

# TIALA - Visual Time Series Alignment Analysis

Günter Jäger<sup>\*1</sup>, Florian Battke<sup>\*2</sup>, Kay Nieselt<sup>\*3</sup>

**Keywords:** Multiple time series alignment, visual analytics, expression analysis, microarrays

## Introduction

The analysis of time series expression data is widely employed for investigating biological mechanisms. In order to successfully compare different time series it is necessary to perform a time series alignment because the experiments can differ in the number of time points, as well as in the time points themselves. We propose a novel visual analytics approach for the analysis of multiple time series experiments in parallel. Our time series alignment analysis tool TIALA [2] allows one to align multiple time series experiments and to visually explore the aligned expression profiles.

## Materials & Methods

TIALA calculates multiple time series alignments either by using user defined time shifts between the provided time series experiments or via applying several statistical methods such as *Euclidean Distance* or our *Time Shift Suggestion Function*, to automatically determine pairwise time shifts. Visualization of the unaligned data sets as well as of the resulting multiple alignment is achieved via several plots that are linked with each other, i.e. objects selected in one visualization are also highlighted in all the other plots. To enhance the display of multiple aligned time series expression profiles, a two- and three-dimensional visualization strategy was implemented. TIALA to our knowledge is the first and only tool for comparing two or more time series of expression data both analytically as well as visually. The large user interactivity and the tight integration into the Mayday framework [1], one of the most feature-rich microarray analysis platforms [3], makes TIALA a powerful visual analytics approach for large scale time series data.

## Results

To exemplify the application of the time series alignment we use two highly resolved data sets from a large-scale transcriptomic study of the soil bacterium *Streptomyces coelicolor* grown under phosphate limited conditions [4]. Two replicates were measured (named F199 and F202 in the original study) and focus is set on the cluster of genes upregulated by phosphate depletion. Comparing this cluster in TIALA reveals a time shift of 1h between the two data sets, shown in figure 1A. This time shift (which is also detected by TIALA's automatic 'time shift suggestion' function) results from technical variations during the fermentation and is not the result of differences in the gene expression of the two replicates. However, after correction for the systematic shift, the expression profiles show a near perfect correlation (figure 1B). The aligned data sets can now be used to study the biological phenomena associated with phosphate depletion (figure 1C). Our application study demonstrates how TIALA as a visual analytical tool for the comparison of time series data can improve the analytical process.

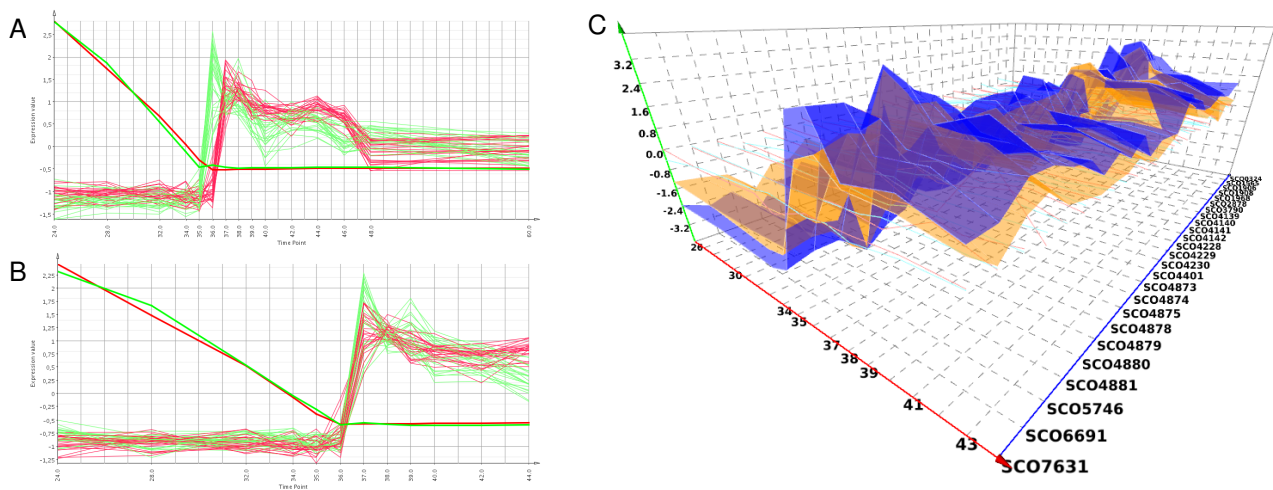


Figure 1: Comparison of the cluster of genes upregulated by phosphate depletion from two *S. coelicolor* replicates F199 (red) and F202 (green). A, the initial alignment clearly shows a time shift of 1h; B, After correction of the time shift, the expression profiles and extracellular phosphate concentration (strong lines) are perfectly correlated. C, 3D view of the aligned time series, identifying the individual genes in this cluster (z-axis).

## References

- [1] F. Battke, S. Symons, and K. Nieselt. Mayday – Integrative analytics for expression data. *BMC Bioinformatics*, 11(1):121, 2010.
- [2] G. Jäger, F. Battke, and K. Nieselt. TIALA – Time Series Alignment Analysis. In *proceedings of BioVis 2011: 1st IEEE Symposium on biological data visualization*, 2011. (to appear)
- [3] A. Koschmieder et al. Tools for managing and analyzing microarray data. *Briefings in Bioinformatics*, 2011.
- [4] K. Nieselt et al. The dynamic architecture of the metabolic switch in *Streptomyces coelicolor*. *BMC Genomics*, 11:10, 2010.

<sup>1</sup>guenter.jaeger@uni-tuebingen.de <sup>2</sup>florian.battke@uni-tuebingen.de <sup>3</sup>kay.nieselt@uni-tuebingen.de

<sup>\*</sup>Integrative Transcriptomics, Center for Bioinformatics Tübingen, University of Tübingen, Sand 14, 72076 Tübingen