TransOrGAN: An Artificial Intelligence Mapping of Rat Transcriptomic Profiles Between Organs, Ages, and Sexes

Ting Li[1], Ruth Roberts[2,3], Zhichao Liu[4], Weida Tong[1*]

Weida.Tong@fda.hhs.gov

[1] Division of Bioinformatics and Biostatistics, National Center for Toxicological Research, US Food and Drug Administration, Jefferson, AR, USA
[2] ApconiX Ltd, Alderley Park, Alderley Edge, SK10 4TG, UK
[3] University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK
[4] Integrative Toxicology, Nonclinical Drug Safety, Boehringer Ingelheim Pharmaceuticals, Inc., Ridgefield, CT 06877, U.S.A

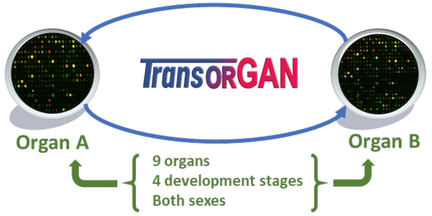
FDA U.S. FOOD & DRUG ADMINISTRATION
NATIONAL CENTER FOR TOXICOLOGICAL RESEARCH

Abstract

Animal studies are required for the evaluation of candidate drugs to ensure patient and volunteer safety. Toxicogenomics is often applied in these studies to gain an understanding of the underlying mechanisms of toxicity, which is usually focused on critical organs such as the liver or kidneys in young male rats. There is a strong ethical reason to reduce, refine and replace animal use (the 3Rs), where the mapping of data between organs, sexes and ages could reduce the cost and time of drug development. Herein, we propose a generative adversarial network (GAN)-based framework entitled TransOrGAN that allows the molecular mapping of gene expression profiles in different rodent organ systems and across sex and age groups. We carried out a proof-of-concept study based on rat RNA-seq data from 288 samples in 9 different organs of both sexes and 4 developmental stages. First, we demonstrated that TransOrGAN could infer transcriptomic profiles between any two of the 9 organs studied, yielding an average cosine similarity of 0.984 between synthetic transcriptomic profiles and their corresponding real profiles. Second, we found that TransOrGAN could infer transcriptomic profiles observed in females from males, with an average cosine similarity of 0.991. Thirdly, we found that TransOrGAN could infer transcriptomic profiles in juvenile, adult, and aged animals from adolescent animals with an average cosine similarity of 0.990, 0.991, and 0.991, respectively. Altogether, TransOrGAN is an innovative approach to infer transcriptomic profiles between ages, sexes, and organ systems, offering the opportunity to reduce animal usage and to provide an integrated assessment of toxicity in the whole organism irrespective of sex or age.

Introduction

Animal models are required for drug safety evaluation and risk assessment. Systemic toxicity considers toxicity on all organs that pose a risk. However, the multiple organ toxicological study is animal-consuming and labor-intensive. To advance the 3Rs principle, we propose TransOrGAN to infer transcriptomic profiles between different organs, ages , and sexes.



Study Design

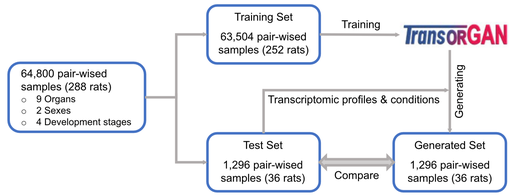


Figure 1 TransOrGAN study design. The 64,800 pairwise transcriptomic profiles between organs under different sex and developmental stage conditions were divided into training and test sets. The 63,504 pairwise transcriptomic profiles were used to develop TransOrGAN. TransOrGAN was then used to generate the 1,296 transcriptomic profiles in the test set. The generated transcriptomic profiles were then compared to the real transcriptomic profiles to evaluate the model performance.

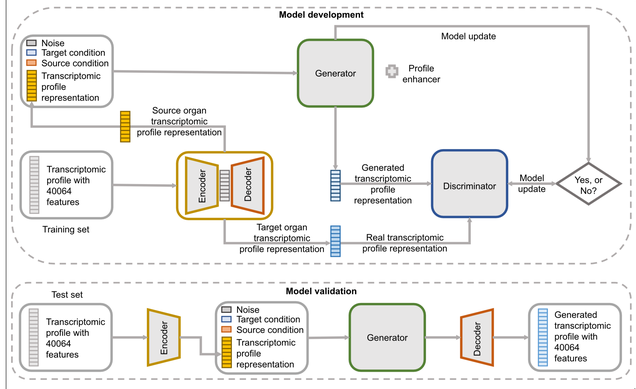


Figure 2 Overview of TransOrGAN Model. A transcriptomic profile with 40,064 genes was first fed into an encoder, and the output of the encoder (transcriptomic profile representation) combined with source organ condition (age, sex, organ, stage), target organ condition (age, sex, organ, stage), and noise as an input to the generator. The generated transcriptomic profile representation is compared with the transcriptomic profile representation of the target organ with the discriminator. The discriminator provided feedback to the generator for improvement. The process is iterative until no further updates are made, yielding the TransOrGAN model. In model validation, a transcriptomic profile with 40,064 genes was fed into the encoder, and the output of the encoder (i.e., transcriptomic profile representation) was combined with source organ condition (age, sex, organ, stage), target organ condition (age, sex, organ, stage), and noise as an input to the TransOrGAN model. The output (generated transcriptomic profile representation) was fed into the decoder to get the generated transcriptomic profile with a total 40,064 genes.

Results

1. The comparison between the generated transcriptomic profiles and the real profiles on the test set using three methods (cosine similarity, RMSE, and UMAP) demonstrated the utility of TransOrGAN in inferring transcriptomic profiles from one organ to another (Figure 3).

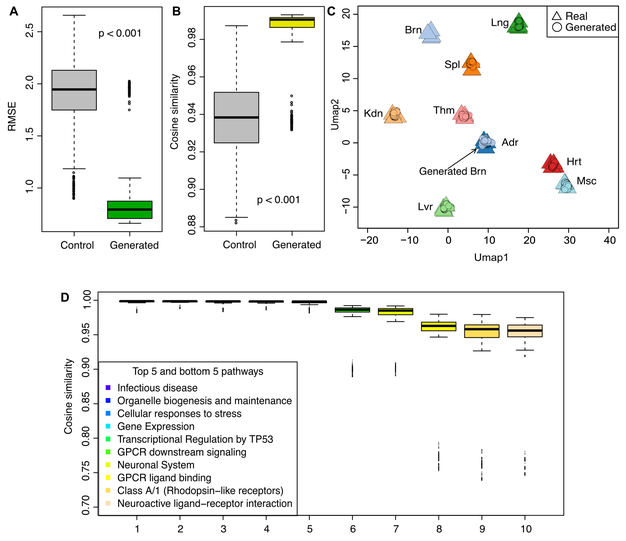


Figure 3 The performance of TransOrGAN on the test set: (A) The root mean square error (RMSE) of the control (real transcriptomic profile vs. real transcriptomic profile) and generated group (generated transcriptomic profile vs real transcriptomic profile); (B) The cosine similarity of the control and generated group; (C) The uniform manifold approximation and projection (UMAP) plot for the real and generated transcriptomic profiles; (D) The distribution of cosine similarity of the top 5 and bottom 5 among 87 toxicity pathway-related gene sets and one gene ontology term.

Conclusion

- TransOrGAN could infer transcriptomic profiles from one organ to other organs (except brain), from male to female, and from adolescent stage to younger and older ages.
- TransOrGAN offered a framework to gain understanding of toxicological mechanisms using toxicogenomics at the systemic level with testing only on a few organs and animals.

2. Higher (the higher the better) cosine similarities were observed for the transcriptomic profiles inferred from male to female (Figure 4). Similar results were observed for the transcriptomic profiles inferred from adolescent to younger and older ages (Figure 5).

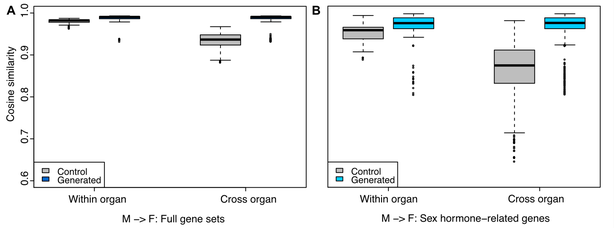


Figure 4 Translating transcriptomic profile from male to female with TransOrGAN. The cosine similarity distribution of the generated group and control on two categories: within organ and cross organ. (A) Full gene sets (i.e., 40,064 genes); (B) Sex hormone-related genes (i.e., 19 genes).

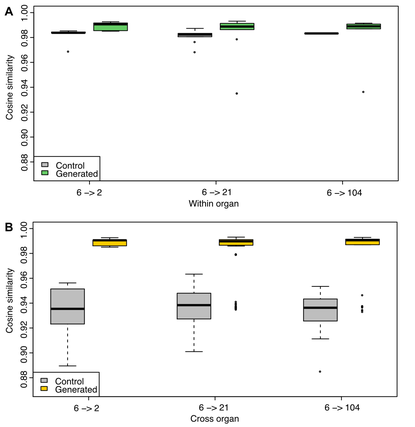


Figure 5 Translating transcriptomic profile from 6 weeks to younger and older developmental stages with TransOrGAN. The cosine similarity distribution of the generated group and control group among three categories: 6 -> 2 indicating from 6 weeks to 2 weeks; 6 -> 21 indicating from 6 weeks to 21 weeks; 6 -> 104 indicating from 6 weeks to 104 weeks. (A) Within organ (i.e., liver); (B) Cross organ (i.e., liver to kidney)

Reference: Li, Ting, Ruth Roberts, Zhichao Liu, and Weida Tong. "TransOrGAN: An Artificial Intelligence Mapping of Rat Transcriptomic Profiles between Organs, Ages, and Sexes." *Chemical Research in Toxicology* (2023).