

C C A T T 0 1 0 0 0  
G A G G A 0 1 1 0 1  
G A A T T 0 0 1 1 0  
A C A A G 0 0 1 0 0  
T A C C A 0 0 1 1 0  
T T A C A 0 1 0 0 0  
A C C T C 0 0 0 1 0  
A A G G A 0 0 0 0 0  
G A T G A 0 1 1 0 0  
T A G A T 0 0 1 0 0  
G A T G A 1 0 1 0 0  
T G T A G 1 0 0 0 0  
T A G T A 0 0 0 0 0  
G A T A T 1 0 0 0 0  
G A G T G 1 1 0 0 0  
A G A T T 1 1 0 0 0  
G A G T A 1 1 0 0 0  
T G A T G 1 1 0 0 0  
A T T A G 1 1 0 0 0  
T A G A T 1 1 0 0 0  
G A G A 1 1 0 0 0  
G T A 1 1 0 0 0  
G A T 1 1 0 0 0  
T A G 1 1 0 0 0  
A G 1 1 0 0 0  
G A 1 1 0 0 0  
A 1 1 0 0 0

# Best Practice

## Monitoring the impact of vaccination against Streptococcus Pneumonia on a national scale

May 7, 2008





## **Monitoring the impact of vaccination against Streptococcus Pneumonia on a national scale**

This paper describes how the deployment of CLC bio's MLST Module has speeded up the process of analyzing large amounts of sequencing data, and how it has been used by the German National Reference Center for Streptococci to open up new research possibilities.

The data analysis required for Multilocus Sequence Typing used to be extremely tedious and time-consuming, resulting in an increasingly high error rate when doing this task for repeated hours on end. Researchers simply got tired and lost focus by doing the same routine task over and over again. A highly skilled researcher could do three complete typings in one hour, but with the CLC MLST Module, one typing only takes a minute. This is a dramatic decrease in time spent and opens up new possibilities to how the Reference Center can employ MLST on a much wider scale and even have additional time for more in-depth research.

### **Introduction**

The German National Reference Center for Streptococci is responsible for the epidemiological monitoring of Streptococcus in Germany. The center monitors the impact of vaccination against Streptococcus Pneumonia, and one of the methods employed is Multi Locus Sequence Typing (MLST).

MLST is used to identify bacterial strains. In the case of Streptococcus, it involves sequencing seven housekeeping genes which are compared against a database. Based on the variation of these seven genes, the bacterial isolate is assigned a sequence type.

Because MLST is based on sequencing, it is straight-forward to compare results, and the method is thus more portable than other diagnostic methods.

When sequencing data are compared to the database, it has to be trimmed, assembled and aligned. This is manual work, and it has to be done for 14 sequences for each isolate (seven genes with a forward and a reverse sequencing read). The mental load on the researcher is heavy because even small errors can lead to completely misleading results.

### **Results**

The CLC MLST Module is an extension of the CLC Combined Workbench which provides a graphical interface to the sequencing data analysis. The MLST Module is specifically designed to support the MLST workflow, taking away the tedious parts of the work and presenting the data for interpretation and inspection by the researcher.

The Reference Center has integrated this technology into their workflow and reaped the benefits in terms of valuable research time (see figure 1 for an illustration of the work flow before and after).

### **Return on investment (ROI)**

With an initial investment in technology and training for a 4-figured amount in Euros, the center has saved time corresponding to about half a position. This means the return on investment is 3-400% already within the first year. The following years will only require minor maintenance fees to keep the software updated, so the benefits are even greater then.

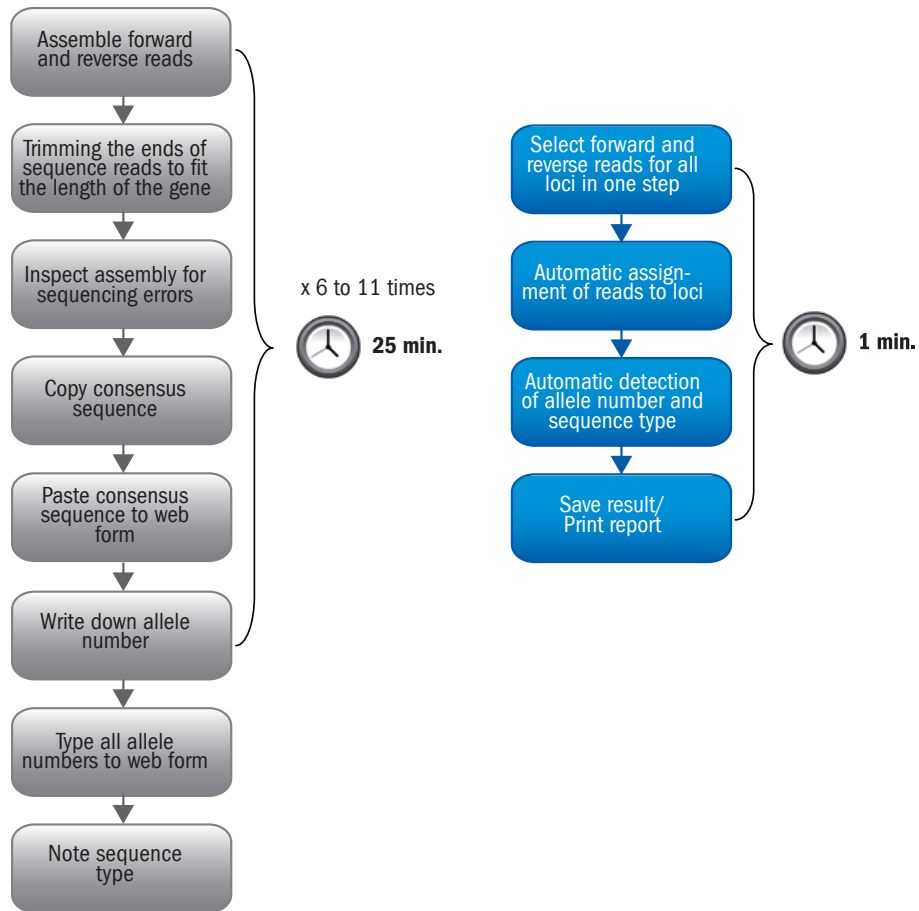


Figure 1: The MLST work flow.

Because the analysis of the sequencing data requires educated personnel, it previously took valuable time from other activities in the center. By integrating the CLC MLST Module into the workflow, new possibilities open up for using MLST. A project involving the typing of 700 isolates has become feasible because the data analysis is now faster and more precise. Thus, the benefits are both financial and in terms of research.

### Conclusions

Many laboratories in the World are involved in the kind of projects conducted at the German National Reference Center for Streptococci, and a lot of them have started to use MLST because of its many benefits. The penetration of the MLST method is however slowed down by the data analysis bottleneck. The German National Reference Center for Streptococci has shown an example of how using innovative technology to analyze the data removes the traditional data analysis obstacles for MLST.



### **Creative Commons License**

All CLC bio's scientific articles are licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License. You are free to copy, distribute, display, and use the work for educational purposes, under the following conditions: You must attribute the work in its original form and "CLC bio" has to be clearly labeled as author and provider of the work. You may not use this work for commercial purposes. You may not alter, transform, or build upon this work.

See <http://creativecommons.org/licenses/by-nc-nd/2.5/> for more about how you may use the contents.

