

# Multomics data analysis using tensor decomposition based unsupervised feature extraction --Comparison with DIABLO--

Y-h. Taguchi<sup>1</sup>[0000-0003-0867-8986]

<sup>1</sup> Department of Physics, Chuo University, Tokyo 112-8551, Japan  
tag@granular.com

**Abstract.** Multomics data analysis is the central issue of genomics science. In spite of that, there are not well defined methods that can integrate multomics data sets, which are formatted as matrices with different sizes. In this paper, I propose the usage of tensor decomposition based unsupervised feature extraction as a data mining tool for multomics data set. It can successfully integrate miRNA expression, mRNA expression and proteome, which were used as a demonstration example of DIABLO that is the recently proposed advanced method for the integrated analysis of multomics data set.

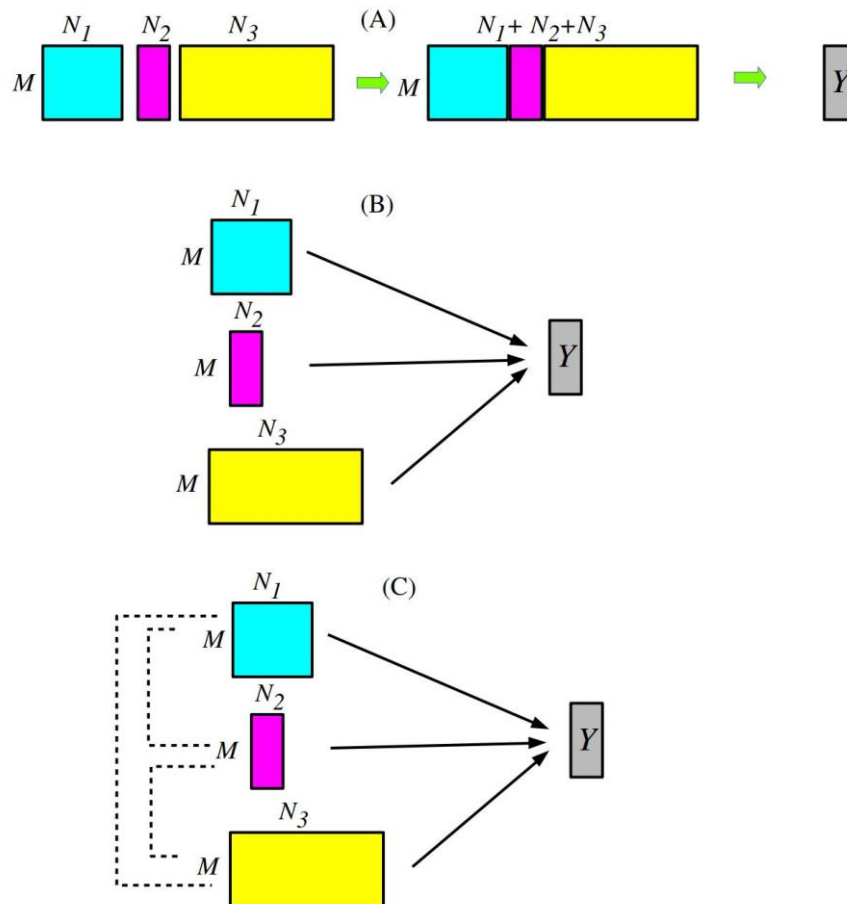
**Keywords:** tensor decomposition, feature selection, multomics data set.

## 1 Introduction

Multomics data, including miRNA expression, mRNA expression, promoter methylation and histone modification, have recently come to be measured over the various biological problems. In contrast to the rapid development of measurement technology, the data analysis pipeline developed very slowly. This is principally because we have never faced the flood of data set until very recently; measuring data set has ever been more expensive than analyzing it. Thus, only limited efforts have been spent for analysis of high dimensional data that has ever been rarely available.

Multomics data is a typical high dimensional dataset; the number of samples (typically  $<10^2$ ) is always less than number of features, i.e., that of mRNAs ( $\sim 10^4$ ), that of miRNAs ( $\sim 10^3$ ) and that of methylation sites ( $\sim 10^5$ ). There are several methods proposed in order to integrate multomics data formatted as matrices with distinct sizes. **Fig. 1** shows typical three strategies to integrate three matrices that share the  $M$  samples for which three distinct features whose numbers are  $N_1$ ,  $N_2$ , and  $N_3$  are measured. **Fig. 1(A)**, contraction, is simply merging three matrices such that they share  $M$  samples as rows while  $N_1$ ,  $N_2$ , and  $N_3$  features are aligned as columns. After merging them, all downstream analyses are performed with assuming only one feature whose number is  $N_1+N_2+N_3$ . Although this strategy looks simple, there can be multiple disadvantages. First of all, if the number of individual features,  $N_1$ ,  $N_2$ , and  $N_3$ , differ from one another so much, features having the smallest number might be neglected.

This might prevent us from considering three features equally. Even if the numbers are almost equal, if some feature has much larger amplitude than others, it also prevents us from dealing with them equally. In this sense, it is very usual to normalize individual features before the contraction. Nevertheless, the normalization might also affect the outcome, because how to normalize them is quite subjective.



**Fig. 1.** Three distinct strategies that integrate multiomics formatted as matrices with distinct dimensions,  $M \times N_1$ ,  $M \times N_2$ , and  $M \times N_3$ . (A) Contraction (B) Ensemble (C) DIABLO.

In contrast to contraction, ensemble strategy shown in **Fig. 1(B)** oppositely does not integrate multiomics data until very last stage. The simplest ensemble strategy is to train individual feature independently, and decision is made upon the majority rule. This strategy is also simple because we do not need any new strategies considering integration of multi omics data set. The disadvantage of this strategy is obvious. There is no justification for giving one vote to individual feature. If the majority of features are useless, i.e., do not have any practical relationship with outcome  $Y$ , voting system

simple increases the noise. In this sense, ensemble can be worse than contraction that can disregard useless features during training.

The third strategy shown in **Fig. 1(C)**, DIABLO [1], which is a part of mixOmics [2] that aims integrated analysis of multiomics data, is more advanced method. In DIABLO strategy, not individual features but sets (pairs) of individual features are employed for training. Then, outcome  $Y$  is predicted based upon ensemble strategy of them. Thus, DIABLO strategy has more ability to learn effective features from available multiomics data. One disadvantage is that how to relate individual features must be designed by human beings; this process can be quite subjective.

In this paper, in contrast to these three strategies in **Fig. 1**, unsupervised strategy based upon tensor decomposition (TD), in short, “TD based unsupervised feature extraction (FE)”, is proposed and is applied to multiomics data set. As can be seen later, TD based unsupervised FE achieves performance competitive with that achieved by DIABLO strategy.

## 2 Materials and Methods

### 2.1 Multiomics data set

Data set used here is included in mixOmics package<sup>1</sup>. As described in the web page “Case study: TCGA”<sup>2</sup>, this data set can be loaded into R using `data('breast.TCGA')` command after installing mixOmics package. It includes 150 samples composed of 45 Basal, 30 Her2 and 75 LumA subtypes, to which 200 mRNAs, 184 miRNAs and 142 proteome expression are measured.

### 2.2 Tensor decomposition

We apply TD with higher order singular value decomposition (HOSVD) [3] algorithm to case I type I tensor [4]. Starting from three matrices,  $x_{i_1j}^{\text{mRNA}} \in \mathbb{R}^{200 \times 150}$ ,  $x_{i_2j}^{\text{miRNA}} \in \mathbb{R}^{184 \times 150}$ , and  $x_{i_3j}^{\text{prot}} \in \mathbb{R}^{142 \times 150}$ , we generate four mode tensor

$$x_{i_1i_2i_3j} = x_{i_1j}^{\text{mRNA}} \cdot x_{i_2j}^{\text{miRNA}} \cdot x_{i_3j}^{\text{prot}} \quad (1)$$

and HOSVD is applied it as

$$x_{i_1i_2i_3j} = \sum_{\ell_1=1}^{200} \sum_{\ell_2=1}^{184} \sum_{\ell_3=1}^{142} \sum_{\ell_4=1}^{150} G(\ell_1, \ell_2, \ell_3, \ell_4) u_{\ell_1i_1}^{\text{mRNA}} u_{\ell_2i_2}^{\text{miRNA}} u_{\ell_3i_3}^{\text{prot}} u_{\ell_4j} \quad (2)$$

where  $u_{\ell_1i_1}^{\text{mRNA}} \in \mathbb{R}^{200 \times 200}$ ,  $u_{\ell_2i_2}^{\text{miRNA}} \in \mathbb{R}^{184 \times 184}$ ,  $u_{\ell_3i_3}^{\text{prot}} \in \mathbb{R}^{142 \times 142}$ , and  $u_{\ell_4j} \in \mathbb{R}^{150 \times 150}$ , are singular value matrices (they are also orthogonal matrices) and  $G \in \mathbb{R}^{200 \times 184 \times 142 \times 150}$  is a core tensor.

<sup>1</sup> <http://www.bioconductor.org/packages/release/bioc/html/mixOmics.html>

<sup>2</sup> <http://mixomics.org/mixdiablo/case-study-tcga/>

### 3 Results

#### 3.1 DIABLO

Here I briefly summarize how well DIABLO works when it is applied to data set described in Sec. 2.1 as denoted in the web page “Case study: TCGA”<sup>2</sup>. The interaction assumed is mRNA-miRNA, mRNA-proteome, and miRNA-proteome.

1. Three subtypes can be discriminated with the success rate of 95 % with using more than two components generated by DIABLO.
2. 15 mRNAs, 18 miRNAs and 7 proteins are selected.
3. Using features selected in step 2, heatmap is drawn. These three subtypes are well separated in the heatmap. Thus, DIABLO successfully selects limited number of features that discriminate three subtypes.

Thus, the point is if TD based unsupervised FE can achieve similar performance as those by DIABLO or not.

#### 3.2 TD based unsupervised FE

As described in Sec. 2.2, HOSVD is applied to case I type I tensor, eq. (1). Then two singular value vectors,  $u_{1j}$  and  $u_{4j}$ , turn out to discriminate three subtypes well (**Fig. 2**). In order to quantitatively evaluate how well these two singular value vector discriminates three subtypes, we apply linear discriminant analysis (LDA) with using leave one out cross validation (LOOCV) (**Table 1**). It achieved as high as 95 % accuracy, which is almost equal to that by DIABLO.

**Table 1.** The confusion table obtained by applying LDA to three cancer subtypes with  $u_{1j}$  and  $u_{4j}$ . LOOCV is used. Row: inference, column: true subtypes. Numbers in bold represent those of correctly predicted samples.

	Basel	Her2	LumA
Basel	<b>42</b>	4	0
Her2	2	<b>25</b>	2
LumA	1	1	<b>73</b>

Next, in order to see if the limited number of mRNAs, miRNAs and proteomes can be selected to discriminate three cancer subtypes well, we select subsets of these three features. In order that, we first need to find which  $u_{\ell_1 i_1}^{\text{mRNA}}$ ,  $u_{\ell_2 i_2}^{\text{miRNA}}$ , and  $u_{\ell_3 i_3}^{\text{prot}}$  are most associated with  $u_{1j}$  and  $u_{4j}$ . This requires the identification of core tensor,  $G(\ell_1, \ell_2, \ell_3, \ell_4)$ , having the largest absolute values with  $\ell_4 = 1, 4$ . **Table 2** shows the list of those  $G(\ell_1, \ell_2, \ell_3, \ell_4)$ s with the descending order of absolute values of  $G(\ell_1, \ell_2, \ell_3, \ell_4)$ s. It is obvious that  $G(\ell_1, \ell_2, \ell_3, \ell_4)$ s with  $1 \leq \ell_1, \ell_2 \leq 2$  and  $1 \leq \ell_3 \leq 4$  have larger absolute values. Then we compute the squared summation of singular value vectors attributed to  $i_1$ th mRNA,  $i_2$ th miRNA, and  $i_3$ th proteome as

$$\sum_{\ell_1=1}^2 (u_{\ell_1 \ell_1}^{\text{mRNA}})^2$$

$$\sum_{\ell_2=1}^2 (u_{\ell_2 \ell_2}^{\text{miRNA}})^2$$

$$\sum_{\ell_3=1}^4 (u_{\ell_3 \ell_3}^{\text{prot}})^2$$

and select top ranked ten mRNAs, miRNAs and proteome with larger squared summation of singular value vectors.

**Table 2.**  $G(\ell_1, \ell_2, \ell_3, \ell_4)$ s sorted descending order of absolute values among those having  $\ell_4 = 1, 4$ .

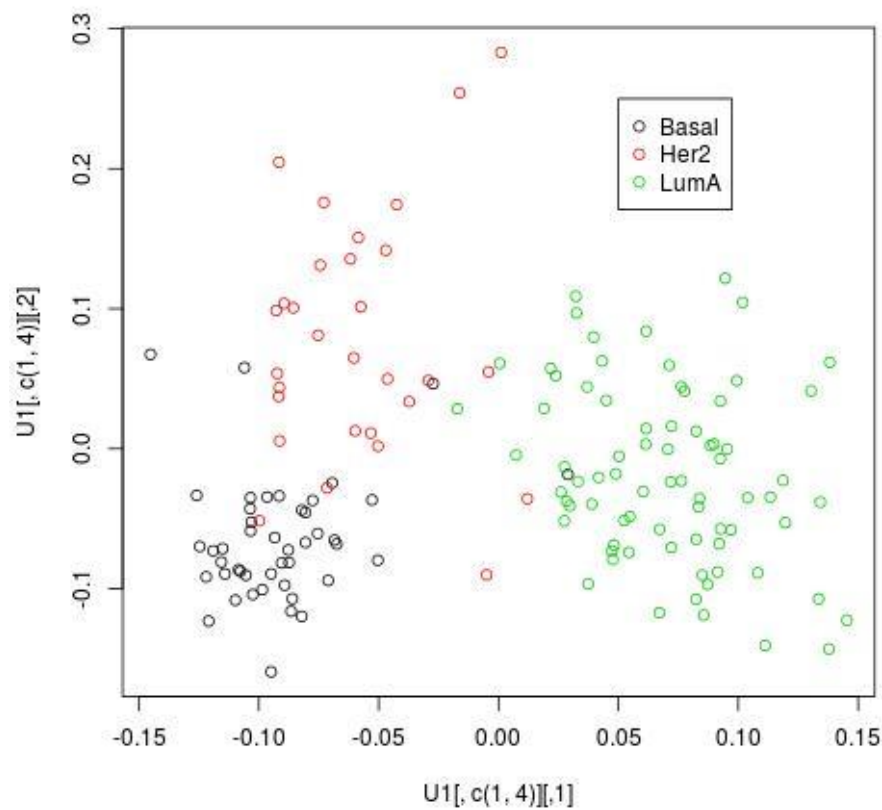
rank	$\ell_1$	$\ell_2$	$\ell_3$	$\ell_4$	$G(\ell_1, \ell_2, \ell_3, \ell_4)$
1	1	1	1	1	-407857.582
2	1	1	4	4	-209720.615
3	2	1	1	4	-20452.480
4	2	1	3	1	-11677.505
5	2	1	4	1	-10428.742
6	2	1	2	1	10157.467
7	1	1	2	1	-8973.774
8	1	2	1	4	8360.976
9	2	1	5	4	-6628.467
10	1	1	3	4	6623.046

**Fig. 3** shows the heatmap generated using selected 10 mRNAs, 10 miRNAs and 10 proteome. It is obvious that three subtypes are clustered well separately; it is competitively well as compared with the heatmap generated by DIABLO<sup>2</sup>. In addition to this, three features that share same profiles are clustered together as in the result by DIABLO (see hierarchical clustering of columns in **Fig. 3**). This suggests that TD based unsupervised FE has ability to select distinct features sharing same profiles over three subtypes as DIABLO achieved.

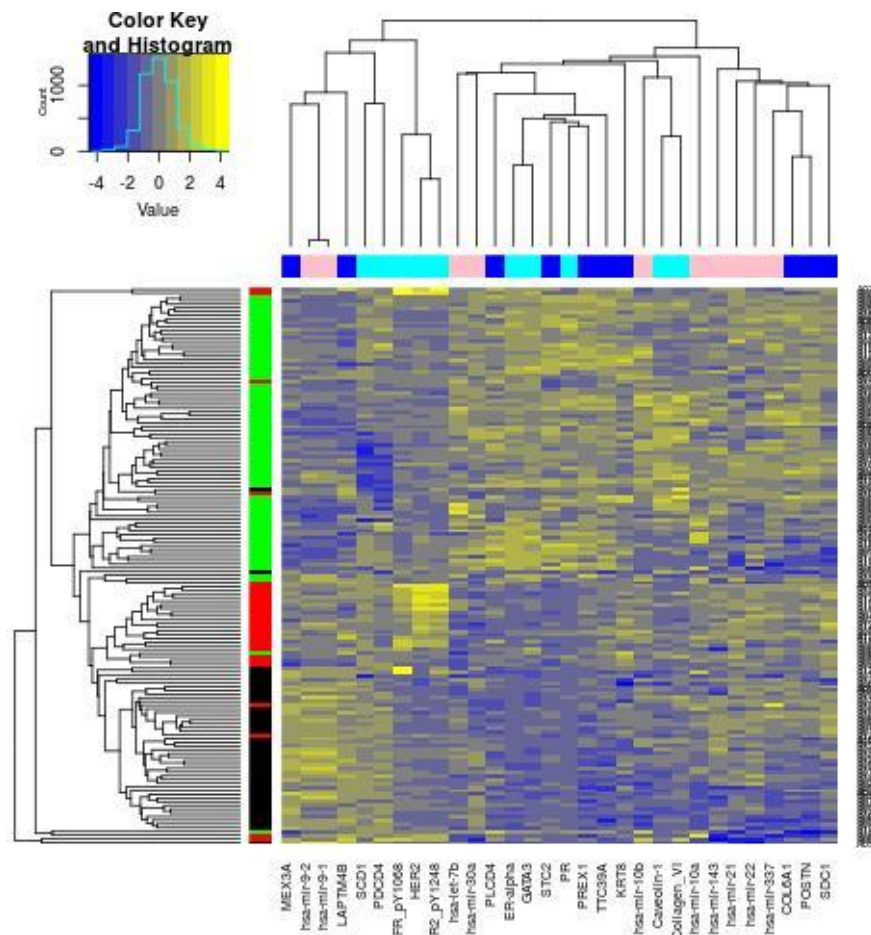
## 4 Discussions

In this paper, I applied recently proposed TD based unsupervised FE, which was applied to wide range of studies [4-35], to multiomics data analysis. Apparently, the performance achieved by TD based unsupervised FE is at most competitive to that of

DIABLO. Thus, one might wonder why TD based unsupervised FE is recommended in this study. There are multiple reasons for this recommendation. First, DIABLO used random number seed to perform analysis. This means, every time we select different random number seed, we inevitably have distinct sets of selected mRNAs, miRNAs and proteome, although overall performance remains unchanged. This might be problematic because we might be interested in identification of disease causing genes. If distinct set of features are selected every time we change random number seed, it might prevent us from interpreting the outcome biologically. In contrast to this, TD based unsupervised FE can give us quite stable outcomes. Not only TD based unsupervised FE does not need random number seed, but also it is quite stable even when samples are resampled [4-35]. In this point, TD based unsupervised FE is more suitable to be employed than DIABLO from the biological point of views.



**Fig. 2.** Scatter plot of  $u_{1j}$  (horizontal axis) and  $u_{4j}$  (vertical axis).



**Fig. 3.** Heatmap of selected 10 mRNAs (blue), 10 miRNAs (pink), and 10 proteome (cyan) in columns and 150 samples (black:Basel, red:Her2, green:Luma). Yellow: expressed expression, blue: depressed expression.

Also from the computational point of views, TD based unsupervised FE is recommended more than DIABLO. From the point of computational time, DIABLO requires more time than TD based unsupervised FE, because DIABLO needs to learn from the data set and labeling while TD based unsupervised FE does not require this process due to unsupervised nature. In this sense, if these two achieve equally, TD based unsupervised FE is more recommended method than DIABLO.

## References

1. Singh, A., Shannon, C.P., Gautier, B., Rohart, F., Vacher, M., Tebbutt, S. J., Cao, K-L.: DIABLO: an integrative approach for identifying key molecular drivers from multi-omics assays, *Bioinformatics*, bty1054, <https://doi.org/10.1093/bioinformatics/bty1054>
2. Rohart, F., Gautier, B., Singh, A., Lê Cao, K-A.: mixOmics: An R package for 'omics feature selection and multiple data integration. *PLoS Comput Biol* 13(11): e1005752 (2017).
3. Symeonidis, P., Zioupos, A.: HOSVD on Tensors and Its Extensions. In: *Matrix and Tensor Factorization Techniques for Recommender Systems*, pp. 81–93. Springer, Heidelberg (2016).
4. Taguchi, Y-h.: Tensor decomposition-based unsupervised feature extraction applied to matrix products for multi-view data processing. *PLoS ONE* 12(8): e0183933 (2017).
5. Ishida, S., Umeyama, H., Iwade, M., Taguchi, Y.H.: Bioinformatic Screening of Autoimmune Disease Genes and Protein Structure Prediction with FAMS for Drug Discovery. *Protein Pept. Lett.* 21, 828–839 (2014).
6. Taguchi, Y.-H.: microRNA-mRNA interaction identification in Wilms tumor using principal component analysis based unsupervised feature extraction. In: *2016 IEEE 16th International Conference on Bioinformatics and Bioengineering (BIBE)*. pp. 71–78 (2016).
7. Murakami, Y., Kubo, S., Tamori, A., Itami, S., Kawamura, E., Iwaisako, K., Ikeda, K., Kawada, N., Ochiya, T., Taguchi, Y.-H.: Comprehensive analysis of transcriptome and metabolome analysis in Intrahepatic Cholangiocarcinoma and Hepatocellular Carcinoma. *Sci. Rep.* 5, 16294 (2015).
8. Taguchi, Y.-H.: Identification of More Feasible MicroRNA-mRNA Interactions within Multiple Cancers Using Principal Component Analysis Based Unsupervised Feature Extraction. *Int. J. Mol. Sci.* 17, 696 (2016).
9. Murakami, Y., Toyoda, H., Tanahashi, T., Tanaka, J., Kumada, T., Yoshioka, Y., Kosaka, N., Ochiya, T., Taguchi, Y. h.: Comprehensive miRNA Expression Analysis in Peripheral Blood Can Diagnose Liver Disease. *PLoS One.* 7, e48366 (2012).
10. Taguchi, Y.-H.: Identification of candidate drugs using tensor-decomposition-based unsupervised feature extraction in integrated analysis of gene expression between diseases and DrugMatrix datasets. *Sci. Rep.* 7, 13733 (2017).
11. Murakami, Y., Kubo, S., Tamori, A., Itami, S., Kawamura, E., Iwaisako, K., Ikeda, K., Kawada, N., Ochiya, T., Taguchi, Y.H.: Comprehensive analysis of transcriptome and metabolome analysis in Intrahepatic Cholangiocarcinoma and Hepatocellular Carcinoma. *Sci Rep.* 5, 16294 (2015).
12. Tamori, A., Murakami, Y., Kubo, S., Itami, S., Uchida-Kobayashi, S., Morikawa, H., Enomoto, M., Takemura, S., Tanahashi, T., Taguchi, Y.-H., Kawada, N.: MicroRNA expression in hepatocellular carcinoma after the eradication of chronic hepatitis virus C infection using interferon therapy. *Hepatol. Res.* 46, (2016).
13. Taguchi, Y.-H., Iwade, M., Umeyama, H., Murakami, Y.: Principal component analysis based unsupervised feature extraction applied to bioinformatics analysis. *Comput. Methods with Appl. Bioinforma. Anal.* 153–182 (2017).



14. Taguchi, Y.H.: Principal Components Analysis Based Unsupervised Feature Extraction Applied to Gene Expression Analysis of Blood from Dengue Haemorrhagic Fever Patients. *Sci. Rep.* 7, 44016 (2017).
15. Taguchi, Y.-H., Iwadate, M., Umeyama, H., Murakami, Y., Okamoto, A.: Heuristic principal component analysis-based unsupervised feature extraction and its application to bioinformatics. (2014).
16. Taguchi, Y.-H.: Principal component analysis based unsupervised feature extraction applied to publicly available gene expression profiles provides new insights into the mechanisms of action of histone deacetylase inhibitors. *Neuroepigenetics.* 8, 1–18 (2016).
17. Taguchi, Y.-H., Murakami, Y.: Universal disease biomarker: can a fixed set of blood microRNAs diagnose multiple diseases? *BMC Res. Notes.* 7, 581 (2014).
18. Taguchi, Y.-H.: Principal component analysis based unsupervised feature extraction applied to budding yeast temporally periodic gene expression. *BioData Min.* 9, 22 (2016).
19. Umeyama, H., Iwadate, M., Taguchi, Y.-H.: TINAGL1 and B3GALNT1 are potential therapy target genes to suppress metastasis in non-small cell lung cancer. *BMC Genomics.* 15, S2 (2014).
20. Taguchi, Y.H., Murakami, Y.: Principal component analysis based feature extraction approach to identify circulating microRNA biomarkers. *PLoS One.* 8, e66714 (2013).
21. Taguchi, Y.-H., Wang, H.: Genetic Association between Amyotrophic Lateral Sclerosis and Cancer. *Genes (Basel).* 8, 243 (2017).
22. Taguchi, Y.-H., Iwadate, M., Umeyama, H.: SFRP1 is a possible candidate for epigenetic therapy in non-small cell lung cancer. *BMC Med. Genomics.* 9, (2016).
23. Taguchi, Y. -H., Iwadate, M., Umeyama, H.: Principal component analysis-based unsupervised feature extraction applied to in silico drug discovery for posttraumatic stress disorder-mediated heart disease. *BMC Bioinformatics.* 16, 139 (2015).
24. Taguchi, Y.-H., Iwadate, M., Umeyama, H.: Heuristic principal component analysis-based unsupervised feature extraction and its application to gene expression analysis of amyotrophic lateral sclerosis data sets. In: *Computational Intelligence in Bioinformatics and Computational Biology (CIBCB), 2015 IEEE Conference on.* pp. 1–10 (2015).
25. Taguchi, Y.-H., Umeyama, H., Iwadate, M., Murakami, Y., Okamoto, A.: Heuristic Principal Component Analysis-Based Unsupervised Feature Extraction and Its Application to Bioinformatics. In: Baoying Wang, Ruowang Li, and William Perizzo (eds.) *Big Data Analytics in Bioinformatics and Healthcare.* pp. 138–162. IGI global (2015).
26. Murakami, Y., Tanahashi, T., Okada, R., Toyoda, H., Kumada, T., Enomoto, M., Tamori, A., Kawada, N., Taguchi, Y.H., Azuma, T.: Comparison of hepatocellular carcinoma miRNA expression profiling as evaluated by next generation sequencing and microarray. *PLoS One.* 9, (2014).
27. Taguchi, Y.-H.: Integrative Analysis of Gene Expression and Promoter Methylation during Reprogramming of a Non-Small-Cell Lung Cancer Cell Line Using Principal Component Analysis-Based Unsupervised Feature Extraction. In: *ICIC2014.* pp. 445–455 (2014).
28. Taguchi, Y.-H.: One-class Differential Expression Analysis using Tensor Decomposition-based Unsupervised Feature Extraction Applied to Integrated Analysis of

- Multiple Omics Data from 26 Lung Adenocarcinoma Cell Lines, In: 2017 IEEE 17th International Conference on Bioinformatics and Bioengineering (BIBE), Washington, DC, pp. 131-138. (2017) doi: 10.1109/BIBE.2017.00-66
29. Taguchi, Y.-H.: Principal Component Analysis-Based Unsupervised Feature Extraction Applied to Single-Cell Gene Expression Analysis. In: ICIC2018. pp. 816–826 (2018).
  30. Taguchi, Y.-H., Wang, H.: Exploring microRNA Biomarker for Amyotrophic Lateral Sclerosis. *Int. J. Mol. Sci.* 19(5), 131 (2018).
  31. Taguchi, Y.-H.: Tensor decomposition-based and principal-component-analysis-based unsupervised feature extraction applied to the gene expression and methylation profiles in the brains of social insects with multiple castes. *BMC Bioinformatics*. 19, S9, 99 (2018).
  32. Taguchi, Y.-H.: Tensor Decomposition-Based Unsupervised Feature Extraction Can Identify the Universal Nature of Sequence-Nonspecific Off-Target Regulation of mRNA Mediated by MicroRNA Transfection. *Cells*. 7(6), 54 (2018).
  33. Taguchi, Y.-H.: Tensor Decomposition-Based Unsupervised Feature Extraction for Integrated Analysis of TCGA Data on MicroRNA Expression and Promoter Methylation of Genes in Ovarian Cancer, In: 2018 IEEE 18th International Conference on Bioinformatics and Bioengineering (BIBE), Taichung, Taiwan, pp. 195-200. (2018) doi: 10.1109/BIBE.2018.00045
  34. Taguchi, Y.-H.: Drug candidate identification based on gene expression of treated cells using tensor decomposition-based unsupervised feature extraction for large-scale data. *BMC Bioinformatics*. 19, S13, 388 (2019).
  35. Taguchi, Y.-H., Wang, H.: Exploring MicroRNA Biomarkers for Parkinson's Disease from mRNA Expression Profiles. *Cells*. 7(12), 245 (2018).

## Appendix: R code for TD based unsupervised FE

```
install.packages("mixOmics") #mixOmics installation
require(mixOmics) #mixOmics load
data("breast.TCGA") #TCGA data loading
data = list(mRNA = breast.TCGA$data.train$mrna,
            miRNA = breast.TCGA$data.train$mirna,
            proteomics = breast.TCGA$data.train$protein)
install.packages("rTensor") #rTensor installation
require(rTensor) #rTensor load
#tensor generation
Z <- array(NA,c(150,200,184,142))
for (i in c(1:150))
{cat(i, " ")
Z[i,,] <- data.matrix(outer(outer(data$mrna[i,],
data$miRNA[i,,"*"]),data$proteomics[i,,"*"])))
#tensor decomposition
HOSVD <- hosvd(as.tensor(Z))
U1 <- HOSVD$U[[1]] #x_{\ell_4,J}
U2 <- HOSVD$U[[2]] #x_{\ell_1,i_1}
```

```

U3 <- HOSVD$U[[3]] #x_{\ell_2,i_2}
U4 <- HOSVD$U[[4]] #x_{\ell_3,i_3}
#150 samples scatter plot (x_{\ell_4,j})
plot(U1[,c(1,4)],col=breast.TCGA$data.train$subtype)
legend(0.05,0.25,names(summary(breast.TCGA$data.train
$subtype)),col=1:3,pch=1)
require(MASS) #MASS load
LD <- lda(U1[,c(1,4)],breast.TCGA$data.train$subtype,CV=T,prior=rep(1/3,3))
#linear discriminant analysis
table(LD$class,breast.TCGA$data.train$subtype) #confusion matrix
ZZ <- order(-abs(HOSVD$Z@data[c(1,4),,]))
[1:20];data.frame(arrayInd(ZZ,dim(HOSVD$Z@data[c(1,4),,])),HOSVD$Z@data[c(1,4),,][ZZ]) #sorting G with ascending order of absolute values of G
P2 <- pchisq(rowSums(scale(U2[,1:2])^2),2,lower.tail=F)
P3 <- pchisq(rowSums(scale(U3[,1:2])^2),2,lower.tail=F)
P4 <- pchisq(rowSums(scale(U4[,1:4])^2),4,lower.tail=F)
U <-
cbind(data$mRNA[,order(P2)[1:10]],cbind(data$miRNA[,order(P3)[1:10]],data
$proteomics[,order(P4)[1:10]])) #select top 10 with larger squared summation
of singular value vectors
install.packages("gplots") #gplots install
require(gplots) #gplot load
#heatmap plot
heatmap.2(scale(U),col=rgb(seq(0,1,by=0.1),seq(0,1,by=0.1),seq(1,0,by=-
0.1)),RowSideColors=c(rep("black",45),rep("red",30),rep("green",75)),hclustfun=
function(x){hclust(x,method="average")},trace="none",ColSideColors=c(rep
("blue",10),rep("pink",10),rep("cyan",10)))

```