

How to use SpolSimilaritySearch

This is a user manual helping researchers how to navigate and search similarities as compared to SITVIT2 database with SpolSimilaritySearch tool.

Keywords: *Mycobacterium tuberculosis*, Database, Spoligotyping, Evolution, Similarity.

Home

Homepage (**Figure 1**) briefly introduces the SpolSimilaritySearch Web tool which is dedicated to study similarities of all spoligotyping patterns recorded in the SITVIT2 database. The main functionalities of the Website are shown in figures below.

The image shows the homepage of the SpolSimilaritySearch tool. At the top, the title 'SpolSimilaritySearch' is displayed in a large, stylized font. Below the title is a navigation bar with five items: 'Home', 'Select the 43 spacers of your spoligotype', 'Geographical Distribution', 'Resources', and 'Contact'. Four callout boxes with arrows point to these navigation items, providing brief descriptions of their functions:

- Home:** Tool allowing Users to Compare Similarity of spoligotypes
- Select the 43 spacers of your spoligotype:** Tool allowing Users to visualize data through interactive maps
- Resources:** List of some resources and links of interest
- Contact:** Contact the authors and virtually visit our Team!

Below the navigation bar, the main heading reads: **Search and Compare Similar spoligotyping patterns contained in the SITVIT2 database**. The main text describes the tool's purpose and capabilities:

THIS WEB TOOL ALLOWS USER TO SEARCH FOR SIMILAR SPOLIGOTYPING PROFILES IN FUNCTION OF A GIVEN QUERY PROVIDED BY THE USER. SPOLIGOTYPING METHOD IS ONE OF THE MOST COMMONLY USED METHODS PERMITTING THE IDENTIFICATION OF MEMBERS OF *MYCOBACTERIUM TUBERCULOSIS* COMPLEX (MTBC). DESPITE KNOWN LIMITATIONS, THIS METHOD ALLOWED A PROBABLE IDENTIFICATION OF MORE THAN 100,000 ISOLATES (COLLECTED FROM 169 COUNTRIES OF ORIGIN OF PATIENT) IN THE SITVIT2 PROPRIETARY DATABASE OF THE INSTITUT PASTEUR DE LA GUADELOUPE. SPOLSIMILARITYSEARCH TOOL INCORPORATES A SIMILARITY SEARCH ALGORITHM ALLOWING USERS TO GET A COMPLETE OVERVIEW OF SIMILAR SPOLIGOTYPE PATTERNS (CONTAINING INFORMATION ON 43 SPACERS THAT COULD BE ABSENT OR PRESENT) EXISTING IN OUR DATABASE, AND ACCORDING TO THEIR SEARCH CRITERIA.

FOR INFORMATION ON HOW TO USE THIS TOOL PLEASE [README!](#)

At the bottom of the page, a copyright notice states: Copyright © 2016, Institut Pasteur de la Guadeloupe. All rights reserved.

Figure 1: Homepage and brief description of the main functionality.

Select the 43 spacers of your spoligotype pattern

This section is the main functionality of this tool. It allows users to select each of the 43 spoligotypes separately in order to catch, thanks to Regular Expression or REGEXP (http://en.wikipedia.org/wiki/Regular_expression), the most similar patterns (in function of the query conveyed) contained within SITVIT2 database. A spacer value can be either an absence (□), a presence (■), or an undetermined or doubtful value (-).

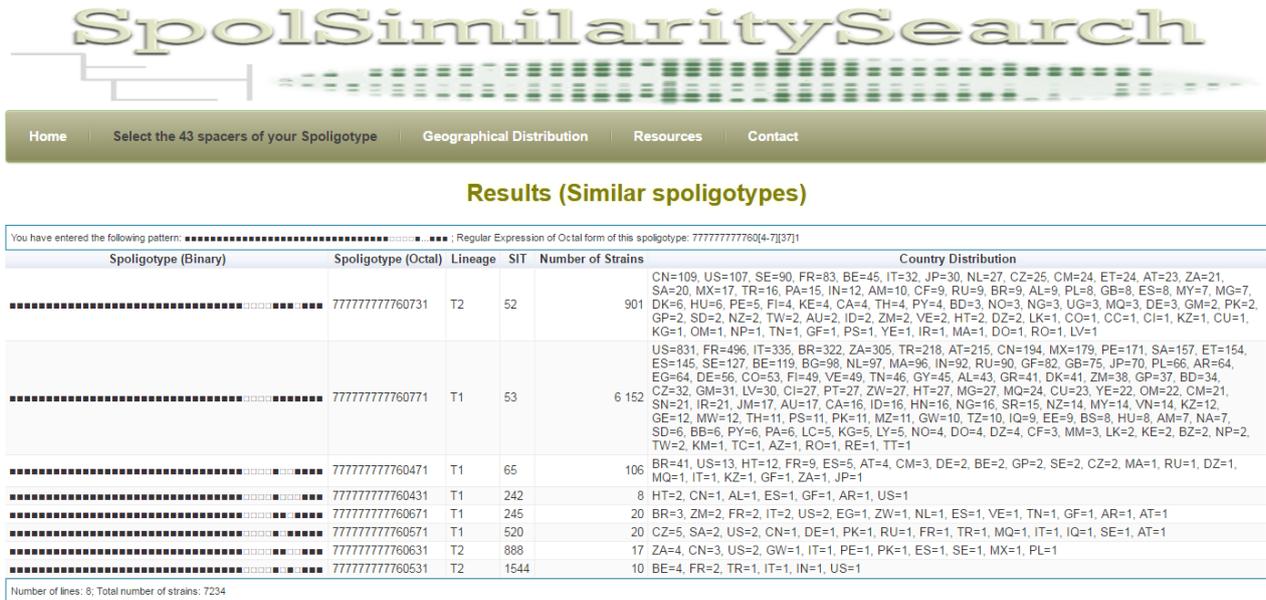


Figure 3: Snapshot of 'Results' obtained in function of our selection or query.

Note that in this Result section dots (.) are equivalent to dashes (-) in the Binary form of spoligotype.

Geographical Distribution

This section allows users to visualize geographical distribution of main lineages and spoligotype patterns (in octal format) through interactive maps.

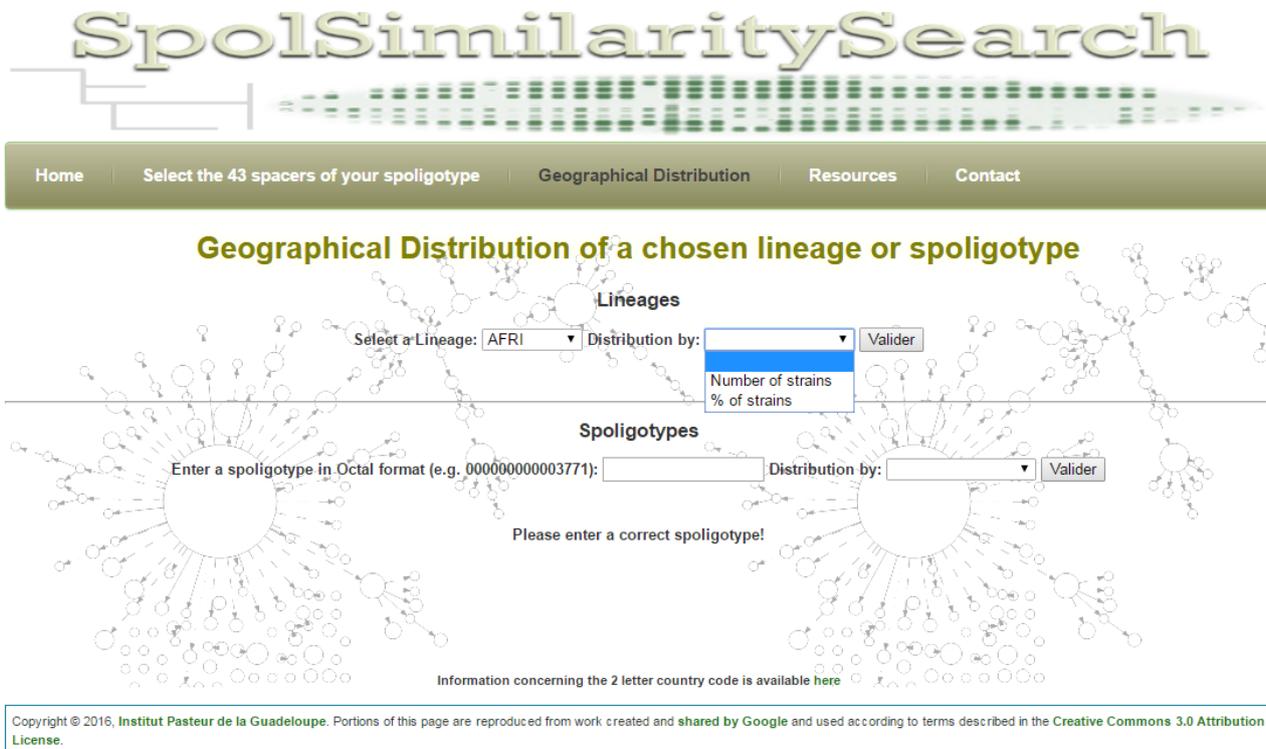


Figure 4: Snapshot of 'Geographical Distribution' page.

The user could visualize distribution by number of strains (i.e. number of strains by country, given the selection), or by percentage (%) of strains (i.e. number of strains given the selection divided by total number of strains according to SITVIT2 by country, and multiplied by 100). When no value of distribution is selected, the default value will be percentage (% of strains).

In the following example, we can visualize distribution (by number of strains) of all isolates belonging to AFRI (*M. africanum*) lineage:

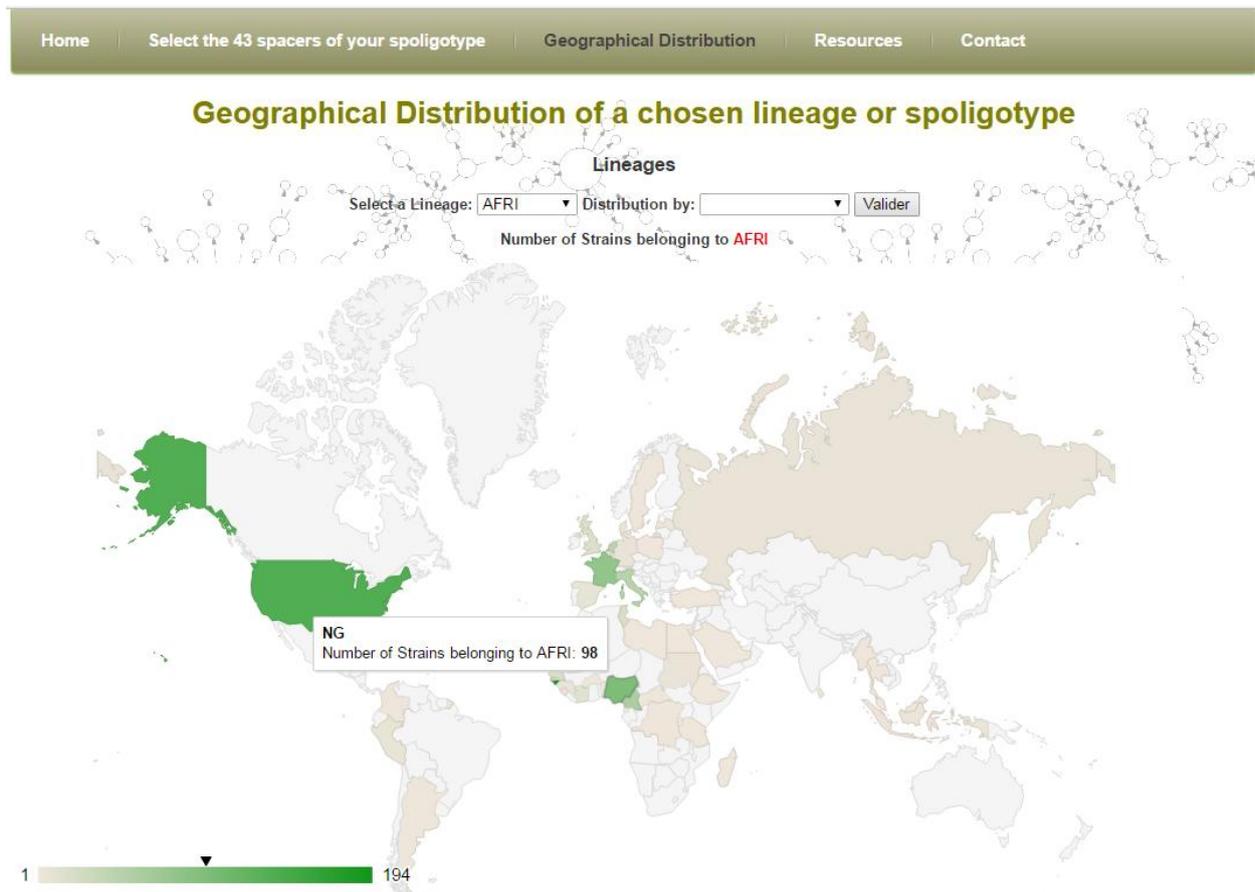


Figure 5 : Global distribution (by number) of strains belonging to AFRI lineage

We can notice that 98 strains belonging to AFRI are located in Nigeria (NG).

In this other example, we can visualize the distribution (by percentage) of strains belonging to spoligotyping pattern 000000000003771 (octal format corresponding to SIT 1 / Beijing):

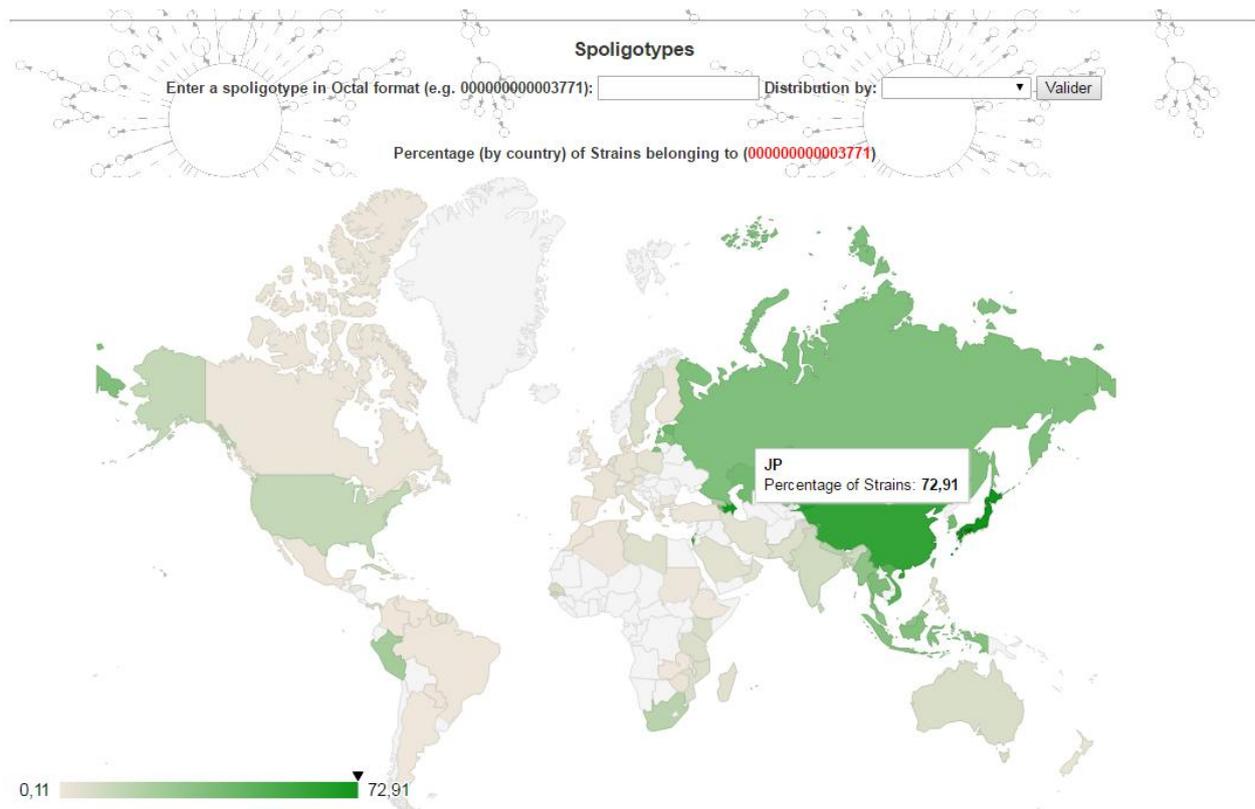


Figure 6: Global distribution (by percentage) of strains belonging to spoligotyping pattern 000000000003771

We can notice that 72.91% of strains belonging to spoligotyping pattern “000000000003771” are located in Japan (JP).

Resources

This section provides a set of other Web resources or links related to analysis of *Mycobacterium tuberculosis* complex (MTBC).

Contact

This section allows to contact authors and supervisor of this tool, and briefly introduces the team working at the lab: “Unité de la Tuberculose et des Mycobactéries, Institut Pasteur de la Guadeloupe”. Users can also subscribe to our Twitter feed, and follow our Facebook page in this section.