

PROBLEMY PROTEOMIKI KLINICZNEJ – TRENDY, NIEBEZPIECZEŃSTWA I PROBLEMY*

CLINICAL PROTEOMICS – TRENDS, PITFALLS AND PROBLEMS

Jerzy SILBERRING

Zakład Neurobiochemii, Wydział Chemii Uniwersytetu Jagiellońskiego, Kraków

Streszczenie: Proteomika kliniczna zajmuje się badaniami na pograniczu chemii, biochemii i medycyny, wykorzystując nowoczesne narzędzia badawcze, jak m.in. techniki mikroekstrakcji i separacji złożonych próbek biologicznych, metody identyfikacji białek za pomocą spektrometrii mas i bioinformatykę. Na współczesną proteomikę składają się zarówno badania podstawowe, jak też badania kliniczne. Pomiędzy tymi dyscyplinami zachodzi ciągła wymiana informacji (translacja), przyczyniając się do stałego postępu wiedzy o określonych schorzeniach. Ten proces jest częścią biologii systemów, czyli zrozumienia, w jaki sposób funkcjonuje organizm człowieka. W rzeczywistości, czyli w praktyce laboratoryjnej, założenie to, jakkolwiek ambitne, napotyka na szereg trudności metodologicznych. Głównym problemem jest tu brak ogólnie przyjętej, standardowej strategii, umożliwiającej monitorowanie określonych schorzeń, a także brak kryteriów jakościowych dla tej, złożonej z szeregu etapów, strategii. Bez przyjęcia takich ogólnych norm, porównywanie wyników uzyskiwanych w różnych laboratoriach nie ma w tej sytuacji żadnego uzasadnienia. Już na etapie kolekcjonowania materiału biologicznego występuje wiele wątpliwości związanych z dynamiką przemian metabolicznych, wpływem: wieku, leków, środowiska, sposobu pobierania i przechowywania próbek itp. na końcowy wynik analiz. Dodatkowo, kilkanaście białek stanowi ponad 96% całkowitego składu białkowego surowicy, a zakres stężeń pomiędzy poszczególnymi składnikami sięga 10^{10} . Jednoczesna analiza ogromnej liczby danych przekracza możliwości standardowego oprogramowania, dostarczanego wraz z aparaturą pomiarową. Kryteria selekcji uzyskanych danych także w sposób zasadniczy wpływają na wynik końcowy, a spektakularnym przykładem może być jedna z niedawnych publikacji, w której autorzy analizowali mieszaninę 6 białek standardowych, otrzymując w końcowym etapie wydruk komputerowy, zawierający identyfikację ok. 800 białek!! Czy w świetle powyższych faktów możliwa jest w ogóle rzetelna i powtarzalna analiza?

Słowa kluczowe: proteomika kliniczna.

Summary: Clinical proteomics is an analytical strategy functioning at the edge between chemistry, biochemistry, and medicine. This approach utilizes modern and ultrasensitive techniques, such as microextraction and separation of complex biological samples, identification with the aid of mass spectrometry, and

*Niniejszy materiał jest fragmentem przygotowywanej książki (A. Kraj i J. Silberring, red.) „Proteomika i Metabolomika”, która ukaze się w 2009 r. nakładem Wydawnictwa Uniwersytetu Warszawskiego.

advanced bioinformatics. Nowadays proteomics is based on basic research and clinical investigations. Between those disciplines there is a constant exchange of information (translation) and this process is part of the systems biology, i.e. understanding how human organism works. In reality, i.e. in a daily laboratory practice this assumption, however ambitious, meets many methodological ambiguities. The major issue is a lack of the commonly accepted, standard (or optimal) analytical strategy to handle biological material, and to monitor specific diseases. There is also an urgent need to evaluate quality criteria for this multi-task approach. Without such widespread norms, data comparison between various laboratories is out of question. Many ambiguities already appear during material collection phase. This is associated with the overall dynamics of such complex biological system as living organism, influence of age, medications, environment, a way the sample is collected and stored, etc. on the final result of the entire analysis. Additionally, several proteins comprise 96% of the protein content in plasma and the dynamic range of molecules reaches nearly 10^{10} , which obscures less abundant components having important biological impact. Thus, simultaneous analysis of a huge amount of raw data is far beyond capabilities of modern instrumentation and software, and, even more important, cannot be fully interpreted by operators and translated to a common biological sense. One spectacular example may here be cited, where authors honestly reported analysis of a mixture of 6 standard proteins, which resulted in a printout containing „identification” of ca. 800 proteins! How the above contributes to the reliable and reproducible analysis?

Key words: clinical proteomics.

WSTĘP

Proteomika kliniczna zajmuje się badaniami na pograniczu chemii, biochemii i medycyny, wykorzystując nowoczesne narzędzia badawcze, jak m.in. techniki mikroekstrakcji i separacji złożonych próbek biologicznych, metody identyfikacji białek za pomocą spektrometrii mas i bioinformatykę. Współczesna proteomika obejmuje zarówno badania podstawowe, jak i badania kliniczne. Pomiędzy tymi dyscyplinami zachodzi ciągła wymiana informacji (translacja), przyczyniając się do stałego postępu wiedzy o określonych schorzeniach. Ten proces jest częścią biologii systemów, czyli zrozumienia, w jaki sposób funkcjonuje organizm człowieka.

CZY PROTEOMIKA POMOŻE NAM W POSZUKIWANIU BIOMARKERÓW?

W rzeczywistości, czyli w praktyce laboratoryjnej założenie to, jakkolwiek ambitne, napotyka na wiele trudności metodologicznych. Głównym problemem jest tu brak ogólnie przyjętej, standardowej strategii, umożliwiającej monitorowanie określonych schorzeń. Brak także kryteriów jakościowych oceniających poszczególne etapy klinicznej analizy proteomicznej. Nie jest bowiem oczywiste, co oznacza dobrze wykonana elektroforeza w żelu poliakrylamidowym i jak powinien przedstawiać się wielowymiarowy rozdział chromatograficzny? Jaki powinien być udział tła (czy w ogóle powinien być?) i w jaki sposób określić, czy wybarwione białko w żelu ma „właściwy” kształt? Bez przyjęcia takich ogólnych norm, porównywanie wyników uzyskiwanych w różnych laboratoriach nie ma w tej sytuacji żadnego uzasadnienia.

PRZYGOTOWANIE PRÓBEK, CZYLI PIERWSZY ETAP PRODUKCJI ARTEFAKTÓW

Proteomika jest znakomitym narzędziem do produkcji ogromnej ilości danych, w znacznym stopniu niedających się zweryfikować. Już na etapie kolekcjonowania materiału biologicznego występuje szereg wątpliwości, związanych z dynamiką przemian metabolicznych, wpływem wieku, leków, środowiska, sposobu pobierania i przechowywania próbek itp. na końcowy wynik analiz [1]. Dobrym przykładem materiału, który stanowi ogromne wyzwanie dla analizy proteomicznej, jest surowica krwi. Kilkanaście białek stanowi ponad 96% jej całkowitego składu białkowego, a różnica stężeń pomiędzy poszczególnymi składnikami sięga 10^{10} . Z kolei usuwanie tych kilkunastu białek za pomocą chromatografii powinowactwa jest niezwykle kosztowne, czasochłonne i najczęściej eliminuje też inne istotne składniki mieszaniny, zaburzając proporcje ilościowe. Jednoczesna analiza ogromnej liczby danych przekracza możliwości standardowego oprogramowania, dostarczanego wraz z aparaturą pomiarową. Wynika stąd jednoznacznie, że kluczowym elementem całej strategii będzie preseparacja składników mieszaniny. Od efektywności tego procesu zależeć będzie końcowy wynik. Nie ma w chwili obecnej metody, która umożliwiałaby rozdział białek z odpowiednią selektywnością, a mieszanina już tylko kilku składników nie jest prawidłowo analizowana przez oprogramowanie.

BIOINFORMATYKA – NAJSŁABSZE OGNIWO PROTEOMIKI

Kryteria selekcji uzyskanych danych także w sposób zasadniczy wpływają na wynik końcowy, a spektakularnym przykładem może być jedna z niedawnych publikacji, w której autorzy analizowali mieszaninę 6 standardowych białek, otrzymując w końcowym etapie wydruk komputerowy, zawierający identyfikację ok. 800 białek!!! Czy w świetle powyższych faktów możliwa jest w ogóle rzetelna i powtarzalna analiza?

Problemem jest tutaj nie tylko brak odpowiednich narzędzi bioinformatycznych, ale przede wszystkim słaba znajomość tej tematyki przez badaczy. Identyfikacja białka jest, przynajmniej intuicyjnie, bardziej zrozumiała niż skomplikowany warsztat informatyczny i chemometryczny. Dotyczy to jednak także drugiej strony, czyli samych bioinformatyków, nie mających najczęściej odpowiedniej wiedzy biochemicznej, aby prawidłowo ocenić sens biologiczny wydruków komputerowych. Kolejnym przykładem jest wykazywanie sporej liczby hipotetycznych białek (ang. *hypothetical proteins*), które komputer identyfikuje nie na podstawie kompletnej sekwencji aminokwasowej w bazie białek, ale na podstawie przetłumaczonej sekwencji nukleotydów w bazie genów. Nikt nie zadaje sobie pytania, czy takie białko w rzeczywistości w ogóle istnieje i czy może mieć jakąkolwiek funkcję biologiczną. Przeważająca większość oprogramowania do analizy wielowymiarowych żeli zakłada rozkład normalny i stosuje test Studenta oraz analizę wariancji (*one-way ANOVA*)

do oceny różnic pomiędzy próbkami. Zwykle taki rozkład nie jest weryfikowany przed rozpoczęciem analizy, a dodatkowo statystyka wymaga co najmniej 3–7 powtórzeń (czyli 3–7 rozdziałów KAŻDEJ próbki), aby test Studenta można było stosować [2]. Zwiększa to znacząco koszty i czasochłonność procedury. W zespole autora przeprowadzono z kolei eksperymenty podwójnego znakowania stabilnymi izotopami, połączone z techniką *shotgun*. W rezultacie wygenerowano ponad 50 000 widm fragmentacyjnych, których analiza była możliwa dopiero po opracowaniu własnego oprogramowania. Należy jednak przyjąć założenie, że eksperyment musi zostać powtórzony co najmniej trzykrotnie i to w całości.

CZY TO JEST COŚ WARTE?

Przytoczmy tu jedno zdanie z wywiadu przeprowadzonego przez czasopismo Science [4] z Amosem Bairochem, twórcą bazy SwissProt, serwera ExPASy i szeregu znakomitych narzędzi bioinformatycznych: „...*much of the data generated by proteomics groups over the past decade is junk*”. Rzeczywiście, obserwując gwałtowny rozwój laboratoriów proteomicznych, to właśnie łatwość pozyskiwania funduszy na „modne” kierunki, niedostatek fachowców, a przede wszystkim brak walidacji uzyskanych wyników spowodowały lawinową „produkcję” danych, których nie można porównać nawet w tym samym laboratorium. Trzeba też pamiętać, że metodami proteomicznymi identyfikujemy średnio ok. 20% sekwencji danego białka. Można oczywiście uznać, że jest to wystarczający dowód na istnienie określonej molekuly, ale nie jest to wartościowa informacja z punktu widzenia kompletnej struktury białka (przynajmniej I-rzędowej), bez podania modyfikacji potranslacyjnych, ew. mutacji itp. A przecież najczęściej modyfikacje odgrywają główną rolę w regulowaniu funkcji molekuł. Wszystkie stosowane obecnie techniki analizują ekspresję białka, a nie jego aktywność. Jaka jest więc relacja np. pomiędzy nadekspresją enzymu a jego aktywnością katalityczną?

Z powyższych względów istotna jest walidacja uzyskanych wyników i potencjalnych „biomarkerów”, które na tym etapie analizy żadnymi biomarkerami nie są. Procedura walidacji jest tak samo złożona jak inne, nieproteomiczne metody identyfikacji biomarkerów i tak samo kosztowna i czasochłonna. Wymagania są tu jednoznacznie sformułowane; należy opracować specyficzną metodę analizy ilościowej (np. ELISA, reakcja enzymatyczna itp.) i zweryfikować uzyskane dane porównując próbki kliniczne, wyznaczyć szereg parametrów, zakres norm itp. Są to standardowe procedury stosowane m.in. w chemii klinicznej czy chemii sądowej i powszechnie znane. Mówimy tu oczywiście wyłącznie o markerach sygnalizujących rozwój choroby, a nie o tworzeniu potencjalnych leków, co często jest mylone.

PRZYSZŁOŚĆ „OMIKI”

Pomimo przewijającego się tu pesymizmu, strategie „omiczne” mają przyszłość, po przyjęciu pewnych zastrzeżeń. Nie można każdej z technik traktować oddzielnie

a główny dogmat biologii molekularnej (gen – transkrypcja – translacja – białko) należy uznać za kierunek rozwoju metodologii, czyli prowadzić analizę całościową i nie pomijać innych dziedzin nauki. Wydaje się, że po początkowym okresie euforii zaczynają pojawiać się publikacje nawiązujące do rzeczywistych problemów i wyzwań proteomiki [5]. Nie ulega natomiast wątpliwości, że narzędzia „omiczne” nie nadają się, przynajmniej na razie, do identyfikacji biomarkerów przy łóżku pacjenta. Mogą natomiast wspomagać ten proces na etapie wstępnej analizy jako środek a nie cel całej strategii. Spektrometria mas jest tylko jednym z elementów tej procedury i z pewnością nie spełnia funkcji *black box*, do którego wprowadza się próbkę, naciska guzik i uzyskuje wynik. Sytuacji z pewnością nie ułatwia powstawanie kolejnych dziedzin „omiki”, chemogenomiki i chemoproteomiki, bazujących na odmiennych założeniach [3].

Jak dotąd nie udało się uzyskać metodami proteomicznymi żadnego wartościowego biomarkera, np. w diagnostyce nowotworów prostaty ciągle najlepszym testem pozostaje oznaczanie PSA, przy wszystkich jego ograniczeniach. Analogiczna sytuacja przedstawia się w zastosowaniach chemii kombinatorycznej, gdzie nakłady stoją w rażącej dysproporcji do uzyskiwanych efektów w przemyśle farmaceutycznym. Może więc warto przemyśleć całą koncepcję i prowadzić poszukiwania w połączeniu z innymi metodami, jak np. metaanalizy wpływu określonych leków na duże grupy pacjentów lub poszukiwanie określonych modyfikacji potranslacyjnych, co daje szansę na zawężenie obszaru badań do rozsądnego zakresu. Zanim więc dojdziemy do celu, upłynie znacznie więcej czasu niż początkowo zakładano.

LITERATURA

- [1] AHMED FE. Sample preparation and fractionation for proteome analysis and cancer biomarker discovery by mass spectrometry. *J Sep Sci* 2009; **32**: 771–798.
- [2] BIRON DG, BRUN C, LEFEVRE T, LEBARBENCHON C, LOXDALE HD, CHEVENET F, BRIZARD J-P, THOMAS F. The pitfalls of proteomics experiments without the correct use of bioinformatics tools. *Proteomics* 2006; **6**: 5577–5596.
- [3] JACOBY E. Chemogenomics – drug discovery's panacea? *Mol Bio Syst* 2006; **2**: 218–220.
- [4] SERVICE RF. Proteomics prime time. *Science* 2008; **321**: 1758–1761.
- [5] YAN L, TONACK S, SMITH R, DODD S, JENKINS RE, KITTERINGHAM N, GREENHALF W, GHANEH P, NEOPTOLEMOS JP, COSTELLO E. Confounding effects of obstructive jaundice in the interpretation of proteomic plasma profiling data for pancreatic cancer. *J Proteome Res* 2009; **8**: 142–148.

Prof. dr hab. Jerzy SILBERRING

Zakład Neurobiochemii, Wydział Chemii Uniwersytetu Jagiellońskiego

ul. R. Ingardena 3, 3, 0-060 Kraków

e-mail: silber@chemia.uj.edu.pl