**Reproducibility Checklist**

## Data

**Abstract**

We have included the real data set in the file "data.Rdata”. It can be loaded via load(“data.Rdata") The data set is described in Section 8 of the main paper. It is in the format of a 747x2 matrix, where

* **Rows**: cells (statistical units);
* **Columns**: named “MDS1”, “MDS2” represent the two biomarkers describing the geniting expressions.

|  |  |  |
| --- | --- | --- |
| i\p | MDS1 | MDS2 |
| OEP01\_N706\_S501 | -0.7139231 | -0.94959181 |
| OEP01\_N701\_S501 | -0.6274044 | -0.09495844 |
| OEP01\_N707\_S507 | -1.0055568 | -0.27521693 |
| OEP01\_N705\_S501 | -0.8102099 | 0.19664808 |
| OEP01\_N709\_S501 | 0.1190077 | -0.03042750 |
| OEP01\_N702\_S505 | -1.1606041 | -0.99415583 |

In the table above, we display the first six rows of the dataset. They can be shown in R via head(Data).

## Code

**Abstract**

The codes are written in R.

The codes comprise two main programs:

1. a) **GARP\_main.R**: This is the main script that runs the analysis by calling the functions included in the other files. This script produces the MCMC samples of the single-cell RNA data analysis and reproduces the results of the analysis summarized in Section 8 of the main.
2. b) **GARP\_fcts.R:** This file contains all the R functions needed to run the main script, including the MCMC function to implement the sampler described in Section 6 of the main manuscript.

**Description**

The codes are included in a zipped file. The main file to run is **GARP\_main.R**. The main file calls functions that are also included in the zipped file. The implementation is highly automated - the main function implementing the GARP model takes in the matrix (as described above) as an argument and a few additional parameters.

The MCMC algorithm takes around 28 minutes on a Lenovo machine with 32 Gb RAM. Alternatively, the user can load the results from a previous run of the algorithm (instructions in the body of the R script). Additional descriptions and instructions are included as detailed comments in the body of the R scripts.

**Optional Information**

The following R libraries are utilized and have to be pre-installed:

* dplyr 1.0.8
* ggplot2 3.3.6
* reshape2 1.4.4
* Rfast 2.0.6
* multiway 1.0-6
* tidyr 1.2.0
* factor.switching 1.3

## Instructions for Use

**Reproducibility**

The file **GARP\_main.R** reproduces the results presented in Section 1 and Section. On top of every plot command in the main R script, there is a header describing where that plot appears in the manuscript.