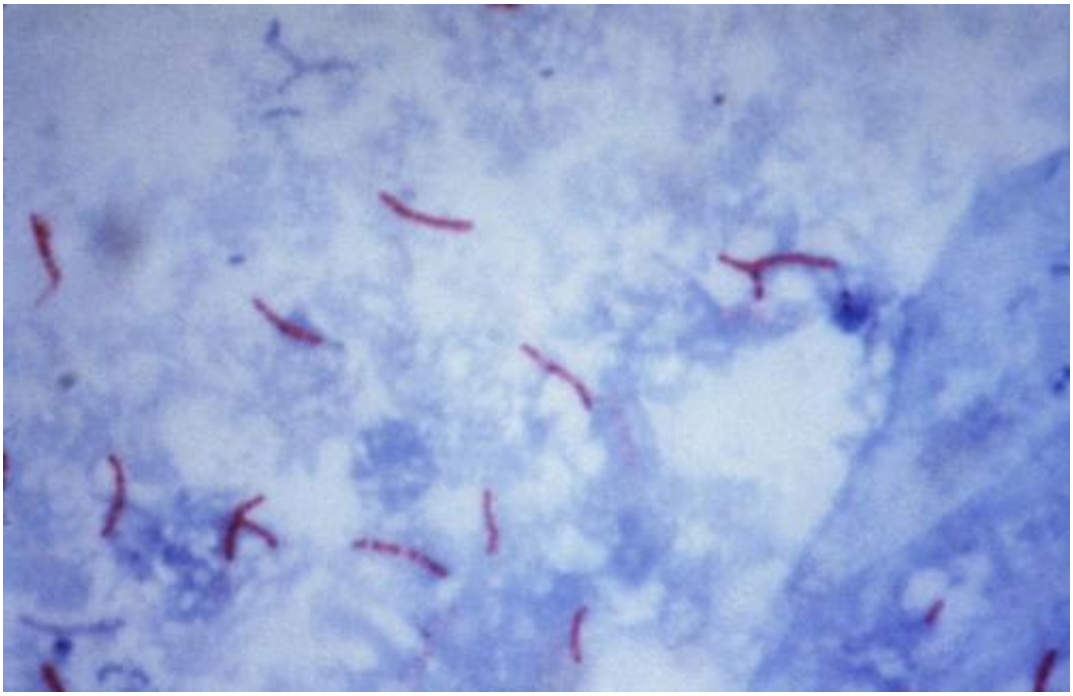
What role do genome variations play in tuberculosis?

September 7 2017



This photomicrograph reveals Mycobacterium tuberculosis bacteria using acid-fast Ziehl-Neelsen stain; Magnified 1000 X. The acid-fast stains depend on the ability of mycobacteria to retain dye when treated with mineral acid or an acid-alcohol solution such as the Ziehl-Neelsen, or the Kinyoun stains that are carbolfuchsin methods specific for M. tuberculosis. Credit: public domain

Tuberculosis is caused by *Mycobacterium tuberculosis*. In 2016, this was the most common causative pathogen for death by infectious diseases. Therefore, investigating the biology of infection and disease

development is important in the quest to end tuberculosis. In this study, the authors conducted an integrative analysis of human and pathogen genome variations in tuberculosis. The study is published in Springer Nature's *Journal of Human Genetics*.

Genome-wide association study (GWAS) is an analysis method in genetics which is used to identify disease-related genome variations by comparing the differences in genotype frequencies in a case and control group. There have been several GWAS regarding susceptibility to tuberculosis. However, previous GWAS have not been based on analysis that accounted for pathogen variations. The possible interplay between the host and pathogen genomic variations is difficult to analyze because both the infected human genome and infecting pathogen genome need to be collected from large numbers of patients. In this research the authors did a systematic exploration of host variations for their association with specific lineages of *Mycobacterium tuberculosis*, which share the same pathogen genome variations.

This is the first GWAS report identifying the host genetic association with tuberculosis after stratification by pathogen variations. Genotype frequency of a single nucleotide polymorphism (SNP) was increased in a group of patients infected by specific lineage of *Mycobacterium tuberculosis* compared with the healthy controls. The increased frequency of the SNP was not observed in a group of patients infected by the other lineage, suggesting the pathogen lineage-specific risk of this human genome variant and importance of analyzing the interaction between the host and pathogen genome variants. The identified SNP locates near CD53 gene, encoding a leukocyte surface glycoprotein and known to have functions in immunity and stress response. Increased gene expression of CD53 was observed in active TB patients and supports its biological roles in susceptibility to tuberculosis.

There have been six major lineages reported worldwide (Gagneux S. et

al., PNAS 2006). The authors have already reported a strain-dependent association of HLA class II genes in tuberculosis (Toyo-oka L et al., HLA 2017), thus heterogeneity of the pathogen genome may be responsible for the inconsistency of previous genetic association studies for tuberculosis.

"Nearly one third of the world's population are infected by *Mycobacterium tuberculosis*, but only 10% of them show symptoms of TB throughout their life." explains lead author Yosuke Omae. "This research lays the foundations for us to identify who will get TB."

More information: Omae Y et al. (2017) Pathogen lineage-based genome-wide association study identified CD53 as susceptible locus in tuberculosis, *Journal of Human Genetics*, DOI: 10.1038/jhg.2017.82

Provided by Springer