

# fuchikoma: Revealing differentially expressed genes using nonlinear dimensionality reduction for single cell RNA-Seq data

Koki Tsuyuzaki<sup>1</sup>, Haruka Ozaki<sup>1</sup>, Mika Yoshimura<sup>1</sup>, and Itoshi Nikaido<sup>1</sup>.

December 14, 2015

<sup>1</sup>Bioinformatics Research Unit, Advanced Center for Computing and Communication, RIKEN.

k.t.the-answer@hotmail.co.jp

## Contents

---

<b>1</b>	<b>Introduction</b>	<b>1</b>
<b>2</b>	<b>Getting started</b>	<b>2</b>
2.1	Package loading . . . . .	2
2.2	Test data . . . . .	2
2.3	Effectivity of non-linear dimentional reduction . . . . .	2
2.4	Limitation of non-linear dimentional reduction . . . . .	5
2.5	How to get help . . . . .	5
<b>3</b>	<b>How to use fuchikoma</b>	<b>5</b>
3.1	fuchikoma : Supervised mode (ANOVA-like) . . . . .	5
3.2	fuchikoma : Supervised mode (cluster specific DEGs) . . . . .	5
3.3	fuchikoma : Unsupervised mode . . . . .	6
<b>4</b>	<b>Setup</b>	<b>6</b>

## 1 Introduction

---

This document provides the way to perform *fuchikoma* package. *fuchikoma* is the method for detection of differentially expressed genes (DEGs) within single-cell RNA-Seq (scRNA-Seq) dataset.

## 2 Getting started

---

### 2.1 Package loading

To install the *fuchikoma* package into your R environment, start R and enter.

```
> # source("http://bioconductor.org/biocLite.R")
> # biocLite("fuchikoma")
```

And then load the *fuchikoma* package.

```
> library("fuchikoma")
```

By `ls` function, we can confirm three functions are loaded from *fuchikoma* package.

```
> ls("package:fuchikoma")
[1] "CatKernel" "fuchikoma" "HSIC"
```

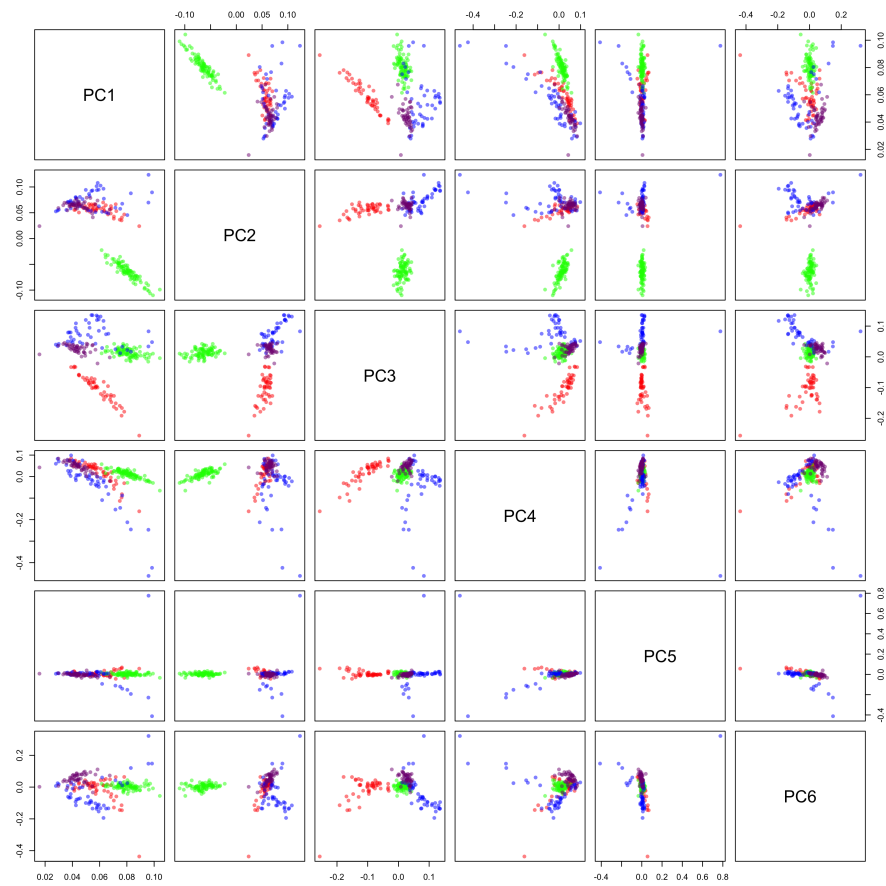
### 2.2 Test data

For demonstration of *fuchikoma*, we prepared MARS-Seq data [5], which is a

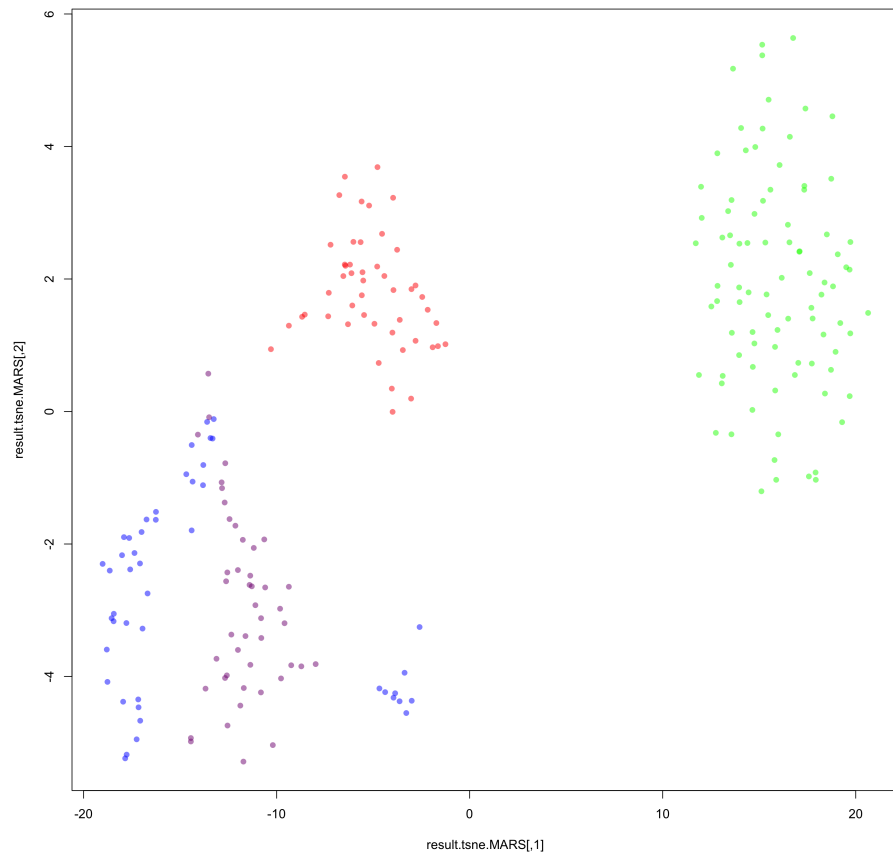
```
> data(MARS)
> data(label.MARS)
> data(result.pca.MARS)
> data(result.tsne.MARS)
> data(result.dmap.MARS)
```

### 2.3 Effectivity of non-linear dimensional reduction

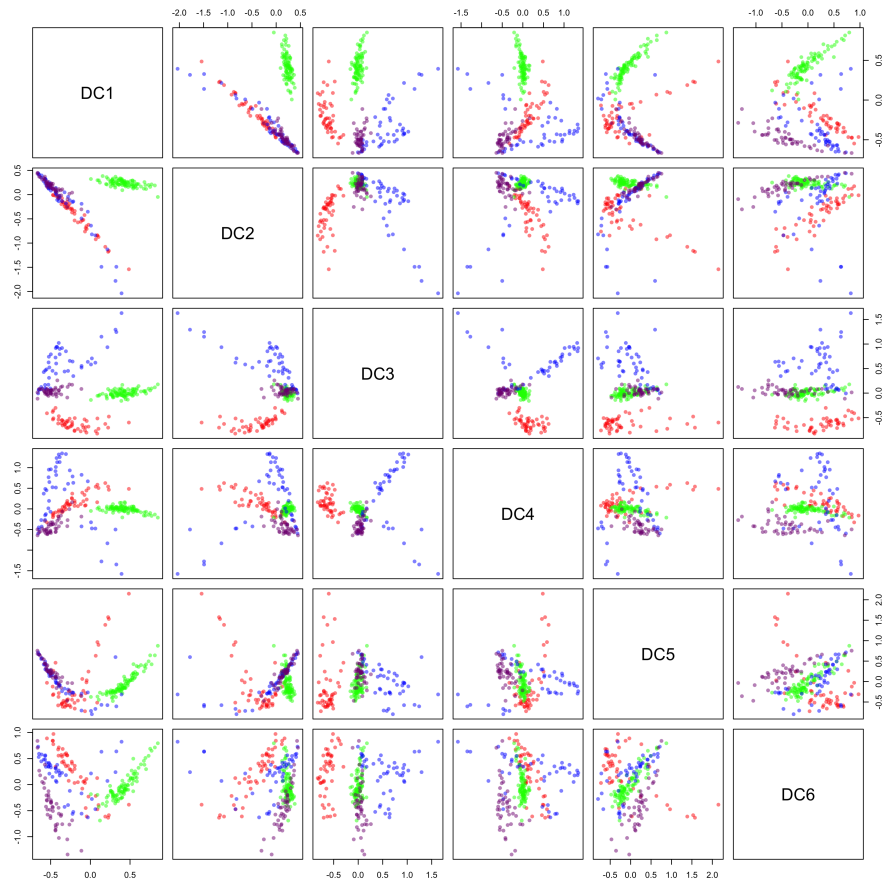
```
> pairs(result.pca.MARS, col=label.MARS, pch=16)
```



```
> plot(result.tsne.MARS, col=label.MARS, pch=16)
```



```
> pairs(result.dmap.MARS, col=label.MARS, pch=16)
```



## 2.4 Limitation of non-linear dimensional reduction

## 2.5 How to get help

All *fuchikoma* questions should be posted to the Bioconductor support site: <https://support.bioconductor.org>, which serves as a repository of questions and answers. See the first question in the list of Frequently Asked Questions (Section 5) for more information about how to construct an informative post.

# 3 How to use fuchikoma

---

## 3.1 fuchikoma : Supervised mode (ANOVA-like)

## 3.2 fuchikoma : Supervised mode (cluster specific DEGs)

### 3.3 fuchikoma : Unsupervised mode

## 4 Setup

---

This vignette was built on:

```
> toLatex(sessionInfo())
```

- R version 3.2.2 (2015-08-14), x86\_64-apple-darwin13.4.0
- Locale: ja\_JP.UTF-8/ja\_JP.UTF-8/ja\_JP.UTF-8/C/ja\_JP.UTF-8/ja\_JP.UTF-8
- Base packages: base, datasets, graphics, grDevices, methods, parallel, stats, utils
- Other packages: Biobase 2.30.0, BiocGenerics 0.16.1, destiny 1.1.0, doParallel 1.0.10, FNN 1.1, foreach 1.4.3, fuchikoma 0.0.8, iterators 1.0.8, Matrix 1.2-3
- Loaded via a namespace (and not attached): BiocStyle 1.8.0, car 2.1-0, chron 2.3-47, class 7.3-14, codetools 0.2-14, colorspace 1.2-6, data.table 1.9.6, DEoptimR 1.0-4, e1071 1.6-7, grid 3.2.2, igraph 1.0.1, lattice 0.20-33, lme4 1.1-10, lmtest 0.9-34, magrittr 1.5, MASS 7.3-45, MatrixModels 0.4-1, mgcv 1.8-9, minqa 1.2.4, nlme 3.1-122, nloptr 1.0.4, nnet 7.3-11, pbkrtest 0.4-2, proxy 0.4-15, quantreg 5.19, Rcpp 0.12.2, RcppEigen 0.3.2.5.1, robustbase 0.92-5, scatterplot3d 0.3-36, sp 1.2-1, SparseM 1.7, splines 3.2.2, tools 3.2.2, vcd 1.4-1, VIM 4.4.1, zoo 1.7-12

## References

---

- [1] Laurens van der Maaten, Geoffrey Hinton *Journal of Machine Learning Research* *Visualizing High-Dimensional Data Using t-SNE*, 9(Nov): 2579-2605, 2008.
- [2] Laleh Haghverdi, Florian Buettner and Fabian J. Theis *Bioinformatics* *Diffusion maps for high-dimensional single-cell analysis of differentiation data*, 31(18): 2989-2998, 2015.
- [3] Le Song, Justin Bedo, Karsten M. Borgwardt, Arthur Gretton and Alex Smola *Bioinformatics* *Gene selection via the BAHSIC family of algorithms*, 23(13): i490-i498, 2007.
- [4] Y-h Taguchi, Mitsuo Iwadata and Hideaki Umeyama *BMC Bioinformatics* *Principal component analysis-based unsupervised feature extraction applied to in silico drug discovery for posttraumatic stress disorder-mediated heart disease*, 16(139), 2015.
- [5] Diego Adhemar Jaitin, Ephraim Kenigsberg, Hadas Keren-Shaul, Naama Elefant, Franziska Paul, Irina Zaretsky, Alexander Mildner, Nadav Cohen, Steffen Jung, Amos Tanay and Ido Amit *Science* *Massively Parallel Single-Cell RNA-Seq for Marker-Free Decomposition of Tissues into Cell Types*, 343(6172): 776-779, 2014.
- [6] Arthur Gretton, Kenji Fukumizu, Choon Hui Teo, Le Song, Bernhard Scholkopf and Alexander J. Smola *NIPS 21 A Kernel Statistical Test of Independence*, 2007.
- [7] Aaditya Ramdas and Leila Wehbe *IJCAI 15 Nonparametric Independence Testing for Small Sample Sizes*, 2015.