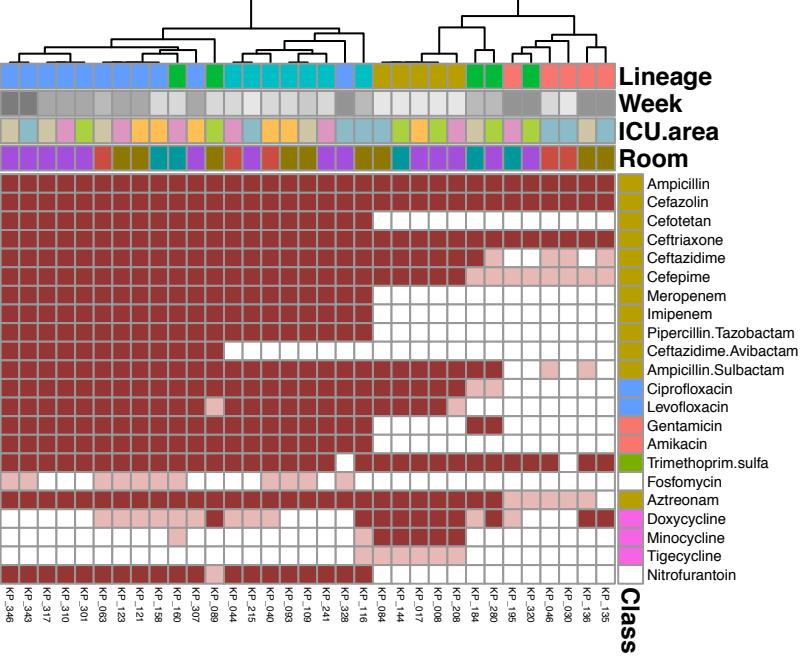
81 **(b)** Overview of *bla*_{NDM-1}, *bla*_{NDM-5}, and *bla*_{NDM-7} on PAK-H surfaces. Each horizontal grey panel represents a
82 PAK-H room. Within each room, the horizontal gridded white lines are the five sampled surfaces. Each vertical
83 white line is one of the collection weeks. Places where the horizontal and vertical white lines intersect represent
84 a sampling. Small circles are *K. pneumoniae* isolates with a *bla*_{NDM} gene. Large boxes are other bacterial isolates
85 with a *bla*_{NDM} gene. These boxes are annotated with the specific species identification. Circles and boxes are
86 colored by gene found in the isolate. *bla*_{NDM-1} is red, *bla*_{NDM-5} is green, and *bla*_{NDM-7} is blue.

87 Source data for all panels is provided in the source data file.

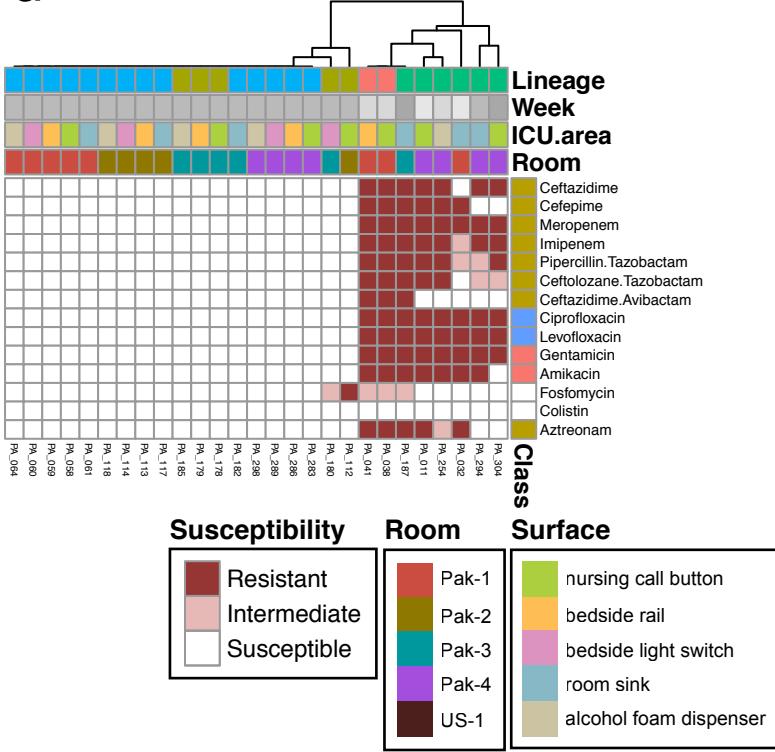
A. baumannii Susceptibility Testing



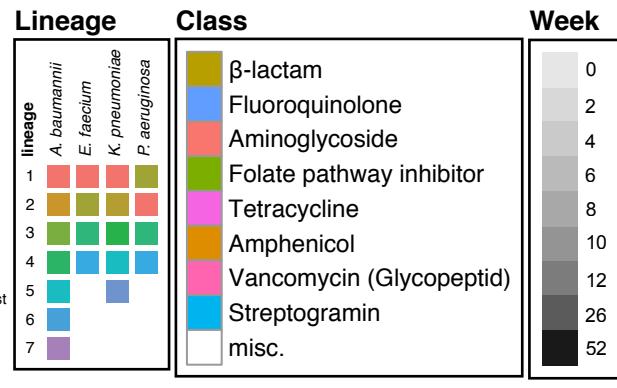
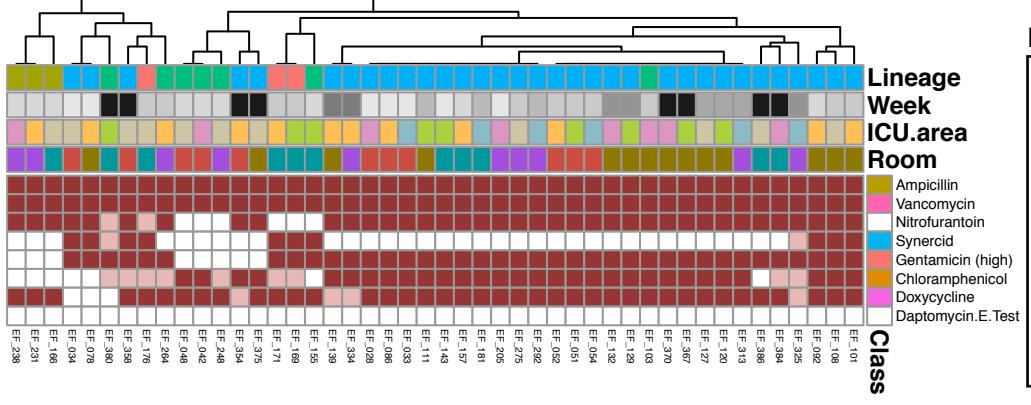
K. pneumoniae Susceptibility Testing



P. aeruginosa Susceptibility Testing



E. faecium Susceptibility Testing



88 **Supplementary Figure 10: Phenotypic susceptibility results for major species**

89 Phenotypic AST results using CLSI guidelines for **(a)** *A. baumannii*, **(b)** *K. pneumoniae*, **(c)** *E. faecium*, **(d)** *P.*
90 *aeruginosa*. Antibiotics are grouped by class on the y-axis and individual isolates are hierarchically clustered by
91 their resistance phenotypes on the x-axis. Red squares indicate phenotypic resistance in an isolate, pink
92 represents intermediate resistance, and white represents susceptibility. Colored annotations are added next to
93 antibiotics for drug class and above the charts for hierBAPS lineage, week, surface, and room. Source data for
94 all panels is provided in the source data file.