Manuscript: Markers of epidemiological success in European MRSA populations
Authors: MACOTRA consortium
https://www.jpiamr.eu/projects/macotra/

SUPPLEMENTARY METHODS 1

DETAILS OF ISOLATE SELECTION

Country-specific strategies were adopted with different inclusion criteria for storage and surveillance. Detailed information for each strategy is given below.

France. Representative isolates from France were selected from the collection of the French National Reference Centre for Staphylococcus (NRC), Lyon, France. This collection consists of methicillin-susceptible *Staphylococcus aureus* (MSSA) and MRSA isolates referred to the NRC by approximately 380 laboratories for microbiological expertise and does not include isolates from clinical studies or cohorts. From 2014 onwards, all isolates in the collection have been subjected to DNA array profiling using 332-loci Alere Staphytype (Alere Technologies GmbH, Jena, Germany) as described elsewhere (Rasigade et al., 2018). Isolates are assigned to multilocus sequence types (STs) and clonal complexes (CCs), as well as specific lineages such as ST8 USA300 (Diep et al., 2006), by comparing whole-array hybridization profiles to previously MLST-typed reference strains in a dedicated database (Monecke et al., 2008).

Isolates with a ST8 USA300 profile were readily classified as sporadic in France based on their limited local spread compared to other countries, in spite of repeated introductions (Glaser et al., 2016). Other isolates were classified as successful or sporadic and stratified across major CCs using the following rationale. Major CCs were defined as those with >20 MRSA isolates in the 2014-17 collection (which had a total of 5,457 isolates including 1,382 MRSA). Eight major CCs were found, namely CC8 (*n = 216*), CC5 (*n = 200*), CC80 (*n = 171*), CC30 (*n =39*), CC22 (*n = 36*), CC1 (*n= 33*), CC88 (*n = 27*) and CC59 (*n = 25*), totalling 747 isolates.

The successful or sporadic classification of isolates was then based on their subtype cluster frequency within each major CC. This rationale for clustering was to ensure that the intra-cluster variability within a CC was constant. First, microarray data were subjected to hierarchical clustering using Ward's method to produce one dendrogram per CC (Ward, 1963). Clusters of isolate subtypes were found in each dendrogram using equal-height tree cutting, where the number of clusters was arbitrarily defined as one-fifth of the number of isolates in the CC, up to a maximum of 10 clusters. This method ensured consistent subtyping of isolates across CCs of varying size. Subtype clusters in each CC were sorted by size. The largest clusters totalling >25%

of CC size were labelled as 'successful' while the smallest clusters totalling >25% of CC size were labelled as 'sporadic'. Other subtype clusters were considered inconclusive and excluded.

The above classification criteria resulted in 316 isolates classified as 'successful' (*CC8 = 97, CC5 = 81, CC80 = 67, CC30 = 21, CC22 = 11, CC1 = 13, CC88 = 18 and CC59 = 8*) and 152 isolates as 'sporadic' (*CC8 = 40, CC5 = 37, CC80 = 42, CC30 = 9, CC22 = 8, CC1 = 4, CC88 = 6 and CC59 = 6*). A final subset of 96 isolates were selected using balanced sampling across CCs, as well as between successful and sporadic isolates within each CC.

Netherlands. ***Type-Ned MRSA database*** - As surveillance and collection of MRSA isolates along with relevant epidemiological data is mandatory in the Netherlands, the Dutch National Institute for Public Health and the Environment (Rijksinstituut voor volksgezondheid en milieu [RIVM]) has been receiving and storing MRSA isolates collected through the national surveillance system. This system includes all Dutch Medical Microbiological Laboratories (MML) associated with general practitioners, regional and university hospitals, long term care facilities, and laboratories in Dutch territories overseas. Only one isolate per person per year is included. These include clinical isolates as well as colonisation isolates, irrespective of the reason for detection, either by contact search or increased risk factors (see below). When both colonisation and clinical isolates are available, a clinical isolate is preferred, but in practice the first isolated MRSA from a person will be included. All data is collected in the Type-Ned MRSA database. This includes MML of submission, all relevant personal data and epidemiological data, such as gender, age and sample site. Patient privacy is guaranteed under the Dutch law.

Following search and destroy (S&D), a policy implemented in the Netherlands since 1988, every patient at risk for MRSA colonisation is screened at hospital or nursing home admission and placed in pre-emptive isolation awaiting culture results. Subsequently, patients with MRSA positive culture are kept in isolation during their hospital stay and offered a treatment to eliminate colonisation, mostly after discharge. Before treatment, household members are tested on transmission and offered an elimination treatment together with the index carrier when positive. Risks for MRSA colonisation were defined by the former Dutch Working party for Infection Prevention (WIP; 1981-2017) and include, among others, contact with an MRSA carrier, recent stay in a hospital abroad and contact with farmed pigs, veal calves or broilers (Werkgroep Infectiepreventie (WIP), 2012). The assumed origin of MRSA acquisition is classified by infection control practitioners, based on the WIP risk categories, and reported in the Type-Ned database. Occasionally, MRSA is isolated from patients not targeted by S&D, for example in a clinical sample (MRSA of unknown origin; MUO (Lekkerkerk et al., 2012; Lekkerkerk et al., 2017). These findings result in contact tracing which aims to screen all exposed contacts to detect and prevent MRSA outbreaks. Sometimes, this results in identifying a MRSA isolate of

different genetic origin than the original MRSA isolate for which the contact search was initiated. These isolates are defined as unexpected findings which start new contact tracings. When no transmission of these unexpected MRSA types is found and their prevalence in the Netherlands is low, we define these MRSA types as unsuccessful (sporadic), as these did not show transmission in a hospital setting, where another MRSA type had spread.

Strain selection

The period 2008-2017 was chosen to ensure overlap in time with the selection period of British and French isolates. During 2008-2017, ±32.000 MRSA isolates were collected through national surveillance. Aside from livestock-associated (LA) MRSA clade MC0398, the following MLVA - Complexes (MCs) were most prevalent: MC0005, MC0008, MC0022, MC0045, MC0030 and MC0001 (Table 1). As the latter six MC corresponded with frequently found MLST-CC in the UK and France, a subset of isolates belonging to these MC were selected for MACOTRA. To narrow our search and account for changes in prevalence over time, we chose to select isolates from sampling years 2008 and 2017 only. We aimed to select 12 isolates for all six MCs consisting of six isolates for each sampling year per MC. During selection, isolates originating from as many different MMLs as possible were chosen. If a further choice was possible, the earliest submitted isolates were preferred. Four independent selection methods (described below and depicted in Figure 1) were used to complete the collection Dutch isolates.

	Prevalence per sampling year										
MLVA-Complex	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	Total
MC0398	41	42	40	40	38	34	30	28	25	25	34
MC0005	15	14	15	15	15	14	18	13	16	13	**15**
MC0008	15	16	14	14	13	16	14	12	12	12	**14**
MC0022	5	5	8	5	6	7	8	12	9	12	**8**
MC0045	8	6	5	8	8	8	11	10	7	6	**8**
MC0030	3	4	4	4	5	5	4	5	6	6	**5**
MC0001	1	2	2	1	2	2	3	4	5	5	**3**
Other MCs	12	12	12	12	14	12	13	16	20	20	15

Table 1. Relative prevalence (%) of included MC in the Netherlands per sampling year

Selection of successful isolates

Individual minimum spanning trees (MST) based on MLVA-types were made for MC0005, MC0008, MC0022, MC0030, MC0045 and MC0001. These MCs are representative of CC5, CC8, CC22, CC30, CC45 and CC1, respectively (expert opinion, Leo M. Schouls). Subsequently, the most prevalent MLVA types were chosen within each MC. From these MLVA types, approximately 8 isolates were selected at random and these were categorized as successful

MRSA as they have been able to persist and spread throughout the study period. MC0001 was considered least successful of the six selected prevalent MCs, hence, only 4 isolates from the most prevalent MLVA types were included as successful isolates. This selection method was named NL1. As LA-MRSA, MC0398 is the most prevalent MC in the Netherlands, a separate MST of MC0398 was used to expand the set of the above successful isolates. Three isolates from each sampling year for the most prevalent MLVA types within MC0398 were selected. This selection method was defined as NL2.

Selection of sporadic isolates

For selecting sporadic isolates, four isolates were selected from rare MLVA types within each MC specific MST used in method NL1. These isolates were categorized as sporadic MRSA. For MC0001, 8 isolates from rare MLVA types were included. The collection was expanded with another six isolates from globally dominant clones, which are not prevalent in the Netherlands. Based on expert opinion, these isolates were categorized as unsuccessful in the Netherlands, as these clones were unable to cause outbreaks in a hospital setting despite repeated introduction (selection method NL3). The added isolates included a pair of ST239 (MC0008) isolates, a pair of USA300 (defined as PVL+ and *spa* type t008) isolates and a pair of MC0080 isolates, with one isolate from 2008 and another from 2017 for each pair.
The set of sporadic isolates was further expanded with six isolates originating from unexpected findings during contact tracing of MRSA outbreaks in Erasmus MC hospital between 2008 and 2017 (selection method NL4). As described above, these MRSA had the chance to spread in a hospital setting but did not show any transmission i.e. were unsuccessful. Furthermore, the MLVA types of these last 6 isolates were present less than 5 times in the Dutch Type-Ned MRSA database between 2008 and 2017.

Additional outbreak isolates

Next-generation sequencing (NGS) of MRSA has been implemented at RIVM since 2017, enabling outbreak investigations based on whole genome MLST (wgMLST). For this approach, 2567 loci of the core and accessory genome were included, and importantly grouping based on wgMLST agreed between NGS groups and MLVA complexes (Leopold et al., 2014). The average allelic distance between NGS groups was 1673 alleles, ranging between 1169 and 1959 alleles. Genetic clusters representing possible outbreak clusters were defined as isolates within a NGS group separated by a maximum of 15 genes. In total, 20 isolates were included from five different genetic clusters (range 1- 12 alleles). From each selected genetic cluster, two isolates were defined as successful. For each genetic cluster two genetically closely related, but outside of the genetic cluster (range 43-288 alleles) were selected as sporadic counterparts.

In total, 109 isolates were included in the Dutch part of the MACOTRA strain collection.

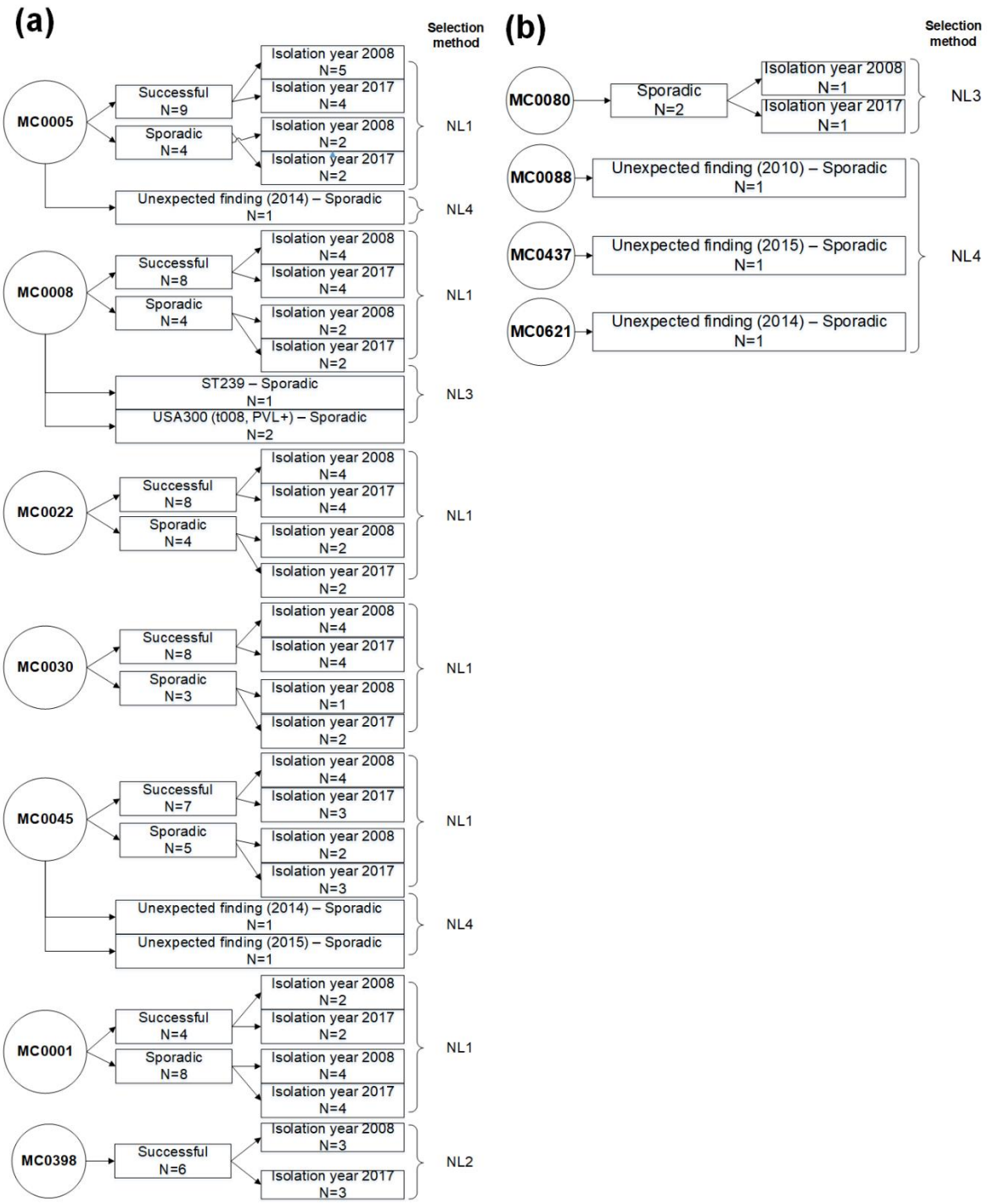

Figure 1. Selection procedure for Dutch isolates. (a) illustrates the selection procedure of successful and sporadic from the most prevalent MLVA-MC found in the Netherlands; includes LA-MRSA clade, MC0398; (b) describes the selection of sporadic isolates from less prevalent MLVA-MC identified in the Netherlands. NL1: based on selection of prevalent and rare MLVA types of 6 prevalent MLVA-MCs; NL2: selection of prevalent MLVA types of MC0398; NL3: selection based on global successful clones; NL4: selection based on unexpected findings in contact tracings. Unexpected findings implies MRSA of unknown origin, no transmission detected in contact tracing, prevalence <0.025% in MRSA Type-Ned database.

United Kingdom*:* For isolates selection, a collection of well characterised isolates from a single London hospital was utilized. These isolates were representative of the region, and were collected both before and after the reduction in incidence of MRSA infection in the UK in 2007 (Knight et al., 2012). These St George's NHS University Hospital Trust isolates were collected from 1999 – 2009 from a range of specimens sent to the diagnostic microbiology laboratory in a large acute teaching hospital servicing south west London. In 1999 and 2003, the dominant clone was CC30, interrupted by the emergence and decline of the ST239 clone, and by 2006 were dominated by CC22. Most isolates were resistant to ciprofloxacin and erythromycin. Additionally, resistance to aminoglycoside, trimethoprim, fusidic acid and tetracycline were seen. All isolates had been subjected to WGS (Kime et al., 2019) and lineages were confirmed. The collection was supplemented with all stored blood culture MRSA isolates collected at St George's between 2013-2016, where CC22 remained the dominant clone.

For the total 173 isolates, all those belonging to CC1, CC5, CC8, CC45, ST239, CC51, CC59 (n=29, 16.8% of the collection) were classified as sporadic owing to their relatively rare occurrence. For CC22 and CC30 isolates, phylogenetic trees of the collections from all three countries were constructed (Figure 2), and we defined 'successful' as those UK isolates that belonged to a cluster of two or more isolates on the tree with a SNP difference of <15 bp. The collection assigned 61 successful and 112 sporadic isolates.

REFERENCES

Diep, B. A., Gill, S. R., Chang, R. F., Phan, T. H., Chen, J. H., Davidson, M. G., Lin, F., Lin, J., Carleton, H. A., Mongodin, E. F., Sensabaugh, G. F., & Perdreau-Remington, F. (2006). Complete genome sequence of USA300, an epidemic clone of community-acquired meticillin-resistant Staphylococcus aureus. The Lancet, 367(9512), 731–739. https://doi.org/10.1016/S0140-6736(06)68231-7

Glaser, P., Martins-Simões, P., Villain, A., Barbier, M., Tristan, A., Bouchier, C., Ma, L., Bes, M., Laurent, F., Guillemot, D., Wirth, T., & Vandenesch, F. (2016). Demography and Intercontinental Spread of the USA300 Community-Acquired Methicillin-Resistant Staphylococcus aureus Lineage. MBio, 7(1), e02183-15. https://doi.org/10.1128/mBio.02183-15

Kime, L., Randall, C. P., Banda, F. I., Coll, F., Wright, J., Richardson, J., Empel, J., Parkhill, J., & O'Neill, A. J. (2019). Transient silencing of antibiotic resistance by mutation represents a significant potential source of unanticipated therapeutic failure. MBio, 10(5). https://doi.org/10.1128/mBio.01755-19

Knight, G. M., Budd, E. L., Whitney, L., Thornley, A., Al-Ghusein, H., Planche, T., & Lindsay, J. A. (2012). Shift in dominant hospital-associated methicillin-resistant Staphylococcus aureus (HA-MRSA) clones over time. Journal of Antimicrobial Chemotherapy, 67(10), 2514–2522. https://doi.org/10.1093/jac/dks245

Lekkerkerk, W. S. N., Sande-Bruinsma, N. van de, van der Sande, M. A. B., Tjon-A-Tsien, A., Groenheide, A., Haenen, A., Timen, A., van den Broek, P. J., van Wamel, W. J. B., de Neeling, A. J., Richardus, J. H., Verbrugh, H. A., & Vos, M. C. (2012). Emergence of MRSA of unknown origin in the Netherlands. Clinical Microbiology and Infection, 18(7), 656–661. https://doi.org/10.1111/j.1469-0691.2011.03662.x

Lekkerkerk, W. S. N., Haenen, A., van der Sande, M. A. B., Leenstra, T., de Greeff, S., Timen, A., Tjon-a-Tsien, A., Richardus, J. H., van de Sande-Bruinsma, N., & Vos, M. C. (2017). Newly identified risk factors for MRSA carriage in The Netherlands. PLOS ONE, 12(11), e0188502. https://doi.org/10.1371/journal.pone.0188502

Leopold, S. R., Goering, R. V., Witten, A., Harmsen, D., & Mellmann, A. (2014). Bacterial whole-genome sequencing revisited: Portable, scalable, and standardized analysis for typing and detection of virulence and antibiotic resistance genes. Journal of Clinical Microbiology, 52(7), 2365–2370. https://doi.org/10.1128/JCM.00262-14

Monecke, S., Slickers, P., & Ehricht, R. (2008). Assignment of Staphylococcus aureus isolates to clonal complexes based on microarray analysis and pattern recognition. FEMS Immunology and Medical Microbiology, 53(2), 237–251. https://doi.org/10.1111/j.1574-695X.2008.00426.x

Rasigade, J. P., Leclère, A., Alla, F., Tessier, A., Bes, M., Lechiche, C., Vernet-Garnier, V., Laouénan, C., Vandenesch, F., Leport, C., Hoen, B., Duval, X., Alla, F., Bouvet, A., Briancxon, S., Cambau, E., Celard, M., Chirouze, C., Danchin, N., … Zinzius, P. Y. (2018). Staphylococcus aureus CC30 lineage and absence of sed, j,r-harboring plasmid predict embolism in infective endocarditis. Frontiers in Cellular and Infection Microbiology, 8(JUN). https://doi.org/10.3389/fcimb.2018.00187

Ward, J. H. (1963). Hierarchical Grouping to Optimize an Objective Function. Journal of the American Statistical Association, 58(301), 236–244. https://doi.org/10.1080/01621459.1963.10500845

Werkgroep Infectiepreventie (WIP). (2012). Meticilline-resistente Staphylococcus aureus (MRSA). https://www.rivm.nl/documenten/wip-richtlijn-mrsa-ziekenhuizen

Figure 2a. CC22 Phylogenetic tree

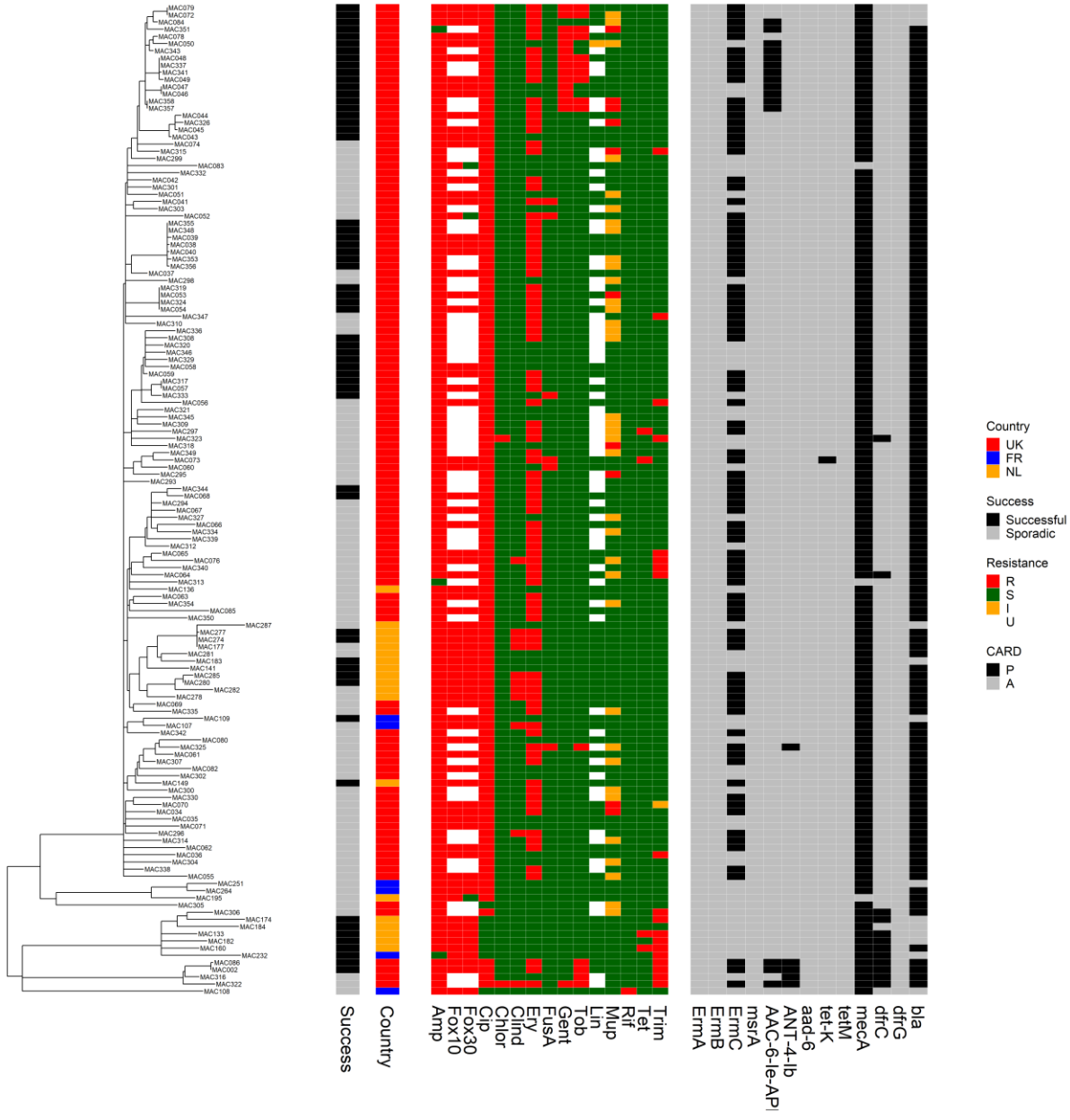

Figure 2b. CC30 phylogenetic tree