Direct DNA Extraction from *Mycobacterium tuberculosis* Frozen Stocks as a Reculture-Independent Approach to Whole-Genome Sequencing

K. Bjorn-Mortensen,\textsuperscript{a,b,d} J. Zallet,\textsuperscript{e,f} T. Lillebaek,\textsuperscript{b} A. B. Andersen,\textsuperscript{c} S. Niemann,\textsuperscript{b,f} E. M. Rasmussen,\textsuperscript{b} T. A. Kohl\textsuperscript{a}

Department of Epidemiology Research, Statens Serum Institut, Copenhagen, Denmark;\ International Reference Laboratory of Mycobacteriology, Statens Serumin Institut, Copenhagen, Denmark;\ Department of Infectious Diseases, Copenhagen University Hospital and University of Southern Denmark, Odense, Denmark;\ Greenland’s Center of Health Research, Nuuk, Greenland; Molecular Mycobacteriology, Forschungszentrum Borstel, Leibniz-Zentrum für Medizin und Biowissenschaften, Borstel, Germany; German Center for Infection Research (DZIF), Borstel, Germany

Culturing before DNA extraction represents a major time-consuming step in whole-genome sequencing of slow-growing bacteria, such as *Mycobacterium tuberculosis*. We report a workflow to extract DNA from frozen isolates without reculturing. Prepared libraries and sequence data were comparable with results from recultured aliquots of the same stocks.

In recent years, studies employing whole-genome sequencing (WGS) of *Mycobacterium tuberculosis* isolates have demonstrated its value in understanding transmission patterns, recurrent tuberculosis (TB), development of drug resistance, and bacterial evolution (1–9). In parallel, improvements of next-generation sequencing platforms and library preparation workflows make it possible to determine the genome sequence from bacterial DNA samples in the time span of 1 week. Reculturing isolates for DNA isolation has been reported as a necessary step in published WGS studies (2, 4, 7, 10). As culturing of slow-growing bacteria, such as *M. tuberculosis*, takes from 1 week to several weeks, it constitutes the main time-consuming process in WGS projects (11–14).

While initially large amounts of DNA were required for reliable WGS library preparation, newly developed library preparation protocols for bacterial samples typically require about 1 to 10 ng of DNA (e.g., Illumina Nextera XT [1 ng], New England Biolabs NEBNext Ultra [5 to 1,000 ng], and Bioo Scientific NEXTflex ChIP-Seq [1 to 10 ng]). Therefore, faster and reliable methods for DNA isolation would enable a considerable decrease in the time needed to perform WGS analyses. In this regard, a recent study proposed a WGS workflow starting from early positive liquid cultures of the MGIT system (15), potentially enhancing the speed of WGS procedures as part of routine diagnostics.

Furthermore, reculturing of isolates from even well-maintained frozen stocks can fail entirely, usually excluding the respective isolate from any further analysis. A recent publication reported failure rates of up to 50% for *M. tuberculosis* glycerol stocks (16). New methods enabling WGS analysis directly from frozen stocks without reculturing can rescue genotype information, especially from historic collections of isolates.

In this study, we investigated and tested a protocol for performing WGS of DNA extracted directly from frozen glycerol stocks, which were all historic isolates from patients diagnosed with fully susceptible TB between 1992 and 2012, circumventing the step of reculturing altogether.

For validation, we sequenced DNA from cultured aliquots of the same frozen stocks in parallel. In total, we included 40 frozen glycerol stocks (1992 to 2012) stored at −80°C at the International Reference Laboratory of Mycobacteriology at the Statens Serumin Institut (Copenhagen). All were processed according to a standard lysis protocol with heat inactivation and sonication as is usually used for PCR (17, 18). Lysates were concentrated with Microcon filters (Merck KGaA, Darmstadt, Germany), followed by a purification with ethanol (EtOH) precipitation and bead clean up (AMPure XP bead; Beckman Coulter, Krefeld, Germany) (15). For culture, aliquots of frozen stocks were put on solid Löwenstein-Jensen (LJ) slants and incubated at 35°C. After approximately 10 weeks of growth, DNA extraction was performed with the traditional cetyltrimethylammonium bromide (CTAB) procedure (19) from visible colonies. For each extraction method, we measured final DNA concentrations with the Qubit 2.0 Fluorometer (Fig. 1; see also the supplemental material for further details). Libraries for WGS were prepared from DNA samples with the Nextera XT kits and run on Illumina next-generation sequencing platforms (MiSeq) as instructed by the manufacturer (Illumina, San Diego, CA, USA). Respective fastq files were submitted to the EMBL EBI ENA short read archive (accession number PRJEB9308).

Reads were mapped to the genome of the *M. tuberculosis* reference strain H37Rv (GenBank accession number NC_000962.3) with the alignment program SARUMAN (20). For variant detection in mapped reads, we employed minimum thresholds of 10× coverage and 75% allele frequency and filtered results for repetitive regions (1, 2). Detected variants were manually curated using the Integrative Genome Viewer software (21) to visualize mapped reads.

Of the 40 selected glycerol stocks, 35 were successfully sequenced from DNA extracted from glycerol stocks and culture.
Five isolates were excluded due to failure to grow ($n = 4$) on any media (LJ, liquid Dubos, blood agar) or continuous failure of library preparation from DNA directly isolated from stock ($n = 2$). For the remaining 35 isolates, the median DNA concentration from direct isolation was 0.67 ng/µl (range, 0.2 to 3.8 ng/µl). DNA extraction with direct lysis of glycerol stocks, DNA concentration, and purification takes less than a day in total. In comparison, most samples grown on LJ medium are normally visible within 3 to 8 weeks, and DNA extraction with the CTAB procedure takes 3 days.

For the two methods, mapping of the genome to the reference strain revealed a genome-wide coverage in the range of 98.8% to 99.5% and medians of 99.3% for glycerol stocks and 99.4% for cultured aliquots. The median percentages of the reference genome fulfilling thresholds for variant detection was 97.9% (range, 96.6% to 98.9%) for glycerol stocks and 98.5% (range, 97.5% to 98.9%) for cultured aliquots.

Analysis of whole-genome sequencing data revealed only two instances where paired genotypes from glycerol stocks or cultured material differed in detected variants. In these cases, sequence data for cultured aliquots exhibited one additional single nucleotide polymorphism (SNP) compared to the H37Rv reference genome. We detected one SNP in the rRNA gene Rvnr01 at position 1471870 for sample GE0078 and one nonsynonymous SNP in the Rv3085 gene-coding region at position 3199069 for sample GE0079 (see Table 1).

Library preparation and whole-genome sequencing succeeded for 38 (95%) of the glycerol stocks with a DNA yield well in the range required for WGS library preparation (range, 0.2 to 3.8 ng/µl). In contrast, culture and DNA extraction succeeded for 36 of the 40 (90.0%) glycerol stocks. In total, DNA was obtained from 39 of the 40 lysed frozen stocks with one or both methods. Reference mapping-based analysis gave excellent results, with at least 96.6% of the reference genome covered at high quality (10× coverage, 75% allele identity) for directly isolated DNA and DNA extracted from cultured stocks.

In conclusion, DNA extraction, library preparation, and sequencing was successful in three samples that exhibited no growth on LJ slants; hence, the proposed method can even be used for samples where reculturing is not possible. As results are comparable between DNA obtained directly from frozen stocks versus DNA extracted from cultured aliquots, this study shows that reculturing of frozen glycerol stocks is not needed in order to perform whole-genome sequencing analysis for epidemiologic research. On the contrary, reculturing might introduce mutations in the genome, which likely happened in two cases for our sample set. In samples with distinct clones, culturing might also change the relative ratio for intrapatient evolving clones (7) and for coinfections (22). This would be even more pronounced for strains carrying fitness-reducing resistance mutations (23). Since direct lysis of frozen stocks, concentration, and purification can be performed in 1 day.

**FIG 1** The workflow of DNA extraction from 40 frozen glycerol stocks with same day direct lysis from frozen glycerol stocks or reculturing of aliquots.
while reculturing and DNA extraction with the CTAB procedure takes several weeks, this finding can facilitate implementation of WGS as a routine clinical tool.

**Nucleotide sequence accession number.** Sequence reads were submitted to the EMBL EBI ENA short read archive under accession number PRJEB9308.

**ACKNOWLEDGMENTS**

We thank Timothy Walker, Viola Schleusener, and Christine Rohde for helpful discussions and Antonina Votintseva and Louise Pankhurst for excellent advice and help with the bead-based DNA purification protocol. We also thank the laboratory staff at the International Reference Laboratory of Mycobacteriology, Statens Serum Institut, Denmark for their help with sample preparation.

All authors contributed to the conception and design of the study, the analyses of data, and writing of the article.

Parts of this work have been supported by the Commission for Scientific Research in Greenland (0605-00062B), Seed Funding Initiative for Collaboration of the Universities of Southern Denmark, Kiel, Hamburg, and Aarhus, Lundbeck Foundation (R151-2013-14628), Novo Nordisk Foundation (7651), the European Union TB-PAN-NET (FP7-223681), Patho-Ngen-Trace (FP7-227864-2) projects, and the German Center for Infection Research (DZIF). The funders had no role in study design, data acquisition, analysis, interpretation, or writing of the article.
REFERENCES


