New *Mycobacterium tuberculosis* Complex Sublineage, Brazzaville, Congo

Sven Malm, Laure S. Ghoma Linguissi, Emmanuel M. Tekwu, Jeannhey C. Vouvoungui, Thomas A. Kohl, Patrick Beckert, Anissa Sidibe, Sabine Rüsch-Gerdes, Igor K. Madzou-Laboum, Sylvie Kwedi, Véronique Penlap Beng, Matthias Frank, Francine Ntoumi, Stefan Niemann

Tuberculosis is a leading cause of illness and death in Congo. No data are available about the population structure and transmission dynamics of the *Mycobacterium tuberculosis* complex strains prevalent in this central Africa country. On the basis of single-nucleotide polymorphisms detected by whole-genome sequencing, we phylogenetically characterized 74 MTBC isolates from Brazzaville, the capital of Congo. The diversity of the study population was high; most strains belonged to the Euro-American lineage, which split into Latin American Mediterranean, Uganda I, Uganda II, Haarlem, X type, and a new dominant sublineage named Congo type (n = 26). Thirty strains were grouped in 5 clusters (each within 12 single-nucleotide polymorphisms), from which 23 belonged to the Congo type. High cluster rates and low genomic diversity indicate recent emergence and transmission of the Congo type, a new Euro-American sublineage of MTBC.

Despite the availability of antituberculous drugs for the past 60 years, tuberculosis (TB) remains a major health threat worldwide. In 2014, the World Health Organization registered 6 million new TB cases, and 1.5 million affected persons died of infection with *Mycobacterium tuberculosis* complex (MTBC) strains, the causative agent of TB (*1*). Congo (Republic of Congo), in Central Africa, has a population of ≈4 million inhabitants and is considered to be a high TB incidence area; incidence is 381 cases/100,000 inhabitants (*1*). Approximately one fourth of the population agglomerates in its capital city, Brazzaville.

Despite the serious situation, precise data on disease dynamics and recent transmission patterns guided by modern molecular epidemiologic tools are only sparsely available. Molecular epidemiology is useful for analyzing MTBC strain diversity and transmission dynamics in low- and high-incidence settings (*2,3*). Furthermore, molecular typing has shown that MTBC has a diverse population structure with manifold lineages that show large differences in geography and pathobiological properties, such as the development and spread of drug resistance (*4,5*).

To address current knowledge gaps, we determined the population structure of MTBC isolates from patients with pulmonary TB in Brazzaville. Samples were collected from patients at the Centre Antituberculeux de Brazzaville during February–June 2011 (*6*). We investigated the population structure and transmission patterns by a combination of classical genotyping and whole-genome sequencing (WGS). Single-nucleotide polymorphisms (SNPs) detected by WGS were used for phylogenetic lineage classification and similarity analysis estimating recent transmission rates. This approach enabled detailed insight into the population structure and phylogeny of MTBC strains circulating in Brazzaville. Moreover, we describe a new predominant sublineage, the Congo type, which most likely forms a larger transmission network in the study area.

Methods

Study Design

The patient population was reported previously (*6*). In brief, 775 consecutive patients seeking care at the Centre Antituberculeux de Brazzaville during February–June 2011 were evaluated for pulmonary TB according to the national diagnostic algorithm (*6*). The institutional ethics committee, Comité d'Ethique pour la Recherche Scientifique, in August 2010 (no. 00000067/DGRST/CERSSA) approved the study. Informed consent for study participation, as well as permission to use isolates from samples provided, were obtained from all enrolled participants.

Author affiliations: Leibniz Center for Medicine and Biosciences, Borstel, Germany (S. Malm, T.A. Kohl, P. Beckert, S. Rüsch-Gerdes, S. Niemann); Fondation Congolaise pour la Recherche Médicale, Brazzaville, Congo (L.S. Ghoma Linguissi, J.C. Vouvoungui, A. Sidibe, S. Kwedi, F. Ntoumi); University Marien Ngouabi, Brazzaville (L.S. Ghoma Linguissi, J.C. Vouvoungui, A. Sidibe, F. Ntoumi); University of Yaoundé I, Yaoundé, Cameroon (E.M. Tekwu, V. Penlap Beng); Institute for Tropical Medicine, University of Tübingen, Tübingen, Germany (E.M. Tekwu, M. Frank, F. Ntoumi); Centre Antituberculeux de Brazzaville, Brazzaville (I.K. Madzou-Laboum); German Center for Infection Research, Tübingen Site, Tübingen (M. Frank); German Center for Infection Research, Borstel Site, Borstel (S. Niemann)

DOI: http://dx.doi.org/10.3201/eid2303.160679

Samples with the highest Ziehl-Neelsen score (semiquantitatively classified in categories 1+, 2+, or 3+ in on-site laboratories based on microscopic findings) out of at least 2 positive sputum samples from 1 patient (n = 211) were shipped to the Research Center Borstel (Borstel, Germany) for culture, drug susceptibility testing, genotyping, and WGS.

Sample Processing, Culture, and Drug Susceptibility Testing

Approximately 5 mL of each specimen was homogenized by digestion for 1 min at room temperature with 1 mL of N-acetyl L-cysteine (25 mg/mL) in phosphate buffer (pH 6.8) and vortexed with several 4-mm glass beads for 30 s. A 5-mL aliquot was decontaminated by using 1% NaOH and concentrated at 4,000 × *g* for 15 min. The sediment was then reconstituted to 2.5 mL by using phosphate buffer pH 6.8 to make the inoculum for smears and cultures. Sputum was cultured by using the conventional Löwenstein-Jensen growth medium followed by determination of mycobacterial species according to standard techniques (*7*). Samples for susceptibility testing of first-line drugs were processed as described previously (*8*). Drug susceptibility testing was performed by using the BACTEC MGIT system (Becton, Dickinson and Company, Franklin Lakes, NJ, USA). Samples without antimicrobial drugs served as growth controls. Genomic DNA was extracted from sputum cultures on Löwenstein-Jensen medium, by using a standard cetyltrimethylammonium bromide–NaCl method (*3*).

Traditional Genotyping

We performed spacer oligonucleotide typing (spoligotyping) as described by Kamerbeek et al. (*9*). PCR-negative controls were included in which distilled, diethyl pyrocarbonate-treated H_2O was added instead of DNA. Genomic DNA of *M. tuberculosis* H37Rv and *M. bovis* BCG were included as controls. In addition to spoligotyping (*10*), we conducted mycobacterial interspersed repetitive unit–variable number tandem repeat (MIRU-VNTR) typing based on 24 loci as described previously (*11*); both 24-loci MIRU-VNTR typing and spoligotyping data analysis was performed by using the tools implemented at the MIRU-VNTRplus website (*12*). Genomic DNA of *M. tuberculosis* H37Rv was included as a positive control. We used the MIRU-VNTRplus nomenclature server (*11*) in addition to the definition of shared spoligotypes to assign a unique MTBC 15-9 number to each 24-loci MIRU-VNTR combination. Data were analyzed further only if all controls showed the expected outcome. All traditional genotyping techniques were performed at the Research Center Borstel, Leibniz Center for Medicine and Biosciences (Borstel, Germany).

WGS and Data Analysis

We prepared libraries for next-generation sequencing from genomic DNA by using the Nextera XT library preparation kit and run with Illumina-supplied reagent kits on the HiSeq and MiSeq systems (Illumina, San Diego, CA, USA), according to the manufacturer's recommendations. For 1 isolate (9679-00), genomic DNA was sequenced by GATC Biotech AG (Konstanz, Germany). WGS of the strains of the study population was conducted at the Research Center Borstel, Leibniz Center for Medicine and Biosciences. NGS data of sequenced isolates was submitted to the EMBL-EBI ENA sequence read archive (PRJEB9545).

We mapped sequence reads to the *M. tuberculosis* H37Rv genome (GenBank accession no. NC_000962.3) with the SARUMAN exact alignment tool (*13*). Genomic coverage was at least 50-fold for all isolates. Customized Perl scripts were used to extract SNPs from mapped reads, requiring a minimum coverage of 10 reads and a minimum allele frequency of 75% as detection thresholds (*14*). We excluded SNPs in resistance-mediating genes and repetitive regions from the phylogenetic analysis (*15*). Moreover, to avoid calling SNPs because of indel-related artifacts, we excluded SNPs within ± 12 nt from each other (*16*). Positions that matched these thresholds in all isolates were considered as valid and used for a concatenated sequence alignment.

We then calculated a pairwise distance matrix from concatenated SNP positions by Perl scripts, with +1 distance between paired isolates for each mismatching base, and plotted data in GraphPad Prism 5 (GraphPad Software Inc., La Jolla, CA, USA). On the basis of the distance matrix, we grouped isolates into putative transmission networks by incrementally accumulating all isolates with a maximum distance of 5 or 12 SNPs, respectively, to the nearest neighbor into 1 group (*16*). On the basis of the WGS data, we classified isolates into known phylogenetic groups according to the set of informative SNP positions published by Coll et al. (*17*).

We calculated the phylogenetic tree using the maximum-likelihood method and the general time reversible (GTR) substitution model, rate heterogeneity, without invariant sites using a gamma distribution as well as bootstrap resampling. Substitution models were tested and trees calculated by using MetaPiga software version 3.1 (*18*) and the maximum-likelihood ratio test (*19*). We applied midpoint rooting with FigTree software version 1.4.2 (http://tree.bio.ed.ac.uk/software/figtree/) and performed formatting by using the online tool Evolview (*20*). Specific SNPs for the Congo-type sublineage, and the more distant undefined strain 8095/11, were extracted by the ancestral states reconstruction method, implemented in the MetaPiga version 3.1 software, for both the specific and the common node. The maximum parsimony tree for visualizing genome-based clusters was calculated with Bionumerics 7.5 software (Applied Maths, Kortrijk, Belgium).

We used the concatenated sequence alignment in a Bayesian coalescent analysis with BEAST version 1.8.2 to infer node ages in the genealogic tree (*21*). A tip dating approach was not possible; therefore, we used a strict molecular clock prior of 1×10^{-7} substitutions per site per year and compared different demographic models and a GTR versus Hasegawa, Kishino and Yano substitution model using a chain length of 10 million and sampling of every 1,000 generations with a burn-in of 10% that resulted in adequate mixing of the Markov chains and effective sample sizes in the thousands. The comparison of the likelihoods of each run with Tracer version 1.5 showed very strong support of the GTR substitution model over Hasegawa, Kishino and Yano ($\log_{10}$ Bayes factors >93) and no preference for a particular demographic model; thus, we used a coalescent constant size model, representing the most straightforward approach. Resulting data were combined in a maximum clade credibility tree by using TreeAnnotator version 1.8.2 to infer node ages and highest posterior density intervals (*21*).

Results

Study Population

Results of sputum smear microscopy, radiographic abnormalities, and HIV infection were reported previously (*6*). In brief, 211 sputum samples of patients with suspected TB were sent to the National Reference Center for Mycobacteria (Borstel, Germany); 75 cultures yielded positive results. Phenotypic characterization identified 1 *M. intracellulare*, 6 *M. africanum*, and 68 *M. tuberculosis* isolates. We excluded the *M. intracellulare* isolate from further analysis. The mean age ± SD of the 74 TB patients was 33.86 ± 11.65; 69 (93%) patients were <50 years of age, and 66% were male. The patients' residences were distributed in the different districts of Brazzaville as follows: southern part, 31% from Makélékélé and 4% from Bacongo; and northern part, 22% from Talangai, 8% from Poto-poto, and 7% from Ouenzé (4%). Of the 74 study participants, 13 (18%) were HIV co-infected.

Drug Susceptibility Patterns

We determined phenotypic drug susceptibility patterns for the first-line anti-TB drugs isoniazid (INH), rifampin (RIF), ethambutol (EMB), and pyrazinamide (PZA) for all strains. In case resistance against 1 of these drugs was detected, streptomycin (STR) and second-line antimicrobial drugs were included in the analysis. Of the 74 MTBC strains, 71 (96%) were fully sensitive to all the first-line anti-TB drugs. Three (4%) isolates were resistant: 1 isolate was resistant against INH and STR, and 2 isolates exhibited a multidrug-resistant (MDR) phenotype with resistances against INH, RIF, PZA, and EMB (Table). The 2 MDR strains underwent susceptibility testing on second-line anti-TB drugs; 1 isolate was resistant to ethionamide. No extensively drug resistant strains were identified.

Population Structure of the MTBC Isolates

For all isolates, we successfully performed classical genotyping and WGS. Overall, we detected 18,059 SNP positions, which we used for further interrogations. On the basis of these analyses, we classified the *M. tuberculosis* strains into the main phylogenetic lineages Euro-American (n = 64); Delhi/Central Asian (n = 2; Coll lineage 3 and sublineage 3.1.1); Beijing (n = 1; Coll sublineage 2.1.1); East African Indian (n = 1; Coll sublineage 1.2.2); and *M. africanum* West African-1 (n = 6; Coll lineage 5). The Euro-American strains split into Latin American Mediterranean (LAM; n = 12; Coll sublineages 4.3.2, 4.3.4.1, 4.3.4.2, 4.3.4.2.1 and 4.3.3); Uganda I (n = 7; Coll sublineage 4.6.1.2); Uganda II (n = 1; Coll sublineage 4.6.1.1); Haarlem (n = 4; Coll sublineage 4.1.2.1); and X type (n = 3; Coll sublineage 4.1.1.1). Eleven Euro-American strains not well classified by MIRU-VNTR/spoligotyping (mainly T) belong to the Coll sublineages 4.1.2, 4.2.2, 4.6, 4.7, 4.8, and 4.9. Although few 4.7 strains were present in the collection analyzed by Coll et al. (*17*), they are dominant in our study collection and thus were termed Congo type (n = 26). These strains form a clear-cut branch in the MIRU/spoligotyping (data not shown) and SNP-based phylogeny (Figures 1, 2). Thus, these strains most likely represent a new Euro-American sublineage circulating in the region.

For an in-depth view on the population structure and to define the position of the Congo-type strains in the MTBC phylogeny, we analyzed the SNPs detected by WGS together with a set of reference strains (65 strains) previously used in the MIRU-VNTRplus dataset (*11*). Our reference collection comprises 3 clinical isolates of the major lineages of the MTBC and the type strains *M. tuberculosis* H37Rv ATCC 27294, *M. bovis* ATCC 19210, and *M. africanum* West African 2 ATCC 25420 (*22*). On the basis of the 18,059 SNP positions, we calculated a maximum-

Table. Description of lineage and associated *rpoB*, *katG*, and *fabG1-InhA* mutation identifiers in drug-resistant *Mycobacterium tuberculosis* isolates, Congo*

Isolate code	Resistance	Gene	Nucleotide change	Amino acid substitutions	Lineage
8032/11	INH, STR	*katG*	AGC-315-ACC	Ser-315-Thr	Congo type
8114/11	INH, RIF, STR, EMB, PZA, ETH	*katG*	AGC-315-ACC	Ser-315-Thr	Uganda I
		rpoB	GAC-516-GTC	Asp-516-Val	
		inhA	−102G/A	NA	
8125/11	INH, RIF, STR, EMB, PZA	*katG*	AGC-315-ACC	Ser-315-Thr	Beijing
		rpoB	GAC-516-TAC	Asp −516- Tyr	

*EMB, ethambutol; ETH, ethionamide; INH, isoniazid; NA, not applicable; PZA, pyrazinamide; RIF, rifampin; STR, streptomycin.

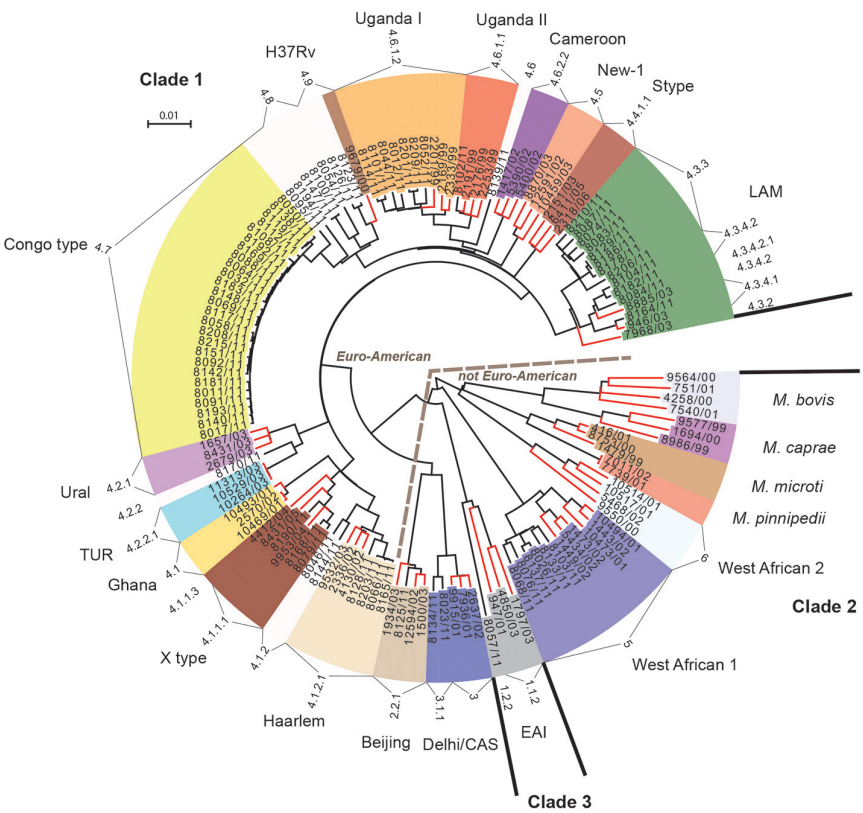

Figure 1. Maximum-likelihood tree of *Mycobacterium tuberculosis* complex isolates from Congo (black branch tips) and 65 reference strains (red branch tips). The tree was calculated by using the general time reversible substitution model with gamma distribution based on single-nucleotide polymorphisms identified by whole-genome sequencing. Models were tested and the tree generated by using MetaPiga software version 3.1 (*18*) and the maximum-likelihood ratio test. Midpoint rooting was performed. Distinct colors were chosen for the lineages identified; leaves with white background represent strains that initially were not assigned to particular lineages because of ambiguous typing patterns from mycobacterial interspersed repetitive unit, restriction fragment length polymorphism, or spoligo analysis (data not shown). The numerical code assigned to the respective lineages at the outer rim of the circular tree shows the Coll-nomenclature inferred from the whole-genome sequencing data. EAI, East African Indian; LAM, Latin American Mediterranean; TUR, Turkish. Scale bar indicates nucleotide substitutions per site.

likelihood tree (Figure 1) that shows that the Congo-type strains cluster most closely with H37Rv, presumably being part of a larger sublineage of H37Rv-related strains mentioned in other settings (*23,24*). Seven additional strains belonging to the Coll lineages 4.8 and 4.9 form a specific branch together with Congo type and H37Rv. We generated a tree containing bootstrap values supporting the branches of the tree (online Technical Appendix Figure 1, https://wwwnc.cdc.gov/EID/article/23/3/16-0679-Techapp1.pdf).

Bayesian coalescent analysis approximated the last common ancestor of all Congo-type strains in our study to 1958 (95% highest posterior density 1947–1967). Thus, the Congo-type strains of this study probably emerged in the past 53 years.

WGS enabled us to identify SNPs specific for the Congo type by extracting the SNPs from the node specific for the Congo-type sublineage only and the common node for the Congo type and the more distant strain 8095/11. We found 49 SNPs unambiguously associated with the Congo-type sublineage (online Technical Appendix Table). The analysis of the pairwise distances revealed a homogenous population of the Congo-type strains with a median pairwise distance of 20 SNPs only (range 2–53 SNPs), whereas this distance was larger among the strains of other lineages (online Technical Appendix Figure 2).

The MDR isolates belonged to the Uganda and the Beijing lineages. By contrast, the strain resistant against INH and STR could be assigned to the Congo-type sublineage. We found no significant association between the Congo type and basic variables, such as sex, age, date of specimen collection, or patients' residence or HIV status.

Cluster Analysis

Among the strains in the study population, we identified 4 groups of strains with pairwise distances within 0–5 SNPs. Such close relationships between the strains indicate recent transmission. Two of the genome clusters are formed by Congo-type strains (2 clusters formed by 2 strains, 15% of the Congo-type strains). The other 2 genome clusters consisted of 2 isolates each, of either the LAM or Coll sublineage 4.8 closely related to H37Rv.

When we used a wider cluster definition of 12 SNPs maximum distance, 30 strains were grouped in 5 clusters ranging in size from 2 to 20 strains (Figure 2). Overall, 23 of the 26 Congo-type strains (88.5%) are in 1 large (n = 20) and 1 smaller (n = 3) cluster. Three (75%) isolates of the Haarlem lineage and 2 (17%) isolates of the LAM lineage are grouped in such clusters.

Discussion

Similar to other countries in Central Africa, Congo has a high incidence of TB. Our aim was to determine the population structure and transmission dynamics of the MTBC strains in Brazzaville.

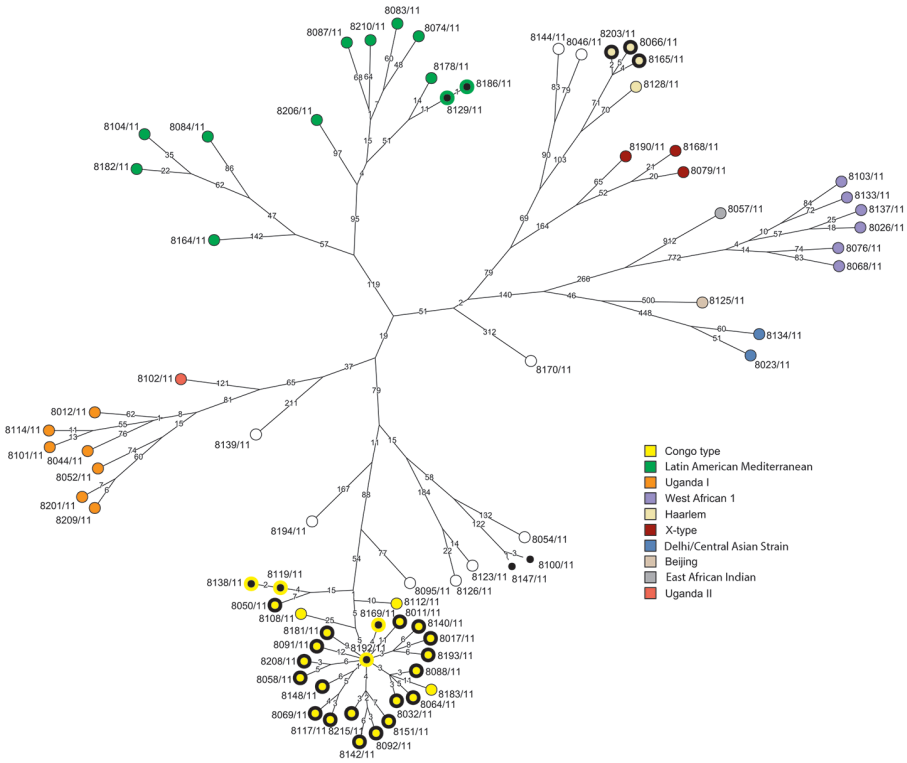

Figure 2. Maximum parsimony tree of *Mycobacterium tuberculosis* complex isolates from Congo. The tree was calculated on the basis of the concatenated single-nucleotide polymorphism (SNP) list. Branch labels indicate SNP distances; node labels represent the strain identifiers. Clusters based on pairwise comparisons were plotted on the tree: clusters ≤5 SNPs, nodes filled in black; clusters ≤12 SNPs, nodes with bold black outlines.

Our genotyping approach showed that all strains investigated were either *M. tuberculosis* or *M. africanum*; *M. tuberculosis* was most prevalent. These findings are in accordance with most recent studies from the other African countries reporting a predominance of *M. tuberculosis* strains (*25–28*). The high prevalence of *M. tuberculosis* detected in our investigation suggests that this predominance of *M. tuberculosis* strains might equally be the case for Congo, mainly driven by the newly described Congo-type sublineage.

The closely related strains of the Euro-American Congo-type sublineage were responsible for 35% of TB cases in the study population and showed a low pairwise genetic distance resulting in a high genome-based cluster rate, indicating ongoing recent transmission. It is tempting to speculate that strains from the Congo type are highly successful in the area and are recently expanding in the region of Brazzaville. On the other hand, strains of other MTBC lineages showed a higher degree of genetic diversity and formed smaller clusters with <3 strains. Such a high diversity is somewhat unexpected because, in a TB-endemic area, only few dominant clones, such as the Congo type, are hypothesized to circulate (*29*). The aforementioned findings point to a particular capacity of strains of the Congo type to spread in the area; for example, because of adaptation to the host population, as already postulated for other MTBC lineages (*5,30*). Accordingly, the diversity in the other MTBC lineages might reflect a higher rate of cases from reactivation of past TB infections, as suggested by a study conducted in South Africa with a high incidence of TB and high strain diversity (*31*).

The dominance of particular highly spreading clones, however, appears to be a more general phenomenon seen in several high-incidence areas. For example, strains of the Beijing lineage dominate in East Asia (*5,32*); the F11 *M. tuberculosis* genotype in Western Cape, South Africa (*31*); the LAM10-Cam family in Cameroon (*25*); and the K family, a sublineage of the Beijing genotype, in South Korea (*33*). In focusing on Africa, recent investigations revealed that Cameroon MTBC strains are responsible for most TB cases in several West Africa countries, such as Ghana and Cameroon (*25,27*); strains of the Uganda sublineage predominate in East Africa (*34*); and strains of the Sierra Leone sublineage predominate in Sierra Leone (*35*). Together, these data indicate marked differences in circulating mycobacterial strains in different Central Africa countries, suggesting a region-specific selection and spread of dominant sublineages of the Euro-American lineage.

Genome analysis enables not only high-resolution description of MTBC population diversity but also improved resolution of strains in recent transmission chains (*14,16*). Consistently, a SNP distance up to 5 SNPs was found in strains from confirmed direct human-to-human transmission, whereas a 12-SNP distance was proposed as a threshold to define larger cluster/transmission networks (*14,16*). Using these thresholds, we found that strains of the Congo-type sublineage formed a genetically homogenous

group with median pairwise genome SNP distances of 20 SNPs (range 2–53 SNPs) and a large number of Congo-type strains in clusters differing by ≤5 SNPs (15%) or a maximum of ≤12 SNPs (88.5%). This finding supports the presence of a larger transmission network of the Congo-type strains that presumably emerged in the past 53 years.

Only 2 of the 74 isolates characterized in this study were MDR, whereas recently Aubry et al. reported a higher MDR rate among MTBC isolates from Brazzaville and Pointe Noire (15 of 46 strains investigated; 10 strains belonging to the same lineage based on MIRU-VNTR and spoligotyping) (*36*). However, 7 of the 15 MDR strains in that study were obtained from retreated TB patients within a short 9-day sampling period. Even though we detected only 2 MDR strains of different lineages in the population in our study, the presence of MDR strains might pose a serious future challenge to public health authorities because these strains might have the potential to spread in the population; in line with this concern, 8 of the MDR strains described by Aubry et al. have been isolated from persons with newly diagnosed cases, which the authors suggested might reflect the transmission of an MDR clone (*36*). Thus, the development of drug-resistant phenotypes among particular lineages circulating in Congo, especially the successful Congo type, should be revisited as part of a larger investigation to evaluate the actual extent of MDR TB and potential pockets of transmission in Congo.

The data from our genome-based investigation of circulating MTBC strains in Congo demonstrate the presence of a new, predominant, and highly transmissible sublineage, the Congo type, which belongs to the Euro-American lineage. Larger molecular epidemiologic studies with respect to sociogeographic data and in addition to traditional contact tracing investigations will be required in Central Africa to gain a better understanding of recent transmission networks, the emergence of dominant lineages, and the prevalence of drug-resistant phenotypes in this wider geographic setting. Such studies may be implemented in large networks, such as the Central Africa Network on Tuberculosis, HIV/AIDS and Malaria (sponsored by the European and Developing Countries Clinical Trials Partnership), with the objective of conducting baseline investigations of TB in its members state: Gabon, Cameroon (*37,38*), and Congo (*6*).

Acknowledgments

We are grateful to the patients who participated in this study. We thank Pembe Issamou Mayengue and Fondation Congolaise pour la Recherche Médicale and Centre Antituberculeux de Brazzaville staff for their technical assistance in collecting samples. We acknowledge Tanja Ubben, Tanja Struve-Sonnenschein, Julia Zallet, and Anja Lüdemann for their technical assistance in typing and sequencing the MTBC strains and Matthias Merker for assistance in the coalescent-based analysis.

This work was supported by the regional network of excellence CANTAM (Central Africa Network on Tuberculosis, HIV/AIDS and Malaria) funded by European and Developing Countries Clinical Trials Partnership (EDCTP). Parts of this work were supported by the German Center for Infection Research and by grants from the European Community's Seventh Framework Program (FP7/2007-2013) under grant agreement no. 278864 in the framework of the European Union PathoNGenTrace project.

L.S.G.L., E.M.T., and J.C.V. are recipients of CANTAM Network training grants. V.P., F.N., S.N., A.S., and M.F. were supported by the German Academic Exchange Service (DAAD) PAGEL Program (Partnerschaften für den Gesundheitssektor in Entwicklungsländern) of the University of Tübingen to attend expert meetings and workshops throughout the duration of the project.

Dr. Malm is a postdoctoral researcher in the Molecular and Experimental Mycobacteriology group at the Research Center Borstel, Leibniz-Center for Medicine and Biosciences, in Borstel. His main research interests are the molecular characterization of virulence determinants of *M. tuberculosis* and the molecular epidemiology of TB.

References

1. World Health Organization. Global tuberculosis report 2015. 20th ed. Geneva: The Organization; 2015.
2. Comas I, Gagneux S. The past and future of tuberculosis research. PLoS Pathog. 2009;5:e1000600.
3. van Soolingen D, Hermans PW, de Haas PE, Soll DR, van Embden JD. Occurrence and stability of insertion sequences in *Mycobacterium tuberculosis* complex strains: evaluation of an insertion sequence–dependent DNA polymorphism as a tool in the epidemiology of tuberculosis. J Clin Microbiol. 1991;29:2578–86.
4. Allix-Béguec C, Fauville-Dufaux M, Stoffels K, Ommeslag D, Walravens K, Saegerman C, et al. Importance of identifying *Mycobacterium bovis* as a causative agent of human tuberculosis. Eur Respir J. 2010;35:692–4. http://dx.doi.org/10.1183/09031936.00137309
5. Merker M, Blin C, Mona S, Duforet-Frebourg N, Lecher S, Willery E, et al. Evolutionary history and global spread of the *Mycobacterium tuberculosis* Beijing lineage. Nat Genet. 2015;47:242–9. http://dx.doi.org/10.1038/ng.3195
6. Linguissi LS, Mayengue PI, Sidibé A, Vouvoungui JC, Missontsa M, Madzou-Laboum IK, et al. Prevalence of national treatment algorithm defined smear positive pulmonary tuberculosis in HIV positive patients in Brazzaville, Republic of Congo. BMC Res Notes. 2014;7:578. http://dx.doi.org/10.1186/1756-0500-7-578
7. Deutsches Institut für Normung. Medical microbiology. Diagnosis of tuberculosis. Part 3: detection of mycobacteria by culture methods. DIN 58943–3. Berlin; Beuth Verlag; 2011.
8. Kent PT, Kubica GP. Public health mycobacteriology: a guide for the level III laboratory. Atlanta: Centers for Disease Control; 1985.
9. Kamerbeek J, Schouls L, Kolk A, van Agterveld M, van Soolingen D, Kuijper S, et al. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. J Clin Microbiol. 1997;35:907–14.
10. Gori A, Bandera A, Marchetti G, Degli Esposti A, Catozzi L, Nardi GP, et al. Spoligotyping and *Mycobacterium tuberculosis*. Emerg Infect Dis. 2005;11:1242–8. http://dx.doi.org/10.3201/eid1108.040982

11. Allix-Béguec C, Harmsen D, Weniger T, Supply P, Niemann S. Evaluation and strategy for use of MIRU-VNTRplus, a multifunctional database for online analysis of genotyping data and phylogenetic identification of *Mycobacterium tuberculosis* complex isolates. J Clin Microbiol. 2008;46:2692–9. http://dx.doi.org/10.1128/JCM.00540-08
12. Weniger T, Krawczyk J, Supply P, Niemann S, Harmsen D. MIRU-VNTRplus: a Web tool for polyphasic genotyping of *Mycobacterium tuberculosis* complex bacteria. Nucleic Acids Res. 2010;38(Web Server issue):W326–331.
13. Blom J, Jakobi T, Doppmeier D, Jaenicke S, Kalinowski J, Stoye J, et al. Exact and complete short-read alignment to microbial genomes using Graphics Processing Unit programming. Bioinformatics. 2011;27:1351–8. http://dx.doi.org/10.1093/bioinformatics/btr151
14. Roetzer A, Diel R, Kohl TA, Rückert C, Nübel U, Blom J, et al. Whole genome sequencing versus traditional genotyping for investigation of a Mycobacterium tuberculosis outbreak: a longitudinal molecular epidemiological study. PLoS Med. 2013;10:e1001387.
15. Comas I, Chakravartti J, Small PM, Galagan J, Niemann S, Kremer K, et al. Human T cell epitopes of *Mycobacterium tuberculosis* are evolutionarily hyperconserved. Nat Genet. 2010;42:498–503. http://dx.doi.org/10.1038/ng.590
16. Walker TM, Ip CL, Harrell RH, Evans JT, Kapatai G, Dedicoat MJ, et al. Whole-genome sequencing to delineate *Mycobacterium tuberculosis* outbreaks: a retrospective observational study. Lancet Infect Dis. 2013;13:137–46. http://dx.doi.org/10.1016/S1473-3099(12)70277-3
17. Coll F, McNerney R, Guerra-Assunção JA, Glynn JR, Perdigão J, Viveiros M, et al. A robust SNP barcode for typing *Mycobacterium tuberculosis* complex strains. Nat Commun. 2014;5:4812. http://dx.doi.org/10.1038/ncomms5812
18. Helaers R, Milinkovitch MC. MetaPIGA v2.0: maximum likelihood large phylogeny estimation using the metapopulation genetic algorithm and other stochastic heuristics. BMC Bioinformatics. 2010;11:379. http://dx.doi.org/10.1186/1471-2105-11-379
19. Posada D, Crandall KA. MODELTEST: testing the model of DNA substitution. Bioinformatics. 1998;14:817–8. http://dx.doi.org/10.1093/bioinformatics/14.9.817
20. Zhang H, Gao S, Lercher MJ, Hu S, Chen W-H. EvolView, an online tool for visualizing, annotating and managing phylogenetic trees. Nucleic Acids Res. 2012;40:W569-72. http://dx.doi.org/10.1093/nar/gks576
21. Drummond AJ, Suchard MA, Xie D, Rambaut A. Bayesian phylogenetics with BEAUti and the BEAST 1.7. Mol Biol Evol. 2012;29:1969–73. http://dx.doi.org/10.1093/molbev/mss075
22. Homolka S, Projahn M, Feuerriegel S, Ubben T, Diel R, Nübel U, et al. High resolution discrimination of clinical Mycobacterium tuberculosis complex strains based on single nucleotide polymorphisms. PLoS ONE. 2012;7:e39855.
23. Tessema B, Beer J, Merker M, Emmrich F, Sack U, Rodloff AC, et al. Molecular epidemiology and transmission dynamics of *Mycobacterium tuberculosis* in northwest Ethiopia: new phylogenetic lineages found in northwest Ethiopia. BMC Infect Dis. 2013;13:131. http://dx.doi.org/10.1186/1471-2334-13-131
24. Olaru ID, Rachow A, Lange C, Ntinginya NE, Reither K, Hoelscher M, et al. Ascertaining in vivo virulence of *Mycobacterium tuberculosis* lineages in patients in Mbeya, Tanzania. Int J Tuberc Lung Dis. 2015;19:70–3. http://dx.doi.org/10.5588/ijtld.14.0403
25. Niobe-Eyangoh SN, Kuaban C, Sorlin P, Cunin P, Thonnon J, Sola C, et al. Genetic biodiversity of Mycobacterium tuberculosis complex strains from patients with pulmonary tuberculosis in Cameroon. J Clin Microbiol. 2003;41:2547–53. http://dx.doi.org/10.1128/JCM.41.6.2547-2553.2003
26. Koro Koro F, Kamdem Simo Y, Piam FF, Noeske J, Gutierrez C, Kuaban C, et al. Population dynamics of tuberculous bacilli in Cameroon as assessed by spoligotyping. J Clin Microbiol. 2013;51:299–302. http://dx.doi.org/10.1128/JCM.01196-12
27. Sidze LK, Tekwu EM, Kuaban C, Assam J-PA, Tedom J-C, Niemann S, et al. Estimates of genetic variability of *Mycobacterium tuberculosis* complex and its association with drug resistance in Cameroon. Adv Infect Dis. 2013;3:55. http://dx.doi.org/10.4236/aid.2013.31007
28. Mbugi EV, Katale BZ, Streicher EM, Keyyu JD, Kendall SL, Dockrell HM, et al. Mapping of *Mycobacterium tuberculosis* complex genetic diversity profiles in Tanzania and other African countries. PLOS One. 2016;11:e0154571.
29. Hermans PW, Messadi F, Guebrexabher H, van Soolingen D, de Haas PE, Heersma H, et al. Analysis of the population structure of *Mycobacterium tuberculosis* in Ethiopia, Tunisia, and the Netherlands: usefulness of DNA typing for global tuberculosis epidemiology. J Infect Dis. 1995;171:1504–13. http://dx.doi.org/10.1093/infdis/171.6.1504
30. Gagneux S, Long CD, Small PM, Van T, Schoolnik GK, Bohannan BJM. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. Science. 2006;312:1944–6. http://dx.doi.org/10.1126/science.1124410
31. Victor TC, de Haas PEW, Jordaan AM, van der Spuy GD, Richardson M, van Soolingen D, et al. Molecular characteristics and global spread of *Mycobacterium tuberculosis* with a Western Cape F11 genotype. J Clin Microbiol. 2004;42:769–72. http://dx.doi.org/10.1128/JCM.42.2.769-772.2004
32. van Soolingen D, Qian L, de Haas PE, Douglas JT, Traore H, Portaels F, et al. Predominance of a single genotype of *Mycobacterium tuberculosis* in countries of east Asia. J Clin Microbiol. 1995;33:3234–8.
33. Park YK, Bai GH, Kim SJ. Restriction fragment length polymorphism analysis of *Mycobacterium tuberculosis* isolated from countries in the Western Pacific Region. J Clin Microbiol. 2000;38:191–7.
34. Asiimwe BB, Koivula T, Källenius G, Huard RC, Ghebremichael S, Asiimwe J, et al. *Mycobacterium tuberculosis* Uganda genotype is the predominant cause of TB in Kampala, Uganda. Int J Tuberc Lung Dis. 2008;12:386–91.
35. Feuerriegel S, Oberhauser B, George AG, Dafae F, Richter E, Rüsch-Gerdes S, et al. Sequence analysis for detection of first-line drug resistance in *Mycobacterium tuberculosis* strains from a high-incidence setting. BMC Microbiol. 2012;12:90. http://dx.doi.org/10.1186/1471-2180-12-90
36. Aubry A, Sougakoff W, Bodzongo P, Delcroix G, Armand S, Millot G, et al. First evaluation of drug-resistant *Mycobacterium tuberculosis* clinical isolates from Congo revealed misdetection of fluoroquinolone resistance by line probe assay due to a double substitution T80A–A90G in GyrA. PLoS One. 2014;9:e95083.
37. Sidze LK, Mouafo Tekwu E, Kuaban C, Assam Assam J-P, Tedom J-C, Eyangoh S, et al. Strong decrease in streptomycin-resistance and absence of XDR 12 years after the reorganization of the National Tuberculosis Control Program in the Central Region of Cameroon. PLoS One. 2014;9:e98374.
38. Tekwu EM, Sidze LK, Assam J-P, Tedom J-C, Tchatchouang S, Makafe GG, et al. Sequence analysis for detection of drug resistance in *Mycobacterium tuberculosis* complex isolates from the Central Region of Cameroon. BMC Microbiol. 2014;14:113. http://dx.doi.org/10.1186/1471-2180-14-113

Address for correspondence: Stefan Niemann, Research Center Borstel, Leibniz Center for Medicine and Biosciences, Parkallee 1, 23845 Borstel, Germany; email: sniemann@fz-borstel.de