RESEARCH

Universal nature of drug treatment responses of drug-tissue-wide model-animal experiments

Y-h. Taguchi[1]* and Turki Turki[2]

*Correspondence:
tag@granular.com
[1]Department of Physics, Chuo
University, 1-13-27 Kasuga,
Bunkyo-ku, Tokyo 112-8551,
JAPAN
Full list of author information is
available at the end of the article

Abstract

Background: Gene expression profiles of tissues treated with drug recently came to be used for the inference of clinical outcomes. Although it is often successful from the application point of view, gene expression altered by the drug is rarely analyzed in detail because of too many number of genes measured.

Method: We apply tensor decomposition (TD) based unsupervised feature extraction (FE) to gene expression profiles of 24 mice tissues treated with 15 drugs.

Results: TD based unsupervised FE unexpectedly identified universal feature of 15 drugs, i.e., the effect of 15 drugs are common to some extent. These drugs affect genes in genes-group wide manner, i.e. dependent upon three set of tissue types (Neuronal tissues, muscle tissues, and gastroenterological tissues). For each tissue group, TD based unsupervised FE identified a few tens to a few hundreds genes that are affected by drug treatment. These genes are distinctly expressive between drug treatment and controls as well as between tissues in individual tissue groups and other tissues. Our various enrichment analysis performed guaranteed that the selected genes attributed to individual tissue groups are appreciated.

Conclusions: Our TD based unsupervised FE is the promising method to perform integrated analysis of gene expression profiles of multiple tissues treated with multiple drugs in fully unsupervised manner.

Keywords: tensor decomposition; gene selection; drug treatment

Background

Drug design is a time-consuming and expensive process. It requires numerous experimental efforts that execute compound investigation in a massive trial and error manner. This is generally because it is very difficult to find genes that cause disease and are targeted for therapy. Even if genes to be targeted are successfully identified, it is also difficult to find some drug candidate compounds that successfully bind to the proteins that the target genes encode.

In order to shorten the period of drug development and to reduce the expense used for the drug development, computer aided methods were invented. Major two computer aided drug design strategies are ligand based drug design (LBDD) and structure based drug design (SBDD). LBDD has various advantages including the less required computational resources and better successful rate for drug design. On the other hand, LBDD has a weak point: less ability to find drug candidate compounds whose structural similarity with known drug is less. In order to fulfill the weak point of LBDD, SBDD has more ability to find drug candidate compounds

that lack structural similarity with known drug, since SBDD tries to screen drug candidate compounds by investigating whether candidate compounds can bind to target proteins. The weak points of SBDD is that it requires massive computational resources. This huge amount of computational resources required prevents us from applying SBDD to many drug candidate compounds whose numbers are often raised to several millions.

Following the reduced prices required for obtaining gene expression profiles, the third computer aided strategy is raised: gene expression profile based drug design. In this strategy, gene expression profiles of tissues/cell lines treated with candidate drug compounds are collected. Then collected profiles are compared with those of tissues/cell lines treated with known drugs compounds. If drug candidate compounds share a gene expression profile to some extent with known drug compounds, they are identified as hopeful drug candidate compounds for target diseases/proteins.

In order to consider the gene expression profile for the drug design, some data bases are established. For example, chemical checker [1] includes gene expression for computer aided drug design while PharmacoDB [2] is fully implemented to consider dose dependence of drug treated cell lines for drug design. Many papers were published to make use of gene expression profiles for computer aided drug design [3, 4]. Huang et al [5] made use of combinatorial analysis of drug treated gene expression for cancer drugs, which were experimentally confirmed *in vitro*. Lee et al. [6] proposed DeSigN that is a robust and useful method for the identification of candidate drugs using an input gene signature obtained from gene expression analysis. Kim et al. [7] performed computational drug repositioning for gastric cancer using reversal gene expression profiles. Wolf et al. [8] analyzed high-throughput gene expression profiles to define drug similarity and predict compound activity. Hodos et al. [9] tried to fill missing observations of gene expression of cells treated with drugs by predicting cell-specific drug perturbation profiles using available expression data from related conditions. Liu et al. [10] performed comparative analysis of genes frequently regulated by drugs based on connectivity map transcriptome data. Pabon et al. [11] predicted protein targets for drug-like compounds using transcriptomics.

In contrast to these successful applications of gene expression profile analysis to computer aided drug design, how individual gene expression is affected via drug treatment is unclear. First of all, the number of gene expression dose dependence is as many as that of gene expression. Thus, it is not easy to invent useful method to integrate and understand these huge number of dose dependence pertaining to individual gene expression profile. For example, Luka et al. [12] employed principal component analysis (PCA) to integrate gene expression profile dose dependence upon combinatorial drug treatment. They reported that they have found convex (not monotonic) dependence upon dose density and identified it as the evidence of cooperative effects caused by dual drug treatments. Nevertheless, the convex dependence upon dose dependence was reported to be observed in a single drug treatment if tensor decomposition (TD) was employed to integrate multiple gene expression profiles of cell lines treated by single drug [13]. Thus, it is primarily important to identify the effective method that can integrate numerous gene expression profiles of tissues/cell lined treated by drugs.

Recently, Kozawa et al. [14] has made use of gene expression profiles of mice tissues treated with drugs in order to predict human clinical outcomes. In this paper, by applying TD based unsupervised feature extraction (FE) to gene expression profiles used in their study, we try to identify what has happened to gene expression profiles of tissues which are treated with individual drugs. In contrast to expectation, the response of gene expression profile towards individual drug treatment is quite universal. We successfully identified multiple sets of genes whose gene expression profiles are altered by individual drug treatments in quite tissues-group specific manners.

Materials and Methods

Figure 1 shows the flow chart of analysis.

Gene expression profiles

Gene expression profiles used in this study were downloaded from gene expression omnibus (GEO) with GEO ID GSE142068. Twenty four profiles named "GSE142068_count_XXXXX.txt.gz" were downloaded where "XXXXX" stands for one of 24 tissues, i.e., AdrenalG, Aorta, BM (Bone marrow), Brain, Colon, Eye, Heart, Ileum, Jejunum, Kidney, Liver, Lung, Pancreas, ParotidG, PituitaryG, SkMuscle, Skin, Skull, Spleen, Stomach, Spleen, Thymus, ThyroidG, and WAT (white adipose tissue), which were treated with 15 drugs, i.e., Alendronate, Acetaminophen, Aripiprazole, Asenapine, Cisplatin, Clozapine, Clozapine, Empagliflozin, Lenalidomide, Lurasidone, Olanzapine, Evolocumab, Risedronate, Sofosbuvir, Teriparatide, and Wild type (WT).

TD based unsupervised FE

In order to apply TD based unsupervised FE [15] to downloaded gene expression profiles, they must be formatted as a tensor. In this analysis, they are formatted as tensor, $x_{ijkm} \in \mathbb{R}^{N \times 24 \times 18 \times 2}$, for N genes, 24 tissues, 18 drug treatments and two replicates. Then HOSVD [15] algorithm was applied to x_{ijkm} and we get TD

$$x_{ijkm} = \sum_{\ell_1 \ell_2 \ell_3 \ell_4} G(\ell_1, \ell_2, \ell_3, \ell_4) u_{\ell_1 j} u_{\ell_2 k} u_{\ell_3 m} u_{\ell_4 i} \tag{1}$$

where $G \in \mathbb{R}^{N \times 24 \times 18 \times 2}$ is core tensor, $u_{\ell_1 j} \in \mathbb{R}^{24 \times 24}$, $u_{\ell_2 k} \in \mathbb{R}^{18 \times 18}$, $u_{\ell_3 m} \in \mathbb{R}^{2 \times 2}$, and $u_{\ell_4 i} \in \mathbb{R}^{N \times N}$, represents singular value matrices that are also orthogonal matrices. x_{ijkm} is supposed to be standardized as $\sum_i x_{ijkm} = 0$ and $\sum_i x_{ijkm}^2 = N$.

In order to understand how gene expression profiles are altered with drug treatment in tissues-group wide manner, we first need to investigate $u_{\ell_1 j}$, $u_{\ell_2 k}$, and $u_{\ell_3 m}$. After identifying which ℓ_1, ℓ_2 and ℓ_3 are biologically interesting, which ℓ_4 is associated with $G(\ell_1, \ell_2, \ell_3, \ell_4)$ that have the largest absolute values with fixed ℓ_1, ℓ_2 and ℓ_3.

Using identified $u_{\ell_4 i}$, P-value, P_i, is attributed to gene i as

$$P_i = P_{\chi^2} \left[> \left(\frac{u_{\ell_4 i}}{\sigma_{\ell_4}} \right)^2 \right] \tag{2}$$

where $P_{\chi^2}[> x]$ is the cumulative probability of χ^2 distribution and σ_{ℓ_4} is the standard deviation. Here we assume that $u_{\ell_4 i}$ obeys Gaussian distribution with zero mean since $\sum_i x_{ijkm} = 0$. P_i is corrected via BH criterion [16] and I, a set of genes i associated with adjusted P-values less than 0.01, is selected. For more detailed explanation of TD based unsupervised FE, see the recently published monograph [15].

t test and Wilcoxn test applied to sets of genes classified based upon tissue groups and drugs groups

In order to see the selected set of genes, I, are expressive distinctly between assigned two tissue groups, J, $\{x_{ijkm}|i \in I, j \in J\}$, and $\overline{J}$, $\{x_{ijkm}|i \in I, j \in \overline{J}\}$, we apply two way t test and Wilcoxn test and compute P-values. We also performed similar to two drug groups, K, $\{x_{ijkm}|i \in I, k \in K\}$, and $\overline{K}$, $\{x_{ijkm}|i \in I, k \in \overline{K}\}$.

Enrichment analysis
Selected genes (gene symbols) are uploaded to Enrichr [17] and Metascape [18] in order to validate various biological meanings of selected genes.

Results
Figure 2 summarizes the results obtained in this study.

Drug treatments specificity
After obtaining TD, eq. (1), we first investigate $u_{\ell_2 k}$ attributed to kth drug. Although the number of drugs tested is as many as 15, because of additional three conditions, the total number of drug treatments is as many as 18. Usually, the first singular value vectors represent uniform values (i.e., the components that are not distinct between samples) [15]. Also in this case, u_{1k} does not represent any drug treatment dependence. It is reasonable since the expression of most of genes is unlikely affected by drug treatment. Then we consider the second and the third singular value vectors, u_{2k} and u_{3k}, attributed to drug treatments (Fig. 3). In contrast to the expectation, drug treatments are quite universal. Most of drug treatments (other than (2), (9), (15) and (17)) are separated from control treatments ((2), (9), (15) and (17)) along one direction (red arrow) while diversity among drug treatment spread perpendicular (blue arrow) to the direction only among drug treatments. This suggests that gene expression profiles of genes are altered similarly independent of drug treatment.

Tissues specificity
We study the relationship of universal drug treatments over individual tissues. In order to see this, we next investigate $u_{\ell_1 j}$ attributed to 24 tissues. Then we have found that several $u_{\ell_1 j}$ are expressive in a tissues-group wide manner (Fig. 4) . It is obvious that the combination of tissue specificity is quite biologically reasonable.

Gene selection
Then we next try to specify singular value vectors attributed to genes, $u_{\ell_4 i}$, that is used for gene selection. In order that, we check which $G(\ell_1, 2, 1, \ell_4)$ and $G(\ell_1, 3, 1, \ell_4)$ have larger absolute values since u_{1m} always exhibits same values between two replicates (Table 1).

For $\ell_1 = 2$, which is supposed to be attributed to neuron specific tissues (u_{2j}), Gs with $\ell_4 = 2$ have larger absolute values. Then u_{2i} is decided to be employed for neuron specific gene selection. For $\ell_1 = 4$, which is supposed to be attributed to muscle specific tissues (u_{4j}), Gs with $\ell_4 = 4$ have larger absolute values. Then u_{4i} is decided to be employed for muscle specific gene selection. For $\ell_1 = 5$, which is supposed to be attributed to gastrointestinal specific tissues (u_{5j}), Gs with $\ell_4 = 5$ have larger absolute values. Then u_{5i} is decided to be employed for muscle specific gene selection. For $\ell_1 = 6$, which is also supposed to be attributed to gastrointestinal specific tissues (u_{6j}), Gs with $\ell_4 = 6, 7$ have larger absolute values. Then u_{6i} and u_{7i} are decided to be employed for muscle specific gene selection.

After computing adjusted P-values, P_i, attributed to genes (see methods), genes associated with adjusted P_i less than 0.01 are selected (Table 2. See also supporting information for lists of selected genes. See Additional file 1). Figure 5 shows the Venn diagram of selected genes. As expected, two sets of genes, Gas1 and Gas2, that are supposed to be gastrointestinal specific are quite common. Other than that, the selected genes are quite distinct with one another. Thus, TD based unsupervised FE successfully identified genes whose expression is affected by drugs in tissues-group specific manner.

Confirmation of differential expression
At first, since we need to check whether selected genes are expressive distinctly between specified tissues and other tissues as well as between drug treatments and controls, we apply statistical test to selected genes (Table 2). It is obvious that for all cases, gene expression is distinct between specified tissues and other tissues as well as between drug treatments and controls. Thus TD based unsupervised FE allowed us to select genes whose expression is coincident with $u_{\ell_2 k}$s in Fig. 3 and $u_{\ell_1 j}$s in Fig. 4.

Biological evaluation
Next, we would like to evaluate selected genes biologically. For this purpose, we first upload genes to Metascape (Fig. 6). At first, we can notice that Gas1 and Gas2 largely share the enriched terms as expected, in spite of that these two gene sets are selected using distinct singular values (u_{5i} and u_{6i}, u_{7i}, respectively). Especially, it is important that two KEGG terms, "mmu04971: Gastric acid secretion" and "mmu04972: Pancreatic secretion" are shared with Gas1 and Gas2, which are supposed to be Pancreas and Stomach specific. On the other hand, various muscle related terms are enriched in Muscle gene set as expected while "GO:0002088: lens development in camera-type eye" is enriched in Neuron gene set. All of these suggest that TD based unsupervised FE selected biologically reasonable genes.

Figure 7 shows the protein protein interaction (PPI) network provided by Metascape. It is obvious that they are highly connected. Thus, TD based unsupervised FE identified sets of genes among within which PPI is enriched. Moreover, Gas1 and Gas2 largely share the PPI network while Neuron and Muscle genes sets form their own PPI network within which PPI is enriched. Thus, from the view point of PPI, TD based unsupervised FE identified biologically reasonable genes.

On might wonder if it is occasional because of one specific data set, Metascape. In order to confirm that the biological reliability is not specific to Metascape, we

uploaded genes selected by TD based unsupervised FE to Enrichr (Table 3). It is obvious that at least one tissue related diseases are detected for four sets of tissue specific genes. Thus biological reliability observed when the sets of genes were uploaded to Metascape is unlikely data set specific but more universal.

Discussion

Although one might wonder why as many as 15 drugs share the genes whose expression is altered by drug treatments, detailed investigation can allow us to interpret it. Table 4 reports the previous drug effects on neuron, muscle and pancreas. It suggests that most of drugs simultaneously have effects on these three groups of tissues.

This is very contrast to the original study [14] where the authors employed fully supervised approach that requires previous knowledge. Although Kosawa et al. [14] also tried to infer human therapy and side effects of drug treatments from gene expression of drug treated tissues, they needed pre-knowledge that we do not need for TD based unsupervised FE. In this sense, our approach has distinct potential that the original study could not achieve.

In addition to the above mentioned biological superiority of TD based unsupervised FE, it also has some methodological superiority as follows. First of all, although we classify 24 tissues into two groups based upon the observation of singular value vectors attributed to tissues, $u_{\ell_1 j}$ (Fig. 4) prior to the identification of differentially expressed genes, it is impossible for other methods to classify 24 tissues into two groups before starting to seek differentially expressed genes, since there are no criterion on how to divide 24 tissues into two groups. It is practically impossible to try all the divisions because of two many possible divisions (more than millions). The same superiority exists on grouping 18 drug treatments into two. One might wonder if it is much easier than classifying tissues, since some of drug treatments are obviously controls. Nevertheless, based upon the second and the third singular values vectors attributed to drug treatments, u_{2k} and u_{3k} (Fig. 3), acetaminophen (APAP) and Sofosbuvir are grouped together with two control treatments. Such a classification can never be proposed without TD.

As least, because of these above mentioned two advantages, TD based unsupervised FE is worthwhile trying, since it might give us completely distinct outcomes that other supervised methods can give us.

Conclusions

In this paper, we applied TD based unsupervised FE [15] to gene expression profiles of 24 mice tissues treated with 15 drugs. Integrated analysis allowed us to identify universal nature of drug treatments in tissues-group wide manner, which is generally impossible to identify with any other supervised strategy that require pre-knowledge.

Declarations

The authors declare that they have no competing interests. All data used in this data were obtained from GEO. This study is supported by KAKENHI 19H05270 that also covered the publication charge. This project was also funded by the Deanship of Scientific Research (DSR) at King Abdulaziz University, Jeddah, under grant no. KEP-8-611-38. The authors, therefore, acknowledge with thanks DSR for technical and financial support. All authors have read and approved the final manuscript. YHT planned and performed the study. YHT and TT discussed the results and wrote the paper.

Author details
[1]Department of Physics, Chuo University, 1-13-27 Kasuga, Bunkyo-ku, Tokyo 112-8551, JAPAN. [2]King Abdulaziz University, Department of Computer Science, 21589 Jeddah, Saudi Arabia.

References
1. Duran-Frigola, M., Pauls, E., Guitart-Pla, O., Bertoni, M., Alcalde, V., Amat, D., Juan-Blanco, T., Aloy, P.: Extending the small molecule similarity principle to all levels of biology. Nature Biotechnology (2020). doi:10.1101/745703. in press
2. Smirnov, P., Kofia, V., Maru, A., Freeman, M., Ho, C., El-Hachem, N., Adam, G.-A., Ba-alawi, W., Safikhani, Z., Haibe-Kains, B.: PharmacoDB: an integrative database for mining in vitro anticancer drug screening studies. Nucleic Acids Research **46**(D1), 994–1002 (2017). doi:10.1093/nar/gkx911. https://academic.oup.com/nar/article-pdf/46/D1/D994/23162705/gkx911.pdf
3. Chengalvala, M.V., , Chennathukuzhi, V.M., Johnston, D.S., Stevis, P.E., Kopf, G.S.: Gene expression profiling and its practice in drug development. Current Genomics **8**(4), 262–270 (2007). doi:10.2174/138920207781386942
4. Bates, S.: The role of gene expression profiling in drug discovery. Current Opinion in Pharmacology **11**(5), 549–556 (2011). doi:10.1016/j.coph.2011.06.009. Anti-infectives/ New technologies
5. Huang, C.-T., Hsieh, C.-H., Chung, Y.-H., Oyang, Y.-J., Huang, H.-C., Juan, H.-F.: Perturbational gene-expression signatures for combinatorial drug discovery. iScience **15**, 291–306 (2019). doi:10.1016/j.isci.2019.04.039
6. Lee, B.K.B., Tiong, K.H., Chang, J.K., Liew, C.S., Rahman, Z.A.A., Tan, A.C., Khang, T.F., Cheong, S.C.: DeSigN: connecting gene expression with therapeutics for drug repurposing and development. BMC Genomics **18**(S1) (2017). doi:10.1186/s12864-016-3260-7
7. Kim, I.-W., Jang, H., Kim, J.H., Kim, M.G., Kim, S., Oh, J.M.: Computational drug repositioning for gastric cancer using reversal gene expression profiles. Scientific Reports **9**(1) (2019). doi:10.1038/s41598-019-39228-9
8. De Wolf, H., Cougnaud, L., Van Hoorde, K., De Bondt, A., Wegner, J.K., Ceulemans, H., Göhlmann, H.: High-throughput gene expression profiles to define drug similarity and predict compound activity. ASSAY and Drug Development Technologies **16**(3), 162–176 (2018). doi:10.1089/adt.2018.845. PMID: 29658791. https://doi.org/10.1089/adt.2018.845
9. Hodos, R., Zhang, P., Lee, H.-C., Duan, Q., Wang, Z., Clark, N.R., Ma'ayan, A., Wang, F., Kidd, B., Hu, J., Sontag, D., Dudley, J.: Cell-specific prediction and application of drug-induced gene expression profiles. In: Biocomputing 2018. WORLD SCIENTIFIC, ??? (2017). doi:10.1142/9789813235533_0004. https://doi.org/10.1142/9789813235533_0004
10. Liu, X., Zeng, P., Cui, Q., Zhou, Y.: Comparative analysis of genes frequently regulated by drugs based on connectivity map transcriptome data. PLOS ONE **12**(6), 1–13 (2017). doi:10.1371/journal.pone.0179037
11. Pabon, N.A., Xia, Y., Estabrooks, S.K., Ye, Z., Herbrand, A.K., Evelyn, S., Biondi, R.M., Assimon, V.A., Gestwicki, J.E., Brodsky, J.L., Camacho, C.J., Bar-Joseph, Z.: Predicting protein targets for drug-like compounds using transcriptomics. PLOS Computational Biology **14**(12), 1–24 (2018). doi:10.1371/journal.pcbi.1006651
12. Lukačišin, M., Bollenbach, T.: Emergent gene expression responses to drug combinations predict higher-order drug interactions. Cell Systems **9**(5), 423–4333 (2019). doi:10.1016/j.cels.2019.10.004
13. 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) (2019). doi:10.1186/s12859-018-2395-8
14. Kozawa, S., Sagawa, F., Endo, S., Almeida, G.M.D., Mitsuishi, Y., Sato, T.N.: Predicting human clinical outcomes using mouse multi-organ transcriptome. iScience, 100791 (2020). doi:10.1016/j.isci.2019.100791
15. Taguchi, Y.-h.: Unsupervised Feature Extraction Applied to Bioinformatics: A PCA Based and TD Based Approach. Springer, Switzerland (2020). doi:10.1007/978-3-030-22456-1. https://doi.org/10.1007/978-3-030-22456-1
16. Burgos, K., Malenica, I., Metpally, R., Courtright, A., Rakela, B., Beach, T., Shill, H., Adler, C., Sabbagh, M., Villa, S., Tembe, W., Craig, D., Van Keuren-Jensen, K.: Profiles of extracellular mirna in cerebrospinal fluid and serum from patients with alzheimer's and parkinson's diseases correlate with disease status and features of pathology. PLOS ONE **9**(5), 1–20 (2014). doi:10.1371/journal.pone.0094839
17. Kuleshov, M.V., Jones, M.R., Rouillard, A.D., Fernandez, N.F., Duan, Q., Wang, Z., Koplev, S., Jenkins, S.L., Jagodnik, K.M., Lachmann, A., McDermott, M.G., Monteiro, C.D., Gundersen, G.W., Ma'ayan, A.: Enrichr: a comprehensive gene set enrichment analysis web server 2016 update. Nucleic Acids Research **44**(W1), 90–97 (2016). doi:10.1093/nar/gkw377. https://academic.oup.com/nar/article-pdf/44/W1/W90/18788036/gkw377.pdf
18. Zhou, Y., Zhou, B., Pache, L., Chang, M., Khodabakhshi, A.H., Tanaseichuk, O., Benner, C., Chanda, S.K.: Metascape provides a biologist-oriented resource for the analysis of systems-level datasets. Nature Communications **10**(1) (2019). doi:10.1038/s41467-019-09234-6
19. Oliveira, J.R.M., Oliveira, M.F.: Primary brain calcification in patients undergoing treatment with the biphosphanate alendronate. Scientific Reports **6**(1) (2016). doi:10.1038/srep22961
20. Harada, A., Ito, S., Matsui, Y., Sakai, Y., Takemura, M., Tokuda, H., Hida, T., Shimokata, H.: Effect of alendronate on muscle mass: Investigation in patients with osteoporosis. Osteoporosis and Sarcopenia **1**(1), 53–58 (2015). doi:10.1016/j.afos.2015.07.005
21. Hung, W.Y.: Contemporary review of drug-induced pancreatitis: A different perspective. World Journal of Gastrointestinal Pathophysiology **5**(4), 405 (2014). doi:10.4291/wjgp.v5.i4.405
22. Ghanem, C.I., Pérez, M.J.., Manautou, J., Mottino, A.D.: Acetaminophen from liver to brain: New insights into drug pharmacological action and toxicity. Pharmacological Research **109**, 119–131 (2016). doi:10.1016/j.phrs.2016.02.020. Country in focus: Pharmacology in Argentina
23. Trappe, T.A., Carroll, C.C., Dickinson, J.M., LeMoine, J.K., Haus, J.M., Sullivan, B.E., Lee, J.D., Jemiolo, B.,

Weinheimer, E.M., Hollon, C.J.: Influence of acetaminophen and ibuprofen on skeletal muscle adaptations to resistance exercise in older adults. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology **300**(3), 655–662 (2011). doi:10.1152/ajpregu.00611.2010. PMID: 21160058. https://doi.org/10.1152/ajpregu.00611.2010

24. Chen, S.-J., Lin, C.-S., Hsu, C.-W., Lin, C.-L., Kao, C.-H.: Acetaminophen poisoning and risk of acute pancreatitis. Medicine **94**(29), 1195 (2015). doi:10.1097/md.0000000000001195

25. Myrick, H., Li, X., Randall, P.K., Henderson, S., Voronin, K., Anton, R.F.: The effect of aripiprazole on cue-induced brain activation and drinking parameters in alcoholics. Journal of Clinical Psychopharmacology **30**(4), 365–372 (2010). doi:10.1097/jcp.0b013e3181e75cff

26. Kiraly, B., Gunning, K.: A case of pancreatitis associated with aripiprazole in the absence of hyperglycemia. The Primary Care Companion to The Journal of Clinical Psychiatry **10**(06), 484–485 (2008). doi:10.4088/pcc.v10n0612e

27. Elsworth, J.D., Groman, S.M., Jentsch, J.D., Valles, R., Shahid, M., Wong, E., Marston, H., Roth, R.H.: Asenapine effects on cognitive and monoamine dysfunction elicited by subchronic phencyclidine administration. Neuropharmacology **62**(3), 1442–1452 (2012). doi:10.1016/j.neuropharm.2011.08.026. Schizophrenia

28. Huo, X., Reyes, T.M., Heijnen, C.J., Kavelaars, A.: Cisplatin treatment induces attention deficits and impairs synaptic integrity in the prefrontal cortex in mice. Scientific Reports **8**(1) (2018). doi:10.1038/s41598-018-35919-x

29. Sakai, H., Sagara, A., Arakawa, K., Sugiyama, R., Hirosaki, A., Takase, K., Jo, A., Sato, K., Chiba, Y., Yamazaki, M., Matoba, M., Narita, M.: Mechanisms of cisplatin-induced muscle atrophy. Toxicology and Applied Pharmacology **278**(2), 190–199 (2014). doi:10.1016/j.taap.2014.05.001

30. Yadav, Y.C.: Effect of cisplatin on pancreas and testies in wistar rats: biochemical parameters and histology. Heliyon **5**(8), 02247 (2019). doi:10.1016/j.heliyon.2019.e02247

31. Li, C.H., Stratford, R.E., de Mendizabal, N.V., Cremers, T.I., Pollock, B.G., Mulsant, B.H., Remington, G., Bies, R.R.: Prediction of brain clozapine and norclozapine concentrations in humans from a scaled pharmacokinetic model for rat brain and plasma pharmacokinetics. Journal of Translational Medicine **12**(1) (2014). doi:10.1186/1479-5876-12-203

32. Reznik, I., Volchek, L., Mester, R., Kotler, M., Sarova-Pinhas, I., Spivak, B., Weizman, A.: Myotoxicity and neurotoxicity during clozapine treatment. Clinical Neuropharmacology **23**(5), 276–280 (2000). doi:10.1097/00002826-200009000-00007

33. Bergemann, N., Ehrig, C., Diebold, K., Mundt, C., v. Einsiedel, R.: Asymptomatic pancreatitis associated with clozapine. Pharmacopsychiatry **32**(02), 78–80 (1999). doi:10.1055/s-2007-979197

34. Lucchetti, J., Fracasso, C., Balducci, C., Passoni, A., Forloni, G., Salmona, M., Gobbi, M.: Plasma and brain concentrations of doxycycline after single and repeated doses in wild-type and app23 mice. Journal of Pharmacology and Experimental Therapeutics (2018). doi:10.1124/jpet.118.252064. http://jpet.aspetjournals.org/content/early/2018/11/05/jpet.118.252064.full.pdf

35. Bendeck, M.P., Conte, M., Zhang, M., Nili, N., Strauss, B.H., Farwell, S.M.: Doxycycline modulates smooth muscle cell growth, migration, and matrix remodeling after arterial injury. The American Journal of Pathology **160**(3), 1089–1095 (2002). doi:10.1016/s0002-9440(10)64929-2

36. Rawla, P., Raj, J.P.: Doxycycline-induced acute pancreatitis: A rare adverse event. Gastroenterology Research **10**(4) (2017)

37. Hayden, M.R., Grant, D.G., Aroor, A.R., DeMarco, V.G.: Empagliflozin ameliorates type 2 diabetes-induced ultrastructural remodeling of the neurovascular unit and neuroglia in the female db/db mouse. Brain Sciences **9**(3) (2019). doi:10.3390/brainsci9030057

38. Jordan, J., Tank, J., Heusser, K., Heise, T., Wanner, C., Heer, M., Macha, S., Mattheus, M., Lund, S.S., Woerle, H.J., Broedl, U.C.: The effect of empagliflozin on muscle sympathetic nerve activity in patients with type ii diabetes mellitus. Journal of the American Society of Hypertension **11**(9), 604–612 (2017). doi:10.1016/j.jash.2017.07.005

39. Kishimoto, M., Yamaoki, K., Adachi, M.: Combination therapy with empagliflozin and insulin results in successful glycemic control: A case report of uncontrolled diabetes caused by autoimmune pancreatitis and subsequent steroid treatment. Case Reports in Endocrinology **2019**, 1–8 (2019). doi:10.1155/2019/9415347

40. Rollin-Sillaire, A., Delbeuck, X., Pollet, M., Mackowiak, M.-A., Lenfant, P., Noel, M.-P., Facon, T., Leleu, X., Pasquier, F., Rhun, E.L.: Memory loss during lenalidomide treatment: a report on two cases. BMC Pharmacology and Toxicology **14**(1) (2013). doi:10.1186/2050-6511-14-41

41. Reece, D., Kouroukis, C.T., LeBlanc, R., Sebag, M., Song, K., Ashkenas, J.: Practical approaches to the use of lenalidomide in multiple myeloma: A canadian consensus. Advances in Hematology **2012**, 1–14 (2012). doi:10.1155/2012/621958

42. Ullenhag, G.J., Mozaffari, F., Broberg, M., Mellstedt, H., Liljefors, M.: Clinical and immune effects of lenalidomide in combination with gemcitabine in patients with advanced pancreatic cancer. PLOS ONE **12**(1), 1–19 (2017). doi:10.1371/journal.pone.0169736

43. Yasui-Furukori, N.: Update on the development of lurasidone as a treatment for patients with acute schizophrenia. Drug Design, Development and Therapy, 107 (2012). doi:10.2147/dddt.s11180

44. Anwar, I.J., Miyata, K., Zsombok, A.: Brain stem as a target site for the metabolic side effects of olanzapine. Journal of Neurophysiology **115**(3), 1389–1398 (2016). doi:10.1152/jn.00387.2015. PMID: 26719086. https://doi.org/10.1152/jn.00387.2015

45. Keyal, N., Shrestha, G., Pradhan, S., Maharjan, R., Acharya, S., Marhatta, M.: Olanzapine overdose presenting with acute muscle toxicity. International Journal of Critical Illness and Injury Science **7**(1), 69–71 (2017). doi:10.4103/2229-5151.201962. http://www.ijciis.org/article.asp?issn=2229-5151;year=2017;volume=7;issue=1;spage=69;epage=71;aulast=Keyal;t=6

46. Kerr, T.A., Jonnalagadda, S., Prakash, C., Azar, R.: Pancreatitis following olanzapine therapy: A report of three cases. Case Reports in Gastroenterology **1**(1), 15–20 (2007). doi:10.1159/000104222

47. Nissen, S.E., Stroes, E., Dent-Acosta, R.E., for the GAUSS-3 Investigators: Efficacy and Tolerability of Evolocumab vs Ezetimibe in Patients With Muscle-Related Statin Intolerance: The GAUSS-3 Randomized Clinical Trial. JAMA **315**(15), 1580–1590 (2016). doi:10.1001/jama.2016.3608. https://jamanetwork.com/journals/jama/articlepdf/2511043/joi160031.pdf

48. Raja, V., Sandanshiv, P., Neugebauer, M.: Risedronate induced transient ocular myasthenia. Journal of Postgraduate Medicine **53**(4), 274 (2007). doi:10.4103/0022-3859.37525

49. Badayan, I., Cudkowicz, M.E.: Profound muscle weakness and pain after one dose of actonel. Case Reports in Medicine **2009**, 1–3 (2009). doi:10.1155/2009/693014

50. Vinogradova, Y., Coupland, C., Hippisley-Cox, J.: Exposure to bisphosphonates and risk of common non-gastrointestinal cancers: series of nested case–control studies using two primary-care databases. British Journal of Cancer **109**(3), 795–806 (2013). doi:10.1038/bjc.2013.383

51. Salman, A.G.: Ocular surface changes with sofosbuvir in egyptian patients with hepatitis c virus infection. Cornea **35**(3), 323–328 (2016). doi:10.1097/ico.0000000000000736

52. Patel, S., Trakroo, S., Sanaka, S., Qureshi, K.: Severe myositis with the use of sofosbuvir/ledipasvir for hepatitis c infection: A case of unexpected interactions. American Journal of Gastroenterology **110**, 333 (2015). doi:10.14309/00000434-201510001-00761

53. Margapuri, J., Jubbal, S.: 902: Acute pancreatitis following treatment with ledipasvir/sofosbuvir for hepatitis virus infection. Critical Care Medicine **47**, 430 (2019). doi:10.1097/01.ccm.0000551651.45733.80

54. Kakaria, P.J., Nashel, D.J., Nylen, E.S.: Debilitating Muscle Cramps after Teriparatide Therapy. Annals of Internal Medicine **142**(4), 310–310 (2005). doi:10.7326/0003-4819-142-4-200502150-00023. https://annals.org/acp/content_public/journal/aim/20083/0000605-200502150-00016.pdf

Figures

Figure 1 Schematic figure of flow chart of analysis performed in this study

Figure 2 Summary of results obtained in this study

Figure 3 Scatter plot between u_{2k} and u_{3k} attributed to drug treatment. Red and blue arrow represent distinction controls and drug treatment and diversity among drug treatment, respectively. (1) Alendronate, (2) APAP, (3) Aripiprazole, (4) Asenapine, (5) Cisplatin, (6) Clozapine, (7) Dox, (8) EMPA, (9) FivePercentSucrose, (10) Lenalidomide, (11) Lurasidone, (12) Olanzapine, (13) Repatha, (14) Risedronate, (15) Sofosbuvir, (16) Teriparatide, (17) WT.No.treated, (18) 5percentCMC0.25percentTween80.

Figure 4 Singular value vectors, $u_{\ell_1 j}$, attributed to tissues. u_{1j}: no tissue specificity. u_{2j}: Brain Eye, Pituitary and Testis, thus mostly neuron specific. u_{3j}: Parotid specific, u_{4j}: Heart and SkMuscle, thus muscle specific, u_{5j} and u_{6j}: stomach and pancreas, thus, gastrointestinal specific.

Figure 5 Venn diagram of genes selected by TD based unsupervised FE. Neuron: genes associated with u_{2j}, which is supposed to be neuron specific. Muscle: genes associated with u_{4j}, which is supposed to be muscle specific. Gas1 and Gas2: genes associated with u_{5j} and u_{6j} respectively, which is supposed to be gastrointestinal specific.

Figure 6 Heatmap of enrichment analysis providede by Metascape.

Figure 7 PPI network provided by Metascape. Red:Gas1, Blue:Gas2, Green:Muscle, Purple:Neuron.

Tables

Table 1 $G(\ell_1, 2, 1, \ell_4)$ and $G(\ell_1, 3, 1, \ell_4)$ for $\ell_1 = 2, 4, 5, 6$. Values in bold correspond to those of ℓ_4s used for gene selection with $u_{\ell_4 i}$.

ℓ_1	2		4	
ℓ_4	$G(2,2,1,\ell_4)$	$G(2,3,1,\ell_4)$	$G(4,2,1,\ell_4)$	$G(4,3,1,\ell_4)$
1	131.248442	19.7819438	-98.4349019	-13.498228
2	**-173.243689**	**-23.9915660**	-4.8528076	1.113899
3	-11.859736	-3.2551088	-0.1595594	-1.116396
4	13.669561	2.4373120	**-81.3734282**	**36.838277**
5	26.610843	-0.3136913	-22.2440356	9.820737
6	-1.275395	4.5339065	-1.3753621	-5.318282
7	-18.306263	15.9791077	21.3673134	-11.230437
8	20.891762	26.5918473	3.9733331	-7.152480
9	21.836494	16.1461476	9.2972447	2.232529
10	11.717415	-12.8960548	1.4137802	-7.748038
ℓ_1	5		6	
ℓ_4	$G(5,2,1,\ell_4)$	$G(5,3,1,\ell_4)$	$G(6,2,1,\ell_4)$	$G(6,3,1,\ell_4)$
1	97.897860	-42.9481806	72.181307	27.3218396
2	9.267391	4.5503920	3.780984	6.3881436
3	-3.744432	0.2003586	2.340165	-0.2130656
4	1.648558	3.4031386	-9.812308	2.8751439
5	**93.027741**	**-56.9322793**	6.435061	8.5776220
6	-57.463765	23.2247109	**-19.332916**	**34.1868710**
7	28.276681	-26.9479131	**30.604535**	**-18.8319412**
8	12.884351	-13.8270607	1.798188	-10.6484624
9	-5.865058	1.0216563	9.581512	0.3507831
10	15.683762	3.7893181	-14.429706	-4.7985105

Table 2 Statistical tests for distinct expression between specified tissues and other tissues as well as that between drug treatments and controls.

				P values by statistical tests		
				Tissues	Drug treatment	
ℓ_1	tissue specificity	# of genes	Specified tissues	t test	Wilcoxon test	t test
2	Neuron	18	Brain, Eye, Pituitary, Testis	2.14×10^{-24}	9.65×10^{-49}	0.22
4	Muscle	51	Heart, SkMuscle	1.99×10^{-55}	2.67×10^{-77}	0.04
5	Gastrointestinal	97	Pancreas, Stomach	8.48×10^{-11}	2.73×10^{-40}	8.13×10^{-22}
6		128		6.67×10^{-8}	8.69×10^{-90}	8.69×10^{-90}

Table 3 Enrichment analysis for "Disease Perturbations from GEO up" and "Disease Perturbations from GEO down" by Enrichr. Diseases in bold correspond to those related to specified tissues. Up to top 10 ranked terms are shown.

Disease Perturbations from GEO up			
Term	Overlap	P-value	Adjusted P-value
Neuron specific genes			
amyotrophic lateral sclerosis DOID-332 mouse GSE3343 sample 685	5/138	1.16×10^{-7}	9.72×10^{-5}
Retinitis Pigmentosa C0035334 mouse GSE128 sample 33	5/338	9.57×10^{-6}	4.01×10^{-2}
Muscle specific			
Polycystic Ovary Syndrome C0032460 human GSE6798 sample 292	21/306	2.87×10^{-25}	2.41×10^{-22}
polycystic ovary syndrome DOID-11612 human GSE8157 sample 880	21/325	1.03×10^{-24}	4.32×10^{-22}
Insulin Resistance DOID-9352 human GSE36297 sample 581	20/290	4.52×10^{-24}	1.26×10^{-21}
Neurogenic Muscular Atrophy C0270948 rat GSE2566 sample 396	18/208	1.98×10^{-23}	4.15×10^{-21}
nemaline myopathy DOID-3191 mouse GSE3384 sample 976	15/150	1.65×10^{-20}	2.77×10^{-18}
psoriasis DOID-8893 mouse GSE27628 sample 822	18/346	2.06×10^{-19}	2.88×10^{-17}
Nemaline Myopathy C0206157 mouse GSE3384 sample 160	16/276	5.20×10^{-18}	6.24×10^{-16}
Muscular Dystrophy C0026850 mouse GSE2507 sample 405	16/278	5.84×10^{-18}	6.13×10^{-16}
cystic fibrosis DOID-1485 mouse GSE3100 sample 1057	17/344	5.91×10^{-18}	5.51×10^{-16}
COPD - Chronic obstructive pulmonary disease C0024117 human GSE475 sample 343	16/289	1.09×10^{-17}	9.11×10^{-16}
Disease Perturbations from GEO down			
Term	Overlap	P-value	Adjusted P-value
Gas1 genes			
pancreatitis DOID-4989 mouse GSE3644 sample 513	36/238	9.21×10^{-45}	7.73×10^{-42}
skin squamous cell carcinoma DOID-3151 human GSE2503 sample 627	37/373	5.24×10^{-39}	2.20×10^{-36}
pancreatic ductal adenocarcinoma DOID-3498 mouse GSE53659 sample 699	26/101	1.24×10^{-38}	3.48×10^{-36}
pancreatic invasive intraductal papillary-mucinous carcinoma DOID-8150 human GSE19650 sample 610	31/248	1.21×10^{-35}	2.54×10^{-33}
cystic fibrosis DOID-1485 mouse GSE769 sample 1058	32/288	3.92×10^{-35}	6.58×10^{-33}
Acute pancreatitis C0001339 mouse GSE3644 sample 376	28/188	2.33×10^{-34}	3.26×10^{-32}
Cystic Fibrosis C0010674 mouse GSE769 sample 428	31/275	3.34×10^{-34}	4.00×10^{-32}
Chronic phase chronic myelogenous leukemia DOID-8552 human GSE5550 sample 456	30/270	7.05×10^{-33}	7.40×10^{-31}
invasive ductal carcinoma DOID-3008 human GSE21422 sample 606	31/304	8.09×10^{-33}	7.54×10^{-31}
Eczema C0013595 human GSE6012 sample 268	26/163	1.06×10^{-32}	8.87×10^{-31}
Gas2 genes			
skin squamous cell carcinoma DOID-3151 human GSE2503 sample 627	51/373	9.37×10^{-55}	7.86×10^{-52}
pancreatitis DOID-4989 mouse GSE3644 sample 513	45/238	1.14×10^{-54}	4.77×10^{-52}
systemic lupus erythematosus DOID-9074 human GSE10325 sample 691	43/210	9.36×10^{-54}	2.62×10^{-51}
systemic lupus erythematosus (SLE) DOID-9074 human GSE36700 sample 512	47/294	1.50×10^{-53}	3.15×10^{-51}
invasive ductal carcinoma DOID-3008 human GSE21422 sample 606	44/304	5.90×10^{-48}	9.90×10^{-46}
Eczema C0013595 human GSE6012 sample 268	37/163	7.47×10^{-48}	1.04×10^{-45}
Malignant Melanoma C0025202 human GSE3189 sample 117	41/250	6.75×10^{-47}	8.09×10^{-45}
Chronic phase chronic myelogenous leukemia DOID-8552 human GSE5550 sample 456	41/270	1.91×10^{-45}	2.00×10^{-43}
Sickle Cell Anemia C0002895 human GSE9877 sample 109	37/197	1.61×10^{-44}	1.50×10^{-42}
Actinic keratosis C0022602 human GSE2503 sample 350	46/429	4.69×10^{-44}	3.94×10^{-42}

Additional Files
Additional file 1 — List of genes selected by TD based unsupervised FE
List of genes shown in Table 2 (xlsx)

Table 4 Previously reported drug effects on neuron (brain and eye), muscle and pancreas tissues. (*):Reported side effects.

	Tissue types		
Drugs	Neuron	Muscle	Pancreas or Stomach
Alendronate	Brain calcification [19]	Muscle mass [20]	Pancreatitis [21]
Acetaminophen (APAP)	Brain [22]	Skeletal muscle [23]	Pancreatitis [24]
Aripiprazole	Brain Activation [25]	Muscle spasms (*)	Pancreatitis [26]
Asenapine	Cognitive and monoamine dysfunction [27]	Muscle rigidity(*)	—
Cisplatin	Prefrontal cortex [28]	Muscle atrophy [29]	Pancreas [30]
Clozapine	Brain [31]	Myotoxicity [32]	Pancreatitis [33]
Doxycycline	Brain [34]	Smooth Muscle [35]	Acute pancreatitis [36]
Empagliflozin	Neurovascular unit and neuroglia [37]	Muscle sympathetic nerve activity [38]	Pancreatitis [39]
Lenalidomide	Memory loss [40]	Muscle cramp [41]	Pancreatic cancer [42]
Lurasidone	Acute schizophrenia [43]	Muscle(*)	—
Olanzapine	Brain stem [44]	Acute muscle toxicity [45]	Pancreatitis [46]
Repatha (Evolocumab)	—	Muscle-related statin Intolerance [47]	—
Risedronate (actonel)	Ocular myasthenia [48]	Muscle weakness [49]	Gastrointestinal cancer [50]
Sofosbuvir	Ocular surface [51]	Myositis [52]	Pancreatitis [53]
Teriparatide	—	Muscle Cramp [54]	Pancreatitis (*)