

Warszawa, 23/8/2017

Dr hab. Dariusz Plewczyński, prof. UW  
Laboratorium Genomiki Funkcjonalnej i Strukturalnej  
Centrum Nowych Technologii  
Uniwersytet Warszawski  
Ul. Banacha 2c, 02-097 Warszawa, Polska

## RECENZJA

rozprawy doktorskiej pani magister Agaty Charzyńskiej

### SENSITIVITY ANALYSIS IN BIOCHEMICAL MODELS

wykonanej w Instytucie Podstaw Informatyki

Polskiej Akademii Nauk

pod kierunkiem promotora

profesor doktor habilitowanej Anny Gambin

Przedstawiona mi do recenzji praca jest owocem udanej analizy teoretycznej stabilności modeli biochemicznych. Budowa i analiza modeli dynamicznych dla systemów biologicznych stanowi część stosunkowo niedawno powstałej, wysoko interdyscyplinarnej dziedziny, jaką jest biologia obliczeniowa. Modelowanie obliczeniowe integruje dostępną wiedzę na temat regulacji ścieżek sygnałowych z danymi eksperymentalnymi dostępnymi w ramach różnych platform biotechnologicznych. Takie podejście stanowi rozwiązanie formalizujące i rozszerzające tradycyjną biologię molekularną.

Rozprawa doktorska mgr Agaty Charzyńskiej ma formę spójnej tematycznie rozprawy napisanej pod kierunkiem Prof. Dr hab. Anny Gambin (Uniwersytet Warszawski, Wydział Matematyki, Informatyki i Mechaniki) liczącej 71 stron z odniesieniami do artykułów opublikowanych przez Kandydatkę w czasopismach naukowych. Przedmiotem mojej oceny, w myśl wymagań Ustawy o stopniach naukowych i tytule naukowym oraz o stopniach i tytule w zakresie sztuki z dnia 14 marca 2003 r. (Dz.U. 2016 poz. 882, z późn. zm.) oraz Rozporządzenia Ministra Nauki i Szkolnictwa Wyższego z dnia 26 września 2016 r. w sprawie szczegółowego trybu i warunków przeprowadzania czynności w przewodzie doktorskim, w postępowaniu habilitacyjnym oraz w postępowaniu o nadanie tytułu profesora (Dz.U. 2016 poz. 1586), jest oryginalność rozwiązanego problemu naukowego, ogólna wiedza teoretyczna Kandydatki w dziedzinie nauk matematycznych w dyscyplinie informatyka, a także umiejętność samodzielnego prowadzenia pracy naukowej. Praca była rozpoczęta na wydziale Matematyki, Informatyki i Mechaniki UW a dokończona została w laboratorium dr hab. Michała Dąbrowskiego w Instytucie Biologii Doświadczalnej PAN.

Lista prac naukowych opublikowanych przez Doktorantkę w trakcie pracy nad doktoratem obejmuje następujące pozycje:

- Charzyńska A., Nałęcz A., Rybiński M., Gambin A.; Sensitivity Analysis of Mathematical Models of Signaling Pathways; *BioTechnologia* 93(3), 291–308, 2012;
- Gambin A., Charzyńska A., Ellert-Miklaszewska A., Rybiński M.; Computational models of JAK1/2-STAT1 signaling; *JAK-STAT* 2:3, e24672, Landes Bioscience, 2013;
- Jetka T., Charzyńska A., Gambin A., Stumpf M.P.H., Komorowski M.; StochDecomp - Matlab package for noise decomposition in stochastic biochemical systems; *Bioinformatics* 30(1), 137–138, 2014;
- Wronowska w., Charzyńska A., Nienałtowski K., Gambin A.; Computational modeling of sphingolipid metabolism; *BMC Systems Biology* 9:47, 2015;
- Charzyńska A., Gambin A.; Improvement of the k-nn Entropy Estimator with Applications in Systems Biology; *Entropy* 18(1), 13, 2016;

Kandydatka uczestniczyła w trzech projektach naukowych (lista podana przez Doktorantkę w języku angielskim):

- „Information technologies: research and interdisciplinary applications.” POKL.04.01.01-00-051/10-00, realized by the Institute of Computer Science PAS (PhD student);
- “Integrative systems biology: inference from heterogeneous data” funded by the National Center for Science, grant No. 2011/01/B/NZ2/00864, realized by the Faculty of Mathematics, Informatics and Mechanics UW, headmaster Prof. Anna Gambin (research participant);
- „Applications of digital technologies in medicine” POIG.02.03.00-00-003/09 realized by the Institute of Biochemistry and Biophysics (research participant);

oraz otrzymała stypendia naukowe z trzech źródeł (również w języku angielskim):

- 2013-2014 Stypendium dla mazowieckich doktorantów projektu systemowego Samorządu Województwa Mazowieckiego z Europejskiego Funduszu Społecznego;
- 2012-2015 Stypendium doktoranckie, Instytut Podstaw Informatyki Polskiej Akademii Nauk;
- 2003-2007 Coroczne stypendium naukowe dla wyróżniających się studentów wydziału Matematyki, Informatyki i Mechaniki UW;

W swojej rozprawie doktorskiej Kandydatka skupiła się na podsumowaniu dostępnych oraz rozwijaniu nowych metod analizy wrażliwości, zaprezentowanych na znanych z literatury bądź stworzonych we współpracy z biologami modelach szlaków sygnałowych i metabolicznych.

Modelowanie matematyczne procesów biologicznych w formie układów dynamicznych stanowi uzupełnienie metod doświadczalnych służących do pomiaru i identyfikacji kluczowych elementów systemu, a przez to do lepszego zrozumienia modelowanego zjawiska. Formułowanie i walidacja modelu opisującego naturalne zjawisko składa się z kilku faz, które powinny być cyklicznie powtórzone w celu lepszego zbadania mechanizmów leżących u podstaw badanego procesu. Pierwszym krokiem jest postawienie

hipotez badawczych. Następnie tworzony jest model, który w formalny sposób opisuje badane zjawisko oraz wspiera projektowania eksperymentów w celu zweryfikowania hipotez badawczych. Formalna walidacja modelu przy pomocy narzędzi statystycznych ma zasadnicze znaczenie dla właściwego projektowania eksperymentów. Ważnym elementem modelowania jest analiza wrażliwości, tj. identyfikacja i analiza najbardziej istotnych elementów, które nawet przy niewielkich zaburzeniach mogą istotnie wpłynąć na wynik modelowania. Weryfikacja modelu następuje poprzez porównanie przewidywań symulacyjnych z wynikami eksperymentalnymi. Należy dokonać wszelkich koniecznych korekt w strukturze modelu oraz oszacowaniu nieznanymi parametrów na podstawie dostępnych danych eksperymentalnych, co ilustruje Figura 1.1 Doktorantki.

Pierwszy rozdział pracy stanowi przekrój i podsumowanie najistotniejszych aspektów pracy doktorskiej.

Kolejny rozdział poświęcony jest teoretycznym aspektom modelowania przy pomocy układów dynamicznych, zarówno w postaci deterministycznej, jak i stochastycznej. Główna część tego rozdziału poświęcona jest metodom analizy wrażliwości stosowanym dla danego typu modelowania. Zawarto tutaj rzetelny przegląd dostępnych współcześnie metod analizy wrażliwości wraz ich klasyfikacją.

Trzeci rozdział poświęcony jest metodzie analizy wrażliwości opartej o wzajemną informację dla zmiennych ciągłych. Pojęcie wzajemnej informacji jest miarą pochodzącą z teorii informacji, przy pomocy której można zbadać poziom zależności dwóch (potencjalnie wielowymiarowych) zmiennych losowych. Estymacja wzajemnej informacji jest możliwa na przykład dzięki wyrażeniu jej za pomocą relacji między entropią zmiennych brzegowych, a entropią zmiennej łącznej. W tym rozdziale Doktorantka prezentuje różne znane estymatory entropii dla zmiennych ciągłych w ujęciu historycznym. Opisuje na przykładach problemy jakie niesie ze sobą estymacja dla wielowymiarowych zmiennych ciągłych o ograniczonym nośniku. Niewątpliwie oryginalnym wkładem w rozwój metody analizy wrażliwości opartej o wzajemną informację jest udoskonalenie przez Kandydatkę estymatora entropii dla wielowymiarowych danych pochodzących z rozkładów ciągłych. W szczególności w rozdziale trzecim przytoczono konstrukcyjny dowód zbieżności estymatora entropii ciągłej, na podstawie którego

zidentyfikowano przyczynę, dla której estymator dla wielowymiarowych zmiennych ciągłych o ograniczonym nośniku jest obciążony. Ponadto podano wzór na korektę dla estymatora entropii (opartego o miarę prawdopodobieństwa na podstawie odległości do k-tego najbliższego sąsiada), dzięki czemu estymacja staje się dokładniejsza. Różne przykłady wielowymiarowych zmiennych losowych o ograniczonym nośniku, ilustrują fakt, że poprawiony estymator znacznie lepiej przybliży właściwą wartość entropii, na losowych próbach danych pochodzących z tych rozkładów.

Czwarty rozdział poświęcony jest analizie wrażliwości kilku wybranych modeli szlaków sygnałowych. Jako ilustrujący przykład klasycznych metod analizy wrażliwości, zaprezentowanych w drugim rozdziale, wykorzystano modele układu receptora błonowego indukowanego ligandem (Shankaran i in. 2007). Analizę wrażliwości przeprowadzono zarówno dla modelu w oryginalnej wersji deterministycznej, jak i przekształconej w wersję stochastyczną przy użyciu równania master (ang. Chemical Master Equation).

W celu zilustrowania nowej metody analizy wrażliwości opartej o entropię przetestowano ją na dobrze znanym przykładzie pętli sprzężenia zwrotnego białka p53 i ligazy Mdm2. Białko p53 antyonkogen, znane także jako białko transkrypcyjne TP53 jest czynnikiem transkrypcyjnym determinującym los komórki w przypadku uszkodzenia DNA. W ponad połowie przypadków raka ludzkiego p53 jest inaktywowane lub całkiem nieobecne, co umożliwia zmutowanej komórce nowotworowym replikację i determinuje ich nieśmiertelność (Geva-Zatorsky i in. 2006).

Ostatnia część tego rozdziału poświęcona jest zaproponowanemu przez Doktorantkę i współpracowników formalnemu modelowi procesów regulacyjnych ścieżkę metaboliczną sfingolipidów. Jest to pierwszy model obliczeniowy metabolizmu sfingolipidów uwzględniający podział na kompartmenty komórkowe (tj. jądro, mitochondria, lizosomy itp.), na podstawie typowej struktury niezróżnicowanej komórki eukariotycznej. Model oparty jest o równania różniczkowe zwyczajne (ang. ODE) opisujące ewolucję stężenia gatunków. Kinetyka modelu opiera się głównie na prawie działania mas (ang. MAL) dla reakcji transportu molekularnego oraz na uproszczonych przy pomocy podejścia Michael-Menten (MM) równaniach dla katalizowanych enzymatycznie reakcji. Modelowana kinetyka obejmuje również inhibicję w

obrębie konkurencyjnych gatunków. W celu sprawdzenia poprawności modelu zastosowano zarówno standardowe, jak i nowe metody analizy wrażliwości. W szczególności użyto metody opartej o analizę rozkładu wariancji (Komorowski i in., 2013). Zaproponowany model stanowi kompleksową, funkcjonalną integrację danych eksperymentalnych z modelowaniem i mamy nadzieję, że przyczyni się do lepszego zrozumienia, jeszcze nie do końca wyjaśnionej, zależności pomiędzy metabolizmem sfingolipidów a różnymi chorobami. Model został opublikowany w pracy w 2015 roku (Wronowska i in., 2015).

Ostatni rozdział stanowi podsumowanie i opisuje kierunki dalszych badań.

### **Ocena końcowa**

Rozprawa przedstawiona przez Doktorantkę, pomimo drobnych usterek językowych, napisana jest interesująco. Kandydatka pokazała, że dobrze rozumie i umiejętnie używa różne metody obliczeniowe, potrafi formułować hipotezy badawcze i je następnie weryfikować. Wyniki badań Doktorantki były poddane szczegółowej ocenie przez recenzentów 5 publikacji, którzy byli wybrani przez edytorów poszczególnych czasopism. Dorobek naukowy magister Agaty Charzyńskiej został już oceniony przez wielu ekspertów. Rozprawa doktorska świadczy o dobrym zrozumieniu przez Kandydatkę stawianych zadań badawczych. Cele rozprawy udało się zrealizować. Doktorantka wykazała, że zna i umiejętnie używa odpowiednio dobrane metody obliczeniowe, formułuje hipotezy badawcze i je weryfikuje, tak więc ogólną wiedzę teoretyczną Doktorantki w dziedzinie nauk matematycznych należy ocenić pozytywnie.

Oceniam, że rozprawa ta spełnia zwyczajowe i ustawowe wymogi, stawiane rozprawom doktorskim, stanowi oryginalne rozwiązanie problemu naukowego, unaocznia ogólną wiedzę teoretyczną kandydata w zakresie matematyki oraz pokazuje umiejętność samodzielnego prowadzenia pracy naukowej. Wnoszę zatem do Rady Naukowej Instytutu Podstaw Informatyki PAN o dopuszczenie Pani magister Agaty Charzyńskiej do dalszych etapów przewodu doktorskiego.



Dr hab. Dariusz Plęwczyński, prof. UW