TB Elimination

Tuberculosis Genotyping

What is tuberculosis (TB) genotyping?

TB genotyping is a laboratory-based approach used to analyze the genetic material (e.g., DNA) of *Mycobacterium tuberculosis*, the bacteria that cause TB disease. The total genetic content is referred to as the genome. Specific sections of the *M. tuberculosis* genome form distinct genetic patterns that help distinguish different strains of *M. tuberculosis*.

Why use TB genotyping?

TB genotyping results, when combined with epidemiologic data, help identify persons with TB disease involved in the same chain of recent transmission. In the same way, TB genotyping helps distinguish between persons whose TB disease is the result of TB infection that was acquired in the past, as compared to recently or newly acquired infection with development of TB disease. TB genotyping is a tool that can add value to conventional contact investigation. Below are some applications:

- Discover unsuspected transmission relationships between TB patients
  - Identify unknown or unusual transmission settings, such as bars or clubs, instead of traditional settings like home and workplace
  - Uncover inter-jurisdictional transmission
- Establish criteria for outbreak-related case definitions
- Identify additional persons with TB disease involved in an outbreak
- Determine completeness of contact investigations
- Detect laboratory cross-contamination event
- Distinguish recent infection (with development of disease) from activation of an old infection

Since TB prevention and control efforts directed at preventing TB transmission are fundamentally different from efforts to prevent activation of latent TB infection, genotyping offers a powerful tool to help direct the application of appropriate efforts. Furthermore, TB genotyping allows us to monitor our progress toward eliminating TB transmission more accurately.

What is the difference between a specimen and an isolate?

A specimen is a clinical sample (e.g., sputum, bronchial wash, urine, blood, or cerebrospinal fluid; tissues from organs or gastric aspirates) from a patient suspected of having TB. Specimens may or may not contain *M. tuberculosis*. If *M. tuberculosis* grows in culture media, it is called an isolate. Only isolates identified as *M. tuberculosis* can be genotyped.

The methods used by the Centers for Disease Control and Prevention (CDC) contract genotyping laboratories will not genotype *Mycobacterium avium* or other non-tuberculous mycobacteria.

What laboratory techniques and methods are used in the National TB Genotyping Service (NTGS)?

CDC contracts with two genotyping laboratories to provide the NTGS. These laboratories routinely use two polymerase chain reaction (PCR) methods on all *M. tuberculosis* isolates: spacer oligonucleotide typing (spoligotyping) and variable-number tandem repeat of mycobacterial interspersed repetitive units (VNTR-MIRU). These methods yield digital results that can be readily analyzed. A detailed description of these methods is available at: http://www.cdc.gov/tb/genotyping/Chap3/3_CDCLab_2Description.htm

- Spoligotyping: Identifies the *M. tuberculosis* genotype based on presence or absence of spacer sequences found in a direct-repeat region of the *M. tuberculosis* genome where 43 identical sequences and 36 base pairs are interspersed by spacer sequences. This highly reproducible method gives results in a standardized 15-digit code that can be easily analyzed and communicated between laboratories and TB programs (e.g., 777777777760771).

- VNTR-MIRU: Distinguishes the *M. tuberculosis* strains by the difference in the number of copies of tandem repeats at specific regions, or loci, of the *M. tuberculosis* genome. Like spoligotyping, this typing method yields results in a standardized code that can be easily analyzed and communicated between laboratories and TB programs. A total of 41 MIRU loci have been reported; however, most laboratories target only 12 loci. Newer versions of the method at the NTGS laboratories now include 24 loci, which may increase discriminatory power. When laboratories employ the 12-loci method, the results are a 12-digit code (e.g., 223225163324); 24-loci method results in a 24-digit code (e.g., 223225163324561333245623).

A third method, IS*6110*-based Restriction Fragment Length Polymorphism (IS*6110*-RFLP), may be utilized upon special request. Details on when to request RFLP is available at: http://www.cdc.gov/tb/genotyping/Chap5/5_Developing_3c_RFLP.htm

- IS*6110*-based RFLP: This method detects variations in a specific section of the *M. tuberculosis* genome called insertion element IS*6110*. The first step of RFLP is purification of DNA from an *M. tuberculosis* isolate. A restriction enzyme is added that cuts the DNA into hundreds of different fragments at specific sequences. The fragments are separated by size on an agarose gel and transferred to a membrane. A probe is then used to detect fragments containing IS*6110*, and the image is captured on film. Each copy of IS*6110* produces one band. IS*6110*-based RFLP patterns containing 7 or more bands provide more specificity in discriminating between isolates than do patterns with 6 or fewer bands.

What is a TB genotype cluster?

When two or more *M. tuberculosis* isolates match by genotyping methods (i.e., same spoligotype and MIRU patterns), they are referred to as a genotype cluster. Patients who are members of the same genotype cluster are assumed to have the same strain, which may be a surrogate for recent transmission. However, genotyping information is only one piece of evidence used to determine transmission patterns. Genotyping information, epidemiologic linkages including spatial (geography) and temporal (time) associations, and drug susceptibility results (phenotype) can help distinguish recent transmission from activation of latent TB infection.

NTGS laboratories assign unique numbers to clusters specifying each spoligotype and MIRU pattern combination. For each combination, a national and a state-specific designation are assigned. For example, all isolates within the United States with an identical spoligotype pattern such as 777776777760601 and MIRU pattern 224325153323 are assigned a national designation such as PCR00015. State-specific designations are also used for the same pattern. For example, in Arkansas, the state-specific designation for this pattern could be AR_010; whereas, in California it could be CA_084; in Colorado it could be CO_016; and in Nevada it could be NV_016. The national designations are designed to help facilitate inter-jurisdictional communication between TB controllers so they don't have to report 15- and 12- digit numbers; whereas, the state designations provide ease of use within the jurisdiction

A detailed description of cluster designations is available at: http://www.cdc.gov/tb/genotyping/Chap3/3_CDCLab_4Results.htm

How can TB genotyping information help identify unsuspected epidemiologic links?

Patients with TB disease caused by the same strain of *M. tuberculosis* will have matching genotype results. If an investigation establishes that two or more TB patients with matching genotypes share known epidemiologic linkages, this provides good evidence that these patients were involved in the same chain of recent transmission.

However, typically 20%–40% of genotype-matched TB cases are not identified as being connected by contact investigation. During traditional contact investigations, public health workers interview TB patients to elicit the names of other people who may have been exposed to them, and venues in which exposure may have taken place. TB patients are often unable or reluctant to name all contacts and all places of potential transmission. A careful review of public health records, genotyping results, contact investigation logs, estimated infectious periods, and re-interview of patients in a genotype cluster can uncover additional potential transmission opportunities.

How can TB genotyping help in source-case investigations for pediatric TB cases?

Pediatric cases are considered sentinel events in that they provide evidence of recent TB transmission. Source-case investigations can potentially identify additional TB cases, confirm suspected transmission links between a suspected source case and the child, or identify new, previously unsuspected, venues of transmission. When a suspected source case is identified, genotyping can support evidence of the transmission link to the child if the isolates from both have matching genotypes. If the investigation yields no suspected source case, a review of all genotyping results from patients living in the same region as the child may identify a genotype match. Also, further interviews of patients with matching genotypes may reveal the true source for the child's TB.

In patients with multiple episodes of TB, how will TB genotyping help distinguish between reactivation of an old infection or recent infection with a new strain?

Reactivation occurs when a patient with TB is treated and cured, but has a subsequent episode of TB disease caused by the same strain of *M. tuberculosis* as the previous episode. Reinfection (with development of disease) is caused by a second infection with a strain that usually has a different genotype from the strain that caused the initial episode of TB disease. Genotyping the initial isolate and the subsequent isolate from the same patient can distinguish between these two possibilities.

How can TB genotyping find or confirm false-positive cultures?

An estimated 2% of all *M. tuberculosis* cultures represent false-positive results. This can occur in even the most proficient laboratories for a variety of reasons (e.g., mislabeling, cross-contamination). A false-positive culture should be suspected when patients have a single culture confirmation, clinical presentation inconsistent with TB, or no clinical improvement despite adequate TB therapy. Public health officials must work closely with laboratorians and clinicians to gather information to verify or refute that suspicion.

A false-positive culture investigation is a multistep process that requires an investigation of the entire path a specimen and isolate take, from collection through the final laboratory report. This process enables the identification of possible common collection or processing points, which, at a given point in time, could have resulted in cross-contamination or mislabeling.

What is universal TB genotyping? Are we there yet?

Universal TB genotyping means submitting to the genotyping laboratories at least one isolate for every culture-confirmed TB case in the country. Universal TB genotyping will provide the best understanding of the epidemiology of TB transmission within a specific TB program's jurisdiction, as well as the entire country. It may uncover the greatest number of unrecognized outbreaks, clusters, and false-positive cultures. In 2004, when the NTGS was initiated, only 47% of all culture-positive TB cases in the United States had genotyping information. In 2007, this proportion had increased to 86%.

How can I receive and manage TB genotyping data for my patients?

The TB Genotyping Information Management System (TB-GIMS) will be launched in early 2009 to track and manage *M. tuberculosis* isolate genotyping data. TB-GIMS links genotyping results to the epidemiologic data from surveillance reports; queries and reports can be generated for comparing genotyping results locally and nationally. In the future, TB-GIMS will generate alerts or notifications of suspected recent transmission and help identify TB clusters that suggest public health action. A more detailed description of TB-GIMS is available at: http://www.cdc.gov/tb/pubs/tbfactsheets/gims.htm.

How much does TB genotyping cost patients or health care providers?

TB genotyping is provided at no cost to patients, healthcare providers, and health departments. CDC also pays for shipping of isolates for genotyping from local laboratories to the genotyping laboratories. Shipping labels are provided from CDC to facilitate the shipping process.

Additional Information

For more information on TB genotyping, please contact tbgenotyping@cdc.gov