

УДК: 57.081.23

ОСОБЕННОСТИ СТРАТЕГИИ “ОТКРЫТОГО” ПОИСКА ДЛЯ ИДЕНТИФИКАЦИИ ПЕПТИДОВ В ПРОТЕОМИКЕ БЕЗ ИСПОЛЬЗОВАНИЯ МОЛЕКУЛЯРНОЙ МАССЫ**Л.И. Левицкий, Ю.А. Бубис, М.В. Иванов, И.А. Тарасова, М.В. Горшков***Институт энергетических проблем химической физики им. В.Л. Тальрозе РАН, Москва, Россия*

Одной из ключевых стадий протеомного анализа является обработка тандемных масс-спектров протеолитических пептидов для их идентификации по геномным базам данных. При этом измеряемые масс-спектры фрагментации сравниваются с теоретическими для пептидов-кандидатов, полученных при гидролизе *in silico* консенсусных белков. Выбор таких кандидатов идет по молекулярной массе, которая должна точно совпадать с измеренной. Это необходимо для сужения «пространства поиска», что повышает специфичность анализа и снижает его вычислительную сложность. Однако ценой специфичности является «потеря» пептидов, масса которых изменилась в результате неучтенных модификаций, либо мутаций в кодирующей генах, редактирования РНК и т.д.

Найти такие пептиды позволяет метод «открытого» поиска, в котором кандидаты практически не фильтруются по молекулярной массе. При этом пространство поиска расширяется на несколько порядков, что ещё недавно делало такой подход практически нереализуемым в силу потери специфичности поиска и временных затрат на обработку данных. Однако, в последнее время был достигнут существенный прогресс в производительности методов идентификации пептидов, а также в точности измерения масс ионов-фрагментов, что позволило широко использовать метод «открытого» поиска.

Основным применением подхода является идентификация пептидов с посттрансляционными и химическими модификациями без априорного знания об этих модификациях для анализируемого образца. При этом, если обычно пространство поиска расширяется экспоненциально с количеством модификаций, то «открытый» подход не влияет на это пространство. Кроме того, он применим, например, для поиска однонуклеотидного полиморфизма в протеогеомике, поскольку допустимое отклонение массы пептида в параметрах поиска превышает массу какого-либо аминокислотного остатка. Таким образом, «открытый» поиск представляет собой заманчивую альтернативу традиционным протеогеоминым подходам, не требуя привлечения персонализированных транскриптомных или геномных данных для идентификации последовательностей, отклоняющихся от консенсусных.

В докладе будут рассмотрены метод «открытого» поиска, его особенности и результаты использования в задачах полнопротеомного анализа.

Работа поддержана грантом РФФИ №18-34-00626.

Ключевые слова: протеомика, масс-спектрометрия, идентификация пептидов, открытый поиск, посттрансляционные модификации.

Литература

1. Aebersold, R., & Mann, M. (2016). Mass-spectrometric exploration of proteome structure and function. *Nature*, 537(7620), 347–355. <http://doi.org/10.1038/nature19949>
2. Chick, J. M., Kolippakkam, D., Nusinow, D. P., Zhai, B., Rad, R., Huttlin, E. L., & Gygi, S. P. (2015). A mass-tolerant database search identifies a large proportion of unassigned spectra in shotgun proteomics as modified peptides. *Nature Biotechnology*, 33(7), 743–749. <http://doi.org/10.1038/nbt.3267>
3. Tarasova, I. A., Chumakov, P. M., Moshkovskii, S. A., & Gorshkov, M. V. (2018). Profiling modifications for glioblastoma proteome using ultra-tolerant database search: Are the peptide mass shifts biologically relevant or chemically induced? *Journal of Proteomics*, (May), 0–1. <http://doi.org/10.1016/j.jprot.2018.05.010>