

InterTransViewer: сравнительный анализ профилей дифференциальной экспрессии генов

Тяпкин А.¹, Омелянчук Н.^{1, 2}, Лавреха В.^{1, 2}, Землянская Е.^{1, 2*}

¹ Институт цитологии и генетики СО РАН, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

* ezemlyanskaya@bionet.nsc.ru

Ключевые слова: транскриптом; метаанализ; ауксин; этилен; *Arabidopsis thaliana* L.

Мотивация и цель: Стремительный рост количества полногеномных экспериментов по изучению изменения экспрессии генов в различных условиях обусловил широкое распространение интегрированного анализа (метаанализа) транскриптомных данных. Интеграция данных может повысить точность статистических оценок, а также позволяет тестировать гипотезы, которые невозможно было проверить в отдельных исследованиях. Для повышения информативности такой интеграции необходимо оптимизировать подбор экспериментов. В данной работе мы предлагаем набор количественных показателей для всестороннего сравнительного описания транскриптомных данных, которые легко могут быть визуализированы и интерпретированы.

Методы и алгоритмы: Предлагаемые показатели включают в себя количество дифференциально экспрессирующихся генов (ДЭГ), долю специфических (уникальных) ДЭГ в каждом наборе данных, меру попарного сходства экспериментов по составу ДЭГ, статистическую оценку однородности профилей ДЭГ. Для автоматического вычисления и визуализации этих показателей мы разработали программу InterTransViewer.

Результаты: Программа InterTransViewer реализована в виде скрипта на языке R (v.4.1.2) и доступна по ссылке (<https://github.com/al-t1/InterTransViewer0/>). Мы применили InterTransViewer для сравнительного описания транскрипционных ответов на обработку фитогормонами у модельного растения *Arabidopsis thaliana* L. Проанализированы 23 ауксин-индуцированных и 16 этилен-индуцированных профилей дифференциальной экспрессии генов. Мы показали, что комплексное рассмотрение предлагаемых характеристик позволяет позиционировать эксперименты в контексте друг друга, оценивать тенденцию к их интеграции/сегрегации, генерировать гипотезы о влиянии нецелевых факторов на исследуемый транскрипционный ответ, выделять потенциально однородные группы экспериментов, статистически оценивать однородность этих групп профилей. Это помогает принять решение о целесообразности использования данных для метаанализа.

Выводы: В целом InterTransViewer позволяет эффективно формировать выборки экспериментов в зависимости от задачи и методов метаанализа.

Финансирование: работа выполнена при поддержке Российского научного фонда, грант № 20-14-00140.

InterTransViewer: a comparative analysis of differential gene expression profiles

Tyapkin A.¹, Omelyanchuk N.^{1, 2}, Lavrekha V.^{1, 2}, Zemlyanskaya E.^{1, 2*}

¹ Institute of Cytology and Genetics, SB RAS, Novosibirsk, Russia

² Novosibirsk State University, Novosibirsk, Russia

* ezemlyanskaya@bionet.nsc.ru

Key words: transcriptome; meta-analysis; auxin; ethylene; *Arabidopsis thaliana* L.

Motivation and Aim: Meta-analysis of transcriptomic data from different experiments has become increasingly prevalent due to a significantly increasing number of genome-wide experiments investigating gene expression changes under various conditions. Data integration provides greater accuracy of statistic estimates and allows testing new hypotheses, which could not be validated in individual studies. To make data integration more informative, it is necessary to optimize the selection of experiments for the analysis. Here, we offer a set of quantitative indicators for a comprehensive comparative description of transcriptomic data, which can be easily visualized and interpreted.

Methods and Algorithms: The proposed indicators include the number of differentially expressed genes (DEGs), the proportion of experiment-specific (unique) DEGs in each data set, the measure of a pairwise similarity of experiments in DEG composition, and a statistic estimate of the homogeneity of DEG profiles. For automatic calculation and visualization of these indicators, we have developed an InterTransViewer program.

Results: InterTransViewer is implemented as an R script (v.4.1.2) and is available via the link (<https://github.com/al-t1/InterTransViewer0/>). We have used InterTransViewer to comparatively describe 23 auxin- and 16 ethylene-induced transcriptomes in *Arabidopsis thaliana* L. We have demonstrated that complex analysis of the proposed indicators allows ranking the experiments in the context of each other, assessing the tendency towards their integration or segregation, generating hypotheses about the influence of non-target factors on the transcriptional response, identifying potentially homogeneous groups of experiments, statistically estimating the homogeneity within the sets of DEG profiles. This helps to decide whether these data are appropriate for meta-analysis.

Conclusion: Overall, InterTransViewer makes it possible to efficiently select experiments for meta-analysis depending on its task and methods.

Funding: The study was supported by the RSF grant No. 20-14-00140.