Tuberculosis (TB) genotyping is a laboratory-based genetic analysis of the bacteria that cause TB disease (i.e., any of the organisms in the *Mycobacterium tuberculosis* complex). When combined with epidemiologic data, TB genotyping has sufficient discriminatory power to help find TB cases likely to be in the same chain of transmission or determine that cases are not related (1). Since 2004, >70,000 *M. tuberculosis* isolates have been genotyped through partnerships between CDC, national genotyping laboratories, and state and local public health departments, with a goal to genotype at least one *M. tuberculosis* isolate for each case of culture-positive TB in the United States. National genotype surveillance coverage, or the proportion of culture-positive TB cases with a genotyped isolate, increased from 51.2% in 2004 to 88.2% in 2010. The TB Genotyping Information Management System (TB GIMS), accessible to public health departments through a secure, online web portal, was launched in 2010. TB GIMS enables systematic collection of genotyping results, which have been available since 2004, and integrates those results with epidemiologic, geographic, demographic, and clinical data collected by the National TB Surveillance System (NTSS) since 1993. Genotyping timeliness, represented by the median time from specimen collection until linked genotyping results and surveillance data are available to TB GIMS users, improved from 22 weeks in July 2010 to 11 weeks in December 2010. These improvements in genotype surveillance coverage and timeliness will improve outbreak detection efforts and enable more in-depth studies of TB epidemiology, leading to better use of limited public health resources.

Analysis of *M. tuberculosis* genotypes* has enhanced TB control with its ability to detect unsuspected transmission links between cases, confirm suspected links, identify unknown or difficult to investigate transmission settings, alert public health departments to possible transmission across geographic reporting areas, identify potential outbreaks (i.e., a group of TB cases with genotype results and epidemiologic links consistent with recent transmission, where control of transmission exceeds locally available resources), refine outbreak case definitions, and identify false-positive TB culture results (2). Cases with indistinguishable genotypes that are close in space and time are considered genotype clusters and might represent an outbreak. Cases with a genotype unique to a given time and place are less likely to be related to another case. The ability to detect outbreaks based on genotyping results depends on sufficient genotype surveillance data (3). Without adequate genotyping data, single genotyped isolates from chains of transmission will appear to be unique.

TB GIMS facilitates systematic data collection of genotyping results and integrates genotyping results with epidemiologic, geographic, demographic, and clinical data collected by NTSS (4). Genotype data are uploaded into TB GIMS by national genotyping laboratories as they become available, and surveillance data from NTSS are uploaded into TB GIMS at least biweekly at CDC. Although TB outbreaks tend to develop slowly (months to years), prompt intervention
during an outbreak can interrupt transmission; thus, timeliness in linking genotyping results to surveillance data (by state and local public health departments) is essential for prompt outbreak detection. TB GIMS sends users automated e-mail alerts when a genotype cluster in their jurisdiction has grown to a higher than expected geospatial concentration in a specific county, compared with the national distribution of that genotype.

For this report, 2004–2010 TB GIMS data from 51 reporting areas in the United States (50 states and the District of Columbia) were analyzed. Measures of geospatial concentration were calculated by the log-likelihood ratio (LLR) for a given genotype cluster in a single county during a 3-year period. A LLR >5.0 indicates a potential outbreak. TB GIMS timeliness data were examined during July–December 2010. For each isolate, the date of arrival at the national genotyping laboratory, date of genotyping result, and date that genotyping results and surveillance data were linked in the system were examined. A record was considered complete when genotyping results (from national genotyping laboratories) and surveillance data (from NTSS) both were entered in the system and linked by public health departments.

Genotype surveillance coverage has increased from 51.2% in 2004 to 88.2% in 2010 (Figure 1). In 2010, 40 (83.3%) of 48 reporting areas had >80% genotype surveillance coverage, compared with 26 (51.0%) of 51 reporting areas in 2004. During 2008–2010, a total of 23,108 TB cases had at least one genotyped isolate; 7,942 (34.4%) were part of 2,184 county-based genotype clusters. Of these clusters, 1,679 (76.9%) clusters consisted of two or three cases, compared with 100 (4.6%) clusters with ≥10 cases (Figure 2). The most common genotype was seen in 932 (4.0%) of all genotyped cases; at least one case with this genotype was identified in 43 of the 51 reporting areas. Among all genotype clusters, 378 (17.3%) were identified as potential outbreaks based on elevated geospatial concentration (LLR >5.0) during 2008–2010.

Since its launch in March 2010, TB GIMS use increased rapidly, with >28,000 logins by 349 state and local users in the first 10 months after release. During July–December 2010, the median time from specimen collection until a complete record was available (with linked genotyping results and surveillance data) in TB GIMS was 16 weeks (interquartile range: 12–22 weeks). This included medians of 8 weeks from specimen collection until the TB isolate arrived at the genotyping laboratory, 1 week for genotyping laboratory turnaround time, and 14 weeks for linking genotyping results with surveillance data in TB GIMS. Median time to a complete record improved from 22 weeks (n = 547 isolates; interquartile range: 15–30 weeks) for specimens collected in July 2010, to 11 weeks (n = 289 isolates; interquartile range: 9–14 weeks) for specimens collected in December 2010.

Reported by
Juliana Grant, MD, Steve Kammerer, MBA, Div of Tuberculosis Elimination, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention; Brian Baker, MD, Lindsay Kim, MD, EIS officers, CDC. Corresponding contributor: Lindsay Kim, lkim@cdc.gov, 404-639-5218.

Editorial Note
TB genotyping has become a commonly used tool for TB outbreak detection and investigations in the United States (B. Baker, MD, CDC, unpublished survey of TB GIMS users, 2010). In one study examining the added value of TB genotyping, 38% of all epidemiologic links were discovered only after using genotyping to connect previously unlinked cases. In addition, of the epidemiologic links discovered during conventional contact investigations, assumed transmission among 29% of case pairs was refuted by genotyping data (5). Only 17.3% of the...
2,184 genotype clusters in the United States during 2008–2010 had geospatial concentrations indicating a potential outbreak. Genotyping data such as these allow local or state public health departments to target interventions among the many cases and epidemiologic clusters identified in a jurisdiction, potentially saving financial and human resources.

Because of the availability of several years of national surveillance data with accompanying high genotype coverage, population-level studies using genotyping data are now possible. Genotyping data can enable a broader understanding of TB epidemiology, which can be applied to future TB control efforts. As TB incidence declines in the United States, TB increasingly is found in harder-to-reach populations and locations, and genotyping data have been used to identify these pockets of transmission that require public health intervention (6). A recent study using national genotyping data found that most TB disease in the foreign-born population in the United States likely has resulted from activation of latent M. tuberculosis infection (rather than recent transmission in the United States), emphasizing the importance of identifying and treating latent infection to decrease the incidence of TB disease in this group (7). Another study used national genotyping data to demonstrate that as much as one fourth of TB cases might represent recent transmission, emphasizing the critical importance of early contact investigations to TB control (8). Finally, TB genotyping data can be used to better understand the development of outbreaks. Nationally, most county-based genotype clusters are small (Figure 2) and do not grow larger. For small genotype clusters (<4 cases) that grow in size and are classified as outbreaks using field data, linked genotyping and surveillance data have been used to identify factors that might increase the likelihood of clusters becoming an outbreak (CDC, unpublished data, 2012).

Analyses of large genotype datasets and prompt outbreak detection now are possible because of substantial increases in genotype surveillance coverage during 2004–2010. Improvements in the timeliness of linking genotyping results to surveillance data will improve TB outbreak detection efforts further. As the United States strives toward TB elimination, genotype surveillance can lead to continued refinement of TB control activities, making the best use of limited public health resources at local, state, and national levels.

Acknowledgments
Clinical laboratories and local and state public health department personnel. Sandy Althomsons, MA, Eleanor Click, MD, Lauren Cowan, PhD, Divia Forbes, MPH, Anne Marie France, PhD, Tracie J. Gardner, PhD, Smita Ghosh, MS, Michael Iademarco, MD, Philip LoBue, MD, Patrick Moonan, DrPH, Thomas Navin, MD, Jagadheeswari Adhimurthy, MS, Jeffrey Chrismon, Joel Haney, Shonna Luten, Div of TB Elimination, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, CDC.

References
What is already known on this topic?

Tuberculosis (TB) genotyping, a laboratory-based genetic analysis of the bacteria that cause TB disease, provides sufficient discriminatory power to confirm or refute links in the chain of transmission. The TB Genotyping Information Management System (TB GIMS), a secure, online web portal accessible by public health authorities, facilitates systematic data collection of genotyping results and integrates genotyping results with epidemiologic, geographic, demographic, and clinical data.

What is added by this report?

Since 2004, >70,000 *Mycobacterium tuberculosis* isolates have been genotyped through partnerships between CDC, national genotyping laboratories, and state and local public health departments. National genotype surveillance coverage, or the proportion of culture-positive TB cases with a genotyped isolate increased from 51.2% in 2004 to 88.2% in 2010. Genotyping timeliness, represented by the median time from specimen collection until linked genotyping results and surveillance data are available to TB GIMS users, improved from 22 weeks in July 2010, to 11 weeks in December 2010.
What are the implications for public health practice?

As the United States strives toward TB elimination, genotype surveillance can lead to more timely outbreak detection and continued refinement of TB control activities, making the best use of limited public health resources at local, state, and national levels.

FIGURE 1. Reported culture-positive tuberculosis (TB) cases and national TB genotype surveillance coverage* by year — United States, 2004–2010

* Proportion of culture-positive TB cases with at least one genotyped isolate.

† Includes 50 states and the District of Columbia.

Alternate Text: The figure above shows reported culture-positive tuberculosis (TB) cases and national TB genotype surveillance coverage by year in the United States, during 2004-2010. Genotype surveillance coverage has increased from 51.2% in 2004 to 88.2% in 2010. In 2010, 40 (83.3%) of 48 reporting areas had >80% genotype surveillance coverage, compared with 26 (51.0%) of 51 reporting areas in 2004.

FIGURE 2. Number of county-based tuberculosis genotype clusters,* by cluster size (number of isolates) — United States, 2008–2010
* Genotype cluster is defined as two or more *Mycobacterium tuberculosis* isolates that share matching genotypes in a county during 2008–2010.

Alternate Text: The figure above shows the number of county-based tuberculosis genotype clusters, by cluster size (number of isolates) in the United States, during 2008-2010. During 2008-2010, a total of 23,108 TB cases had at least one genotyped isolate; 7,942 (34.4%) were part of 2,184 county-based genotype clusters. Of these clusters, 1,679 (76.9%) clusters consisted of two or three cases, compared with 100 (4.6%) clusters with ≥10 cases.

Use of trade names and commercial sources is for identification only and does not imply endorsement by the U.S. Department of Health and Human Services.

References to non-CDC sites on the Internet are provided as a service to *MMWR* readers and do not constitute or imply endorsement of these organizations or their programs by CDC or the U.S. Department of Health and Human Services. CDC is not responsible for the content of pages found at these sites. URL addresses listed in *MMWR* were current as of the date of publication.

All *MMWR* HTML versions of articles are electronic conversions from typeset documents. This conversion might result in character translation or format errors in the HTML version. Users are referred to the electronic PDF version (http://www.cdc.gov/mmwr) and/or the original *MMWR* paper copy for printable versions of official text, figures, and tables. An original paper copy of this issue can be obtained from the Superintendent of Documents, U.S. Government Printing Office (GPO), Washington, DC 20402-9371; telephone: (202) 512-1800. Contact GPO for current prices.

**Questions or messages regarding errors in formatting should be addressed to mmwrrq@cdc.gov.**