A Quick Guide to CellTrans

Thomas Buder

2017-01-12

What is CellTrans?

CellTrans is an R package to quantify stochastic cell state transitions from appropriately prepared data from FACS and flow cytometry experiments (Buder et al. 2017).

CellTrans can be utilized to (i) infer the transition probabilities between different cell states, (ii) predict cell line compositions at a certain time, (iii) predict equilibrium cell state compositions, and, (iv) estimate the time needed to reach this equilibrium.

Installing CellTrans

CellTrans is freely available via GitHub, see my GitHub page. In order to install it, directly type within R

```r
if (!require("devtools")) install.packages("devtools")
library("devtools")
install_github("tbuder/CellTrans")
```

Note that you need the `devtools` package in order to install packages from GitHub and that it might be necessary to adjust your connection settings if you use a proxy server.

Then, load the package with

```r
library("CellTrans")
```

A comprehensive example

To demonstrate how CellTrans is used directly within R, we utilize FACS data from experiments with two different cancer cell subpopulations in colon cancer (Yang et al. 2012). The data describe the evolution of the proportion of cancer stem cells (CSC) and non-cancer stem cells (NCSC) during cell line experiments in time. Two experiments have been performed with pure initial cell states in the beginning and the evolution of the proportions has been measured each 2 days for 24 days.

Data preparation

In order to analyze data from FACS and flow cytometry experiments in CellTrans, they have to be appropriately prepared in text files. These text files contain line-wise the cell state proportions after the corresponding time of the two cell line experiments, see here. For example, the content of `W6.txt` is

```
0.7323 0.2677
0.4740 0.5260
```

meaning that the experiment with pure CSC cells in the beginning evolved to 73.23% CSC and 26.77% NCSC cells after 6 days and the experiment with pure NCSC cells evolved to 47.4% CSC and 52.6% NCSC, respectively.

Reading the data and configuring CellTrans

All relevant information for the analysis and visualization of the data with CellTrans can be read by utilizing the function `readExperimentalData()`. it is recommended to store these data in a variable, e.g.

```
my_CellTransdata<-readExperimentalData()
```

This command opens a sequence of windows asking for the details of the experimental setup, see Figures 1-6.

Figure 1: In our example, we have two different cell states.

Figure 2: A name can be assigned to each cell state.

Figure 3: The unit of time of the experiments has been *days*.

Figure 4: There are 12 time points of measurement...

Figure 5: ... which are specified in the next step: day 2, 4,...,24.

Figure 6: The initial cell state distributions of the experiments in the example are purified cell lines. If the experimental setup is different, an individual matrix containing row-wise the initial cell state compositions can be provided.
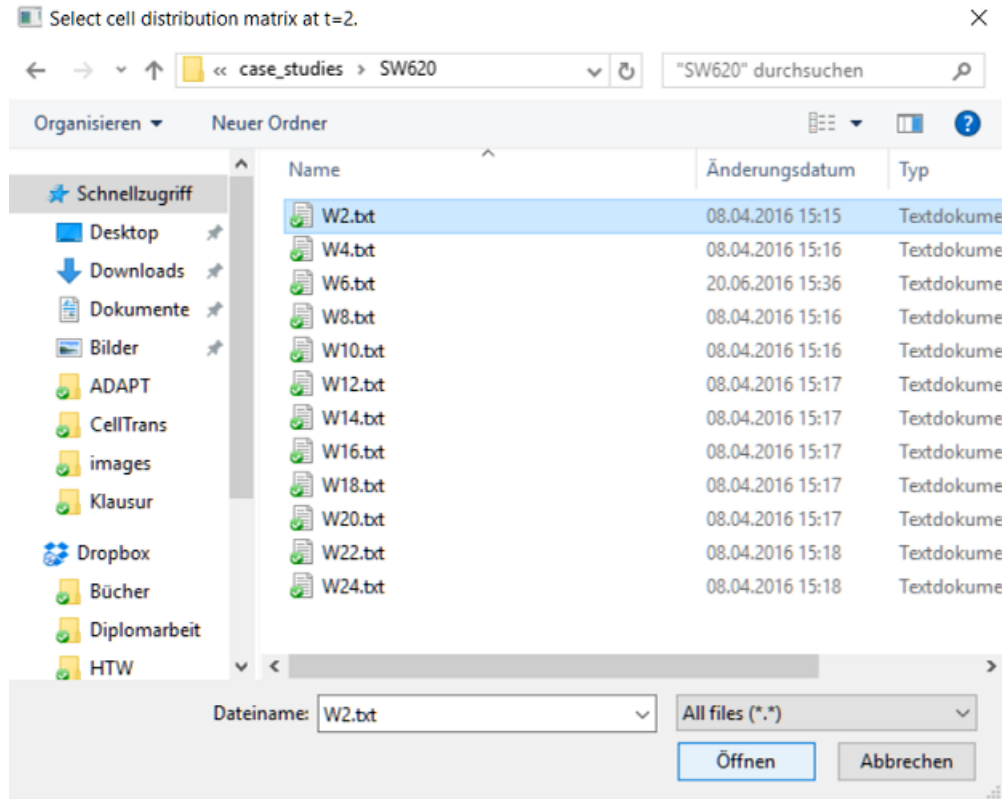
Figure 7: Finally, the experimental data matrices corresponding to the time points have to be selected.

Now, all relevant data for the analysis are stored in `my_CellTransdata` which can be used for the analysis and visualization of the data.

Deriving cell state transition probabilitites

The function `celltransitions` prints an estimate of the transition probabilities between the cell states and an estimation of the equilibrium composition. Applying this function to the read data, i.e.

```
celltransitions(my_CellTransdata)
```

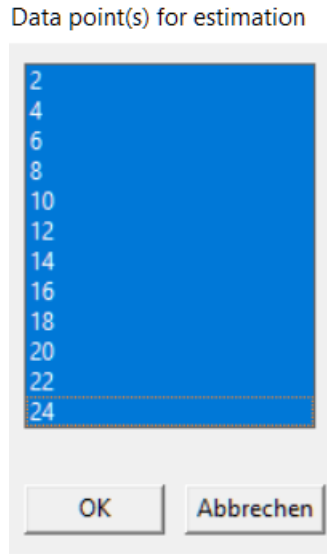allows to select the data points considered for the estimation.

Figure 8: Here, we consider all data points for the estimation.

CellTrans prints the estimated transition probabilities between both cell states and the predicted equilibrium distribution.

```
[1] "Results of CellTrans"
[1] "###############################"
[1] "used timepoints:"
 [1] "2"  "4"  "6"  "8"  "10" "12" "14" "16" "18" "20" "22" "24"
Markov Chain
 A  2 - dimensional discrete Markov Chain defined by the following states:
 Cancer Stem Cells, Non-Cancer Stem Cells
 The transition matrix  (by rows)  is defined as follows:
                    Cancer Stem Cells Non-Cancer Stem Cells
Cancer Stem Cells             0.9455135            0.05448652
Non-Cancer Stem Cells         0.1029900            0.89701002


[1] "predicted equilibrium distribution"
     Cancer Stem Cells Non-Cancer Stem Cells
[1,]         0.6540022             0.3459978
[1] "#########################################"
>
```

Estimating time to equilibrium

In order to estimate the time needed from a specific initial cell state composition to the equilibrium composition, the function

```
timeToEquilibrium(input, initialDistribution, tol)
```

can be utilized. The first parameter represents the read data from the function `readExperimentalData()` and the second parameter describes the initial cell state proportion. This initial cell state composition has to be described as a vector in R. The third parameter `tol` describes the maximum allowed deviation of the proportion of each cell state from the equilibrium distribution.

Let's go the example data stored in `my_CellTransdata` and consider 50% of CSC and 50% of NCSC as initial

composition. This composition can be described in R as vector `c(0.5, 0.5)`. Setting a tolerance deviation of 0.01 yields an expected time of 16 days to equilibrium:

```
> timeToEquilibrium(my_CellTransdata,c(0.5,0.5),0.01)
[1] "Expected time to equilibrium: 16 days"
[1] "Predicted cell state proportions at this time:"
          [,1]       [,2]
[1,] 0.6440747 0.3559253
>
```

Creating Plots

Another feature of CellTrans is the creation of plots of your data and the corresponding predictions. The function

```
celltrans_plotPDF(input)
```

allows to choose the data points conidered for the estimation and asks whether the experimentally measured equilibrium should also be plotted. If yes, a text file containing your experimentally measured equilibrium proportions has to be selected and is then plotted at a time point of your choice.
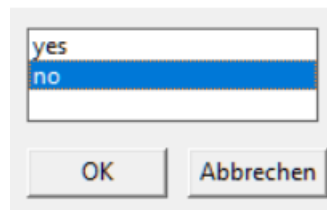

Figure 9: Here, the experimental equilibrium shall not be plotted.

Subsequently, the experiments for which a plot shall be created can be selected. For each selected experiment one plot is created with the corresponding evolution of the cell state proportions and the predicted evolution estimated by CellTrans. Furthermore, you can choose which cell states shall be included in the plot.
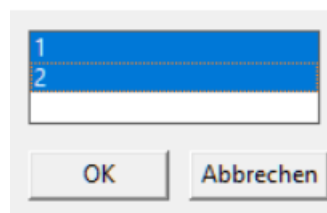

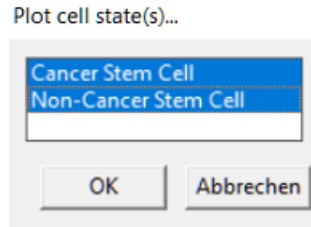Figure 10: This selection creates a plot for both cell line experiments....

Figure 11: ... containing the data and predictions for both cell states.

Then, you can choose the range of the y-axis and select a directory where the plots are saved.
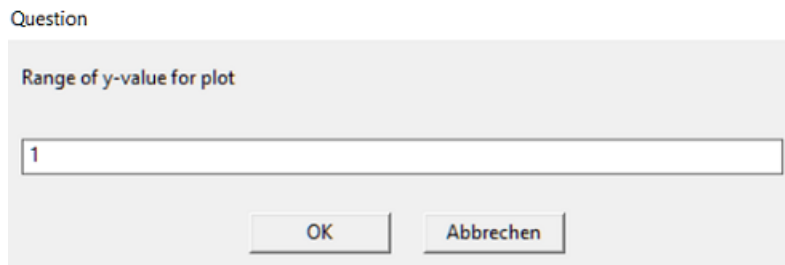

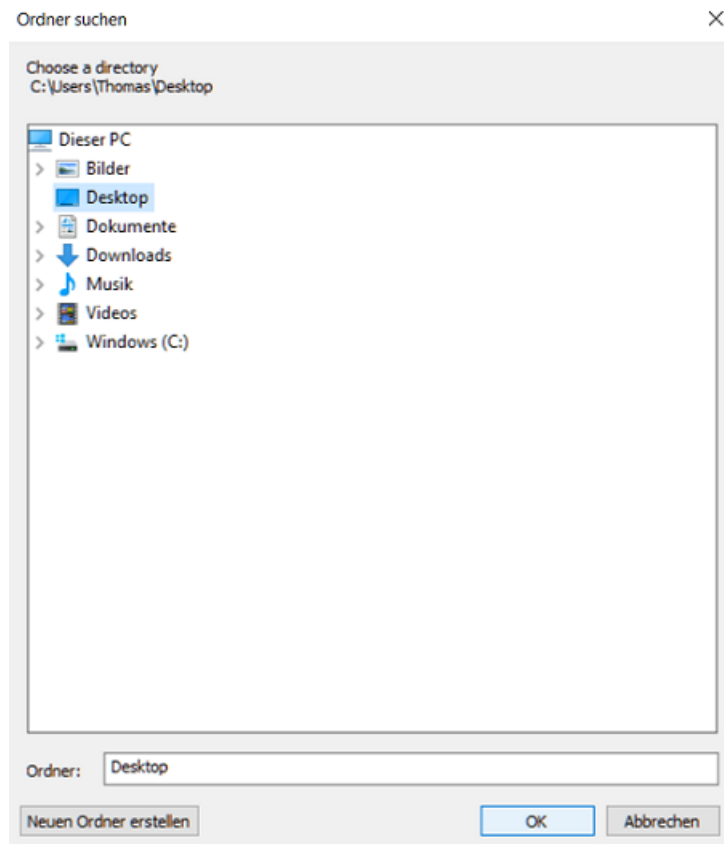
Figure 12: We select a range of 1 for the y-axis.

Figure 13: Here, the pdf files are saved on the desktop.

Finally, two pdfs are created with the data of both experiments and the corresponding predictions of CellTrans using all data points.
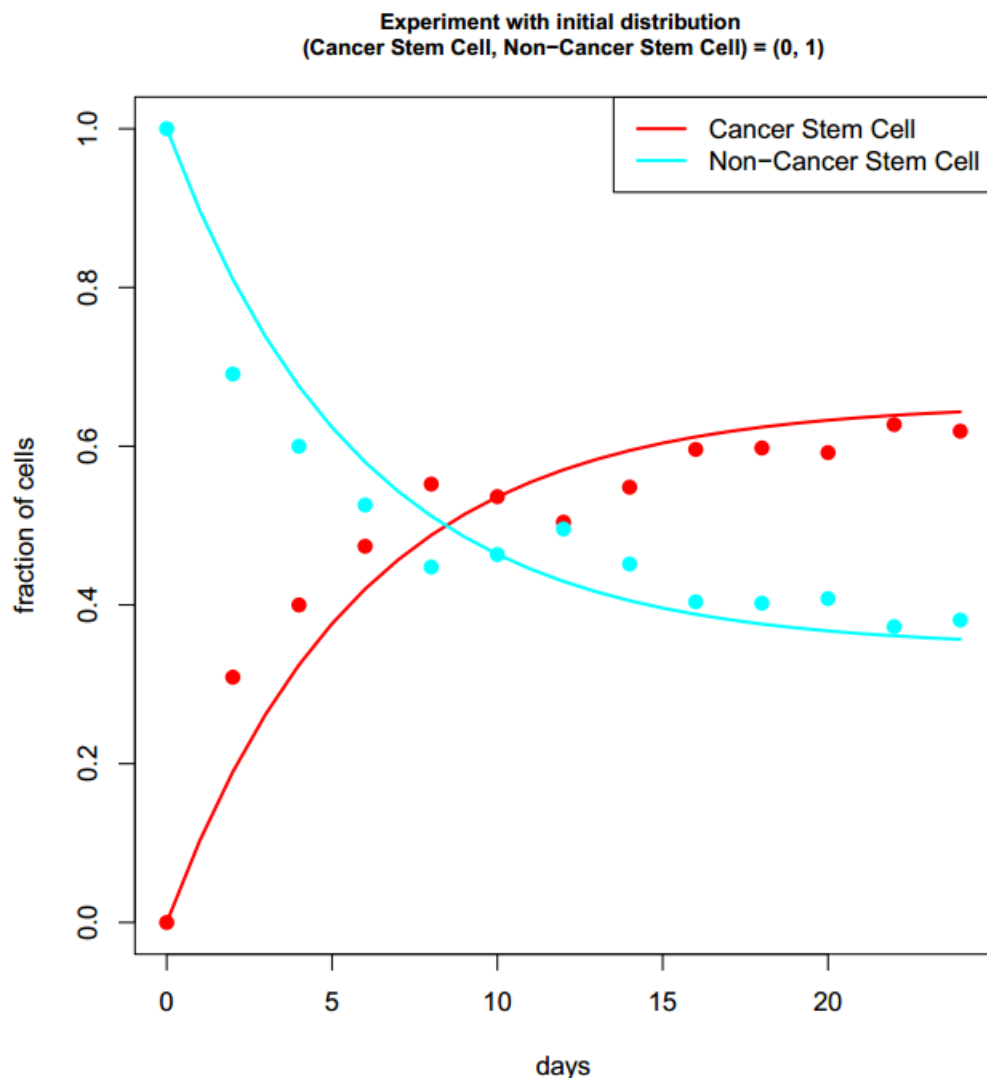

Figure 14: The plot for experiment 2 and both cell types is saved as `plot_2.pdf`.

The function `celltrans_plot()` offers the same options, but directly produces a plot in R without creating a pdf.

References

Buder, T, A Deutsch, M Seifert, and A Voss-Böhme. 2017. "CellTrans: An R Package to Quantify Stochastic Cell State Transitions." *Submitted*.

Yang, G, Y Quan, W Wang, Q Fu, J Wu, T Mei, J Li, et al. 2012. "Dynamic Equilibrium Between Cancer Stem Cells and Non-Stem Cancer Cells in Human Sw620 and Mcf-7 Cancer Cell Populations." *British Journal of Cancer* 106 (9). Nature Publishing Group: 1512–9.