

© ГЛАЗЫРИН Ю. Е., ВЕПРИНЦЕВ Д. В., БЕРЕЗОВСКИЙ М. В.

УДК 577.1

DOI: 10.20333/2500136-2021-2-87-89

Применение полнопротеомного анализа клеток крови, плазмы и мочи для дифференциальной диагностики заболеваний

Ю. Е. Глазырин¹, Д. В. Вепринцев¹, М. В. Березовский²¹Федеральный исследовательский центр «Красноярский научный центр Сибирского отделения Российской академии наук», Красноярск 660036, Российская Федерация²Университет Оттавы, Оттава K1N6N5, Канада

Резюме. Рассматривается методика создания концепций новых неинвазивных дифференциальных диагностических систем на основе протеомного профилирования клеток крови, плазмы и мочи. При этом одновременно используется относительная количественная информация обо всех белках, обнаруженных в пробах. В представленных примерах методами машинного обучения стало возможным разделить по протеомным данным плазмы крови группы больных диабетической нефропатией и гломерулонефритом, а по протеомным данным мочи отделить от этих групп больных гипертонической нефропатией. С помощью регрессионного анализа удалось построить линейную модель, способную оценивать момент необходимости инициации терапии в будущем для больных на ранней стадии хронического лимфолейкоза.

Ключевые слова: протеомное профилирование, масс-спектрометрия, дифференциальная диагностика, машинное обучение, линейная регрессия.

Конфликт интересов. Авторы декларируют отсутствие явных и потенциальных конфликтов интересов, связанных с публикацией настоящей статьи.

Для цитирования: Глазырин ЮЕ, Вепринцев ДВ, Березовский МВ. Применение полнопротеомного анализа клеток крови, плазмы и мочи для дифференциальной диагностики заболеваний. *Сибирское медицинское обозрение.* 2021;(2):87-89. DOI: 10.20333/2500136-2021-2-87-89

Application of full proteomic analysis of blood cells, plasma and urine for differential diagnosis

Y. E. Glazyrin¹, D. V. Vepintsev¹, M. V. Berezivski²¹Federal Research Center "Krasnoyarsk Science Center of the Siberian Branch of the Russian Academy of Sciences", Krasnoyarsk 660036, Russian Federation²University of Ottawa, Ottawa K1N6N5, Canada

Abstract. The paper considers a methodology for creating concepts for new non-invasive differential diagnostic systems based on proteomic profiling of blood cells, plasma and urine. In this case, the relative quantitative information about all proteins found in the samples is used simultaneously. In the presented examples, using machine learning methods, it became possible to separate groups of patients with diabetic nephropathy and glomerulonephritis according to proteomic data of blood plasma, and to separate patients with hypertensive nephropathy from these groups using proteomic data of urine. With the help of regression analysis, it was possible to construct a linear model capable of assessing the moment of the need to initiate therapy in the future for patients at an early stage of chronic lymphocytic leukemia.

Key words: proteomic profiling, mass spectrometry, differential diagnosis, machine learning, linear regression.

Conflict of interest. The authors declare the absence of obvious and potential conflicts of interest associated with the publication of this article.

Citation: Glazyrin YE, Vepintsev DV, Berezovski MV. Application of full proteomic analysis of blood cells, plasma and urine for differential diagnosis. *Siberian Medical Review.* 2021; (2):87-89. DOI: 10.20333/2500136-2021-2-87-89

Введение

Дифференциальная диагностика заболеваний направлена на определение стадий развития заболевания, способствует выявлению причин заболеваний различного происхождения со сходными симптомами [1]. Жидкостная биопсия – неинвазивный метод современной диагностики, позволяющий с помощью отбора биологических жидкостей у пациента и последующего инструментального анализа получить необходимую диагностическую информацию. Протеомное профилирование с помощью масс-спектрометрии высокого разрешения позволяет получить количественные показатели относительной распространённости белков в клетках крови или биологических жидкостях организма. Присутствие и уровень

отдельных традиционных биомаркеров нередко не отражает реальной картины на разных стадиях развития заболевания, поэтому важно комплексное рассмотрение ситуации с учетом анализа всего протеома, который в полном объеме выступает как диагностический признак. Современные методы статистической обработки информации позволяют комплексно оценивать многокомпонентные системы, состоящие из нескольких сотен или тысяч белков, сравнивать их между собой в зависимости от диагноза или степени развития заболевания, выявляя диагностические закономерности.

Материал и методы

Для разработки новых неинвазивных диагностических систем на основе протеомного профилирования

биологических жидкостей необходим правильный набор первичных экспериментальных групп пациентов, имеющих точный диагноз, либо информацию о текущей стадии заболевания, полученную достоверными способами ранее (как правило, с помощью традиционной тканевой биопсии). Дальнейшие стадии разработки новой тест-системы включают отбор, накопление и хранение проб биологических жидкостей, выделение и лизис клеток из крови, деплетирование плазмы крови для удаления широко распространённых (мажорных) белков, ультрафильтрацию мочи для концентрации белков. Подготовка проб к масс-спектрометрическому анализу включает энзимное расщепление белков на пептиды и очистку образцов. Разделение сложных пептидных смесей производится с помощью ультра высокоэффективной жидкостной хроматографии. Масс-спектрометрический анализ производится с помощью приборов высокого разрешения, как правило, использующих в качестве анализатора масс высокочувствительную орбитальную ловушку. Идентификация обнаруженных в пробах белков по масс-спектрам и сравнительный количественный анализ проводится с помощью коммерческих или свободно распространяемых [2] биоинформационных программных пакетов. С помощью статистической обработки результатов протеомного профилирования, включающей регрессионный анализ и применение алгоритмов машинного обучения, возможно построение новых классифицирующих диагностических моделей, которые можно в дальнейшем использовать при интерпретации результатов протеомного анализа биоматериалов, полученных от новых пациентов, требующих подтверждения или дальнейшей классификации диагноза.

Результаты и обсуждение

Подобные работы, использующие результаты протеомного профилирования клеток, плазмы крови и мочи для разработки концепций новых дифференциальных диагностических систем с применением современных статистических методов, были недавно выполнены в КрасГМУ им. проф. В. Ф. Войно-Ясенецкого [3, 4].

Так, для разделения групп пациентов с хроническими болезнями почек (ХБП), имеющих различное происхождение, было проведено сравнительное протеомное профилирование проб мочи и плазмы крови для трех групп пациентов [3]. Дальнейшая интерпретация количественных протеомных профилей проводилась с помощью методов машинного обучения. Тестовые группы больных диабетической нефропатией и гломерулонефритом были разделены с точностью (accuracy) 96 % по результатам протеомного анализа плазмы крови с помощью метода ближайших соседей. При сравнении протеомов мочи методом «один

против всех» удалось с высокой точностью (100 % правильных решений для имеющейся тестовой группы) отделить группу больных гипертонической нефропатией от остальных больных ХБП.

Попытки сравнить реальную и прогнозную скорости развития хронического лимфолейкоза (ХЛЛ) были сделаны с помощью анализа протеомных профилей лимфоцитов из крови больных ХЛЛ, не получавших лечение на момент отбора проб [4]. С учетом полученных позже сведений о моментах начала применения терапии для этих больных, с помощью линейного регрессионного анализа была построена прогнозная модель, учитывающая одновременный вклад относительной экспрессии большого количества белков, статистически выраженной в главных компонентах (рис.). В дальнейшем, подставляя в существующую модель параметры протеомных профилей, полученные при анализе клеток крови новых больных, становится возможной оценка необходимости начала применения терапии в будущем для вновь диагностированных пациентов с начальными стадиями ХЛЛ.

Предсказанные (predicted) значения лежат в плоскости модели близко к реальным (real) образцам. Взято из работы [4].

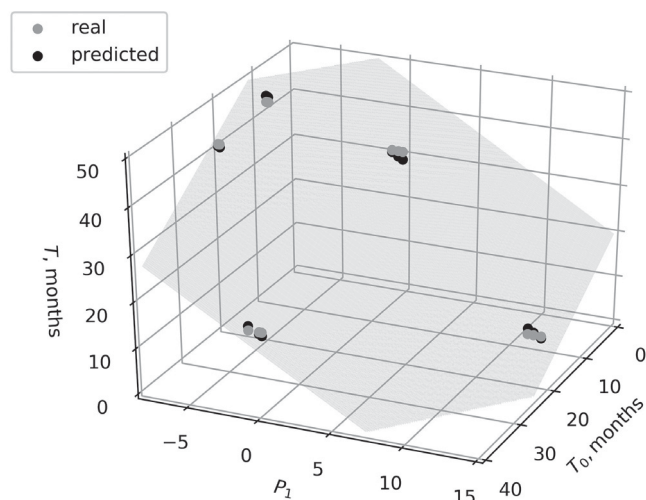


Рисунок. Плоскость регрессионной модели, описывающей зависимость моментов отбора проб (T_0) и начала применения терапии (T) от проекции вектора количественных показателей распространенности белков на первую главную компоненту (P_1).

Figure. The plane of the regression model describing the dependence of the moments of sampling (T_0) and initiation of therapy (T) on the projection of the vector of quantitative indicators of the prevalence of proteins on the first principal component (P_1).

Заключение

Представленные подходы по созданию концепций новых дифференциальных диагностических систем с помощью протеомики и статистики являются универсальными. Они применимы для выявления протеомных различий и прослеживания статистических и биологических закономерностей на уровне протеома при сравнении любых групп пациентов, которые возможно первоначально сформировать с помощью классификации альтернативными достоверными методами, чаще всего требующими более сложного подхода и/или инвазивного вмешательства.

Литература / References

1. Siegenthaler W. Differential Diagnosis in Internal Medicine: From Symptom to Diagnosis. : Thieme; 2011. 1140 p.
2. Cox J, Mann M. MaxQuant Enables High Peptide Identification Rates, Individualized p.p.b.-range Mass Accuracies and Proteome-Wide Protein Quantification. *Nature Biotechnology*. 2008; (26):1367–1372. DOI:10.1038/nbt.1511
3. Glazyrin YE, Veprintsev DV, Ler IA, Rossovskaya ML, Varygina SA, Glizer SL, Zamay TN, Petrova MM, Minic Z, Berezovski MV, Kichkailo AS. Proteomics-Based Machine Learning Approach as an Alternative to Conventional Biomarkers for Differential Diagnosis of Chronic Kidney Diseases. *International Journal of Molecular Sciences*. 2020; 21(13):4802. DOI:10.3390/ijms21134802

4. Bakhtina VI, Veprintsev DV, Zamay TN, Demko IV, Mironov GG, Berezovski MV, Petrova MM, Kichkailo AS, Glazyrin YE. Proteomics-Based Regression Model for Assessing the Development of Chronic Lymphocytic Leukemia. *Proteomes*. 2021; 9(1):3. DOI:10.3390/proteomes9010003

Сведения об авторах

Глазырин Юрий Евгеньевич, к.б.н., научный сотрудник, Красноярский научный центр СО РАН; адрес: Российская Федерация, 660036, г. Красноярск, Академгородок, стр. 50; тел.: +7(391)2201893; e-mail: yury.glazyrin@gmail.com, <http://orcid.org/0000-0002-2826-575>.

Веprinцев Дмитрий Владимирович, к.ф.-м.н., старший научный сотрудник, Красноярский научный центр СО РАН; адрес: Российская Федерация, 660036, г. Красноярск, Академгородок, стр. 50; тел.: +7(391)2907988; e-mail: d_veprintsev@mail.ru.

Березовский Максим Валентинович, PhD, профессор, Университет Оттавы; адрес: Канада, K1N6N5, г. Оттава, ул. Марии Кюри, Холл Д'Иорю; тел.: +1(613)5625800; e-mail: Maxim.Berezovski@uottawa.ca, <https://orcid.org/0000-0003-0514-599X>.

Author information

Yury E. Glazyrin, Cand. Biol. Sci., Researcher, Krasnoyarsk Scientific Center of the SB RAS; Address: 50, Akademgorodok, Krasnoyarsk, Russian Federation 660036; Phone: +7(391)2201893; e-mail: yury.glazyrin@gmail.com, <http://orcid.org/0000-0002-2826-575>.

Dmitry V. Veprintsev, Cand. Phis.-Math. Sci., Senior Researcher, Krasnoyarsk Scientific Center of the SB RAS; Address: 50, Akademgorodok, Krasnoyarsk, Russian Federation 660036; Phone: +7(391)2907988; e-mail: d_veprintsev@mail.ru.

Maxim V. Berezovski, PhD, Professor, University of Ottawa; Address: 10 Marie-Curie, D'Iorio Hall, Ottawa, Ontario, Canada K1N 6N5; Phone: +1(613)5625800; e-mail: Maxim.Berezovski@uottawa.ca, <https://orcid.org/0000-0003-0514-599X>.

Дата поступления: 16.02.2021

Дата рецензирования: 18.03.2021

Принята к печати: 31.03.2021

Received 16 February 2021

Revision Received 18 March 2021

Accepted 31 March 2021