

Цис-элементы, регулирующие экспрессию генов в ответ на ключевые фитогормоны у *Arabidopsis thaliana* L.

Долгих В.А.^{1, 2*}, Землянская Е.В.^{1, 2}, Левицкий В.Г.^{1, 2}

¹ Институт цитологии и генетики СО РАН, Новосибирск, Россия

² Новосибирский государственный университет, Новосибирск, Россия

* dolgikh@bionet.nsc.ru

Ключевые слова: регуляция транскрипции; *цис*-элементы; транскрипционные факторы

Мотивация и цель: Транскрипционные факторы (ТФ) управляют экспрессией генов-мишеней, связываясь с короткими участками в промоторных областях, называемых *цис*-элементами. Последовательности *цис*-элементов, с которыми связывается определенный ТФ, как правило, являются частично вырожденными, и их спектр обобщается в виде мотивов. Достаточно часто для эффективной работы ТФ необходимо формирование гетеродимеров с ТФ-партнерами. В зависимости от ТФ в составе комплекса гетеродимер может садиться на ДНК в результате взаимодействия только одного из ТФ с соответствующим *цис*-элементом или обоих ТФ с так называемым композиционным элементом, состоящим из двух *цис*-элементов, которые могут быть разделены спейсером или частично перекрываться друг с другом. Из-за комбинаторной сложности транскрипционной регуляции ее экспериментальные исследования часто ограничиваются установлением отдельных *цис*-элементов в регуляторных районах отдельных генов. Альтернативным подходом, дающим возможность системного выявления и функционального исследования *цис*-элементов, является совместный анализ полногеномных данных разного типа, описывающих разные аспекты транскрипционной регуляции экспрессии генов (например, ChIP-seq и RNA-seq). Потенциал такого подхода не реализован полностью, и потому создание соответствующего биоинформатического инструментария и его использование для системного изучения регуляции экспрессии генов являются актуальной задачей, к решению которой мы обратились в данной работе.

Методы и алгоритмы: Мы разработали конвейер SFmotif (сокращение от *structure and function of motif*), предназначенный для совместного анализа данных ChIP-seq и RNA-seq с целью комплексного изучения механизмов регуляции транскрипции. Он состоит из трех модулей, позволяет извлекать обогащенные в пиках ChIP-seq мотивы и оценивать их влияние на транскрипционную активность генов. Конвейер реализован в виде R-пакета и запускается из командной строки Linux. SFmotif интегрирует как существующие программные решения (например, Homer [1], MCOT [2]), так и оригинальные функции и модули, написанные на языке R. ChIP-seq данные по связыванию ТФ ABF1, BES1, MYC2 и ARR10, а также индуцированные этими гормонами транскриптомы были взяты из открытых источников [3, 4].

Результаты: Мы применили разработанный нами конвейер SFmotif для исследования *цис*-элементов, регулирующих транскрипционный ответ генов на четыре ключевых фитогормона: абсцизовую кислоту, брассиностероиды, жасмонаты и цитокинин. Активация их сигнальных путей приводит к активации

небольшого числа специфических ТФ, однако позволяет запускать разнообразные транскрипционные ответы в зависимости от условий. Это подразумевает наличие дополнительных механизмов, модулирующих активность соответствующих ТФ, и делает актуальной задачу их исследования. Используя ChIP-seq данные по связыванию ABF1, BES1, MYC2 и ARR10 (которые регулируют ответ на абсцизовую кислоту, брассиностероиды, жасмонаты и цитокинин соответственно), а также индуцированные этими гормонами транскриптомы, нам удалось выявить обогащение мотивов, схожих с известными сайтами связывания этих ТФ. Все они были ассоциированы с повышением экспрессии генов в ответ на действие гормонов. Помимо этих мотивов, в пиках ChIP-seq были обогащены мотивы, соответствующие, вероятно, сайтам связывания потенциальных ТФ-партнеров. Для ТФ ABF1 и MYC2 были обнаружены функциональные композиционные элементы, с которыми, вероятно, эти ТФ могут связываться в составе гетеродимеров с ТФ-партнерами. Примечательно, что классический сайт связывания ТФ ARR10 не является наиболее обогащенным, в то же время для ARR10 обнаружено наибольшее количество мотивов для потенциальных ТФ-партнеров по сравнению с другими ТФ. Это может свидетельствовать о том, что функциональность ТФ ABF1, BES1 и MYC2 регулируется в первую очередь нуклеотидным контекстом сайта связывания, а в активности ARR10 принципиальную роль играют ТФ-партнеры.

Выводы: Разработан программный конвейер SFmotif для комплексного изучения механизмов регуляции активности генов. Предсказано, что фитогормоны могут реализовывать разные стратегии транскрипционной регуляции с различным вкладом транскрипционных факторов партнеров.

Финансирование: Исследование поддержано грантом РНФ № 20-14-00140.

Cis-elements regulating gene expression in response to the key phytohormones in *Arabidopsis thaliana* L.

Dolgikh V.^{1, 2*}, Zemlyanskaya E.^{1, 2}, Levitsky V.^{1, 2}

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

² Novosibirsk State University, Novosibirsk, Russia

* dolgikh@bionet.nsc.ru

Key words: transcription regulation; *cis*-elements; transcription factors

Motivation and Aim: Transcription factors (TFs) control the expression of their target genes by binding to the short sequences in promoter regions called *cis*-elements. *cis*-elements bound by a certain TF usually have degenerate nucleotide sequences, and their whole set is usually generalized as a motif. Quite often, an effective TF functioning requires formation of heterodimers with partner TFs. Depending on the TFs in the complex, the heterodimer can bind to DNA due to an interaction of only one TF with the corresponding *cis*-element or both TFs with the so-called composite element consisting of two *cis*-elements, which can be either separated by a spacer or partially overlap with each other. Due to the combinatorial complexity of the transcriptional regulation, its experimental studies are often limited to the discovery of individual *cis*-elements in the

regulatory regions of individual genes. Alternatively, an integrated analysis of different genome-wide data, which describe different aspects of transcriptional regulation of gene expression (for example, ChIP-seq and RNA-seq) makes systematic identification and functional studies of *cis*-elements possible. The potential of this approach is far from being thoroughly realized, and therefore the development of appropriate bioinformatics tools and their use for the systematic studies of the regulation of gene expression is a relevant task, which we have addressed in this work.

Methods and Algorithms: We have developed the SFmotif pipeline (short for “structure and function of motif”), which was designed for the integrated analysis of ChIP-seq and RNA-seq data to comprehensively study the mechanisms of transcriptional regulation. It consists of three modules and allows extracting the motifs enriched in ChIP-seq peaks and evaluate their influence on the transcriptional activity of genes. The pipeline is implemented as an R package and it is launched from the Linux command line. SFmotif integrates both existing software solutions (e. g. Homer [1], MCOT [2]), as well as original functions and modules written in the R language. ChIP-seq data on the binding of TFs ABF1, BES1, MYC2 and ARR10, as well as the transcriptomes induced by these hormones were taken from publicly available repositories [3, 4].

Results: We used SFmotif pipeline to study *cis*-elements, which regulate the transcriptional response of genes to four key phytohormones: abscisic acid, brassinosteroids, jasmonates, and cytokinin. Activation of their signaling pathways leads to the activation of a limited number of specific TFs, which, however, trigger a variety of transcriptional responses depending on conditions. This implies the presence of additional mechanisms, which modulate the activity of the corresponding TFs. Therefore, investigation of these mechanisms is relevant. Using ChIP-seq data on the binding of ABF1, BES1, MYC2 and ARR10 (which regulate the responses to abscisic acid, brassinosteroids, jasmonates and cytokinin, respectively), as well as the transcriptomes induced by these hormones, we were able to identify an enrichment of motifs, which are similar to the known binding sites for these TFs. All of them were associated with the increase of gene expression in response to hormones. Additionally, other motifs were enriched in the ChIP-seq peaks, which likely correspond to the binding sites of the potential partner TFs. For ABF1 and MYC2, functional composite elements were predicted, which likely are the binding sites for these TFs being a part of heterodimers with the partner TFs. It is noteworthy that the classic TF binding site for ARR10 was not the most enriched, at the same time, ARR10 had the largest number of the motifs for the potential partner TFs compared to other TFs. This may indicate that the functionality of TFs ABF1, BES1 and MYC2 is regulated primarily by the nucleotide context of their binding site, while partner TFs play a fundamental role in the activity of ARR10.

Conclusion: SFmotif pipeline has been developed for a comprehensive study of the mechanisms of regulation of gene activity. It was predicted that phytohormones can implement different strategies of transcriptional regulation with different contributions of partner transcription factors.

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

Список литературы/References

1. Heinz S., Benner C., Spann N., Bertolino E., Lin Y.C., Laslo P., Cheng J.X., Murre C., Singh H., Glass C.K. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. *Mol Cell*. 2010;38(4):576-89. doi 10.1016/j.molcel.2010.05.004
2. Levitsky V., Zemlyanskaya E., Oshchepkov D., Podkolodnaya O., Ignatieva E., Grosse I., Mironova V., Merkulova T. A single ChIP-seq dataset is sufficient for comprehensive analysis of motifs co-occurrence with MCOT package. *Nucleic Acids Res*. 2019;47(21):e139. doi 10.1093/nar/gkz800
3. Yin L., Zander M., Huang S.C., Xie M., Song L., Saldierna Guzmán J.P., Hann E., Shanbhag B.K., Ng S., Jain S., Janssen B.J., Clark N.M., Walley J.W., Beddoe T., Bar-Joseph Z., Lewsey M.G., Ecker J.R. Transcription factor dynamics in cross-regulation of plant hormone signaling pathways. *bioRxiv*. 2023. doi 10.1101/2023.03.07.531630
4. Xie M., Chen H., Huang L., O'Neil R.C., Shokhirev M.N., Ecker J.R. A B-ARR-mediated cytokinin transcriptional network directs hormone cross-regulation and shoot development. *Nat Commun*. 2018;9(1):1604. doi 10.1038/s41467-018-03921-6