

Nr 3 Październik 2018

# BioInfoWorld

## Filogenetyka



Kolekcyjne Biologiczne

W tym numerze:

Parsymonia

Bio.Phylo

Półróżna porcja modeli

# BIOINFO WORLD

Nr 3 Październik/Listopad

**Czasopismo Koła Naukowego Bioinformatyki BIT**

ISSN 2449 -5913

Wydanie online

**Opiekun naukowy**

dr Jacek Śmietański

**Redakcja**

Paulina Kania

Rafał Szkotak

**Okładka**

Kamil Malisz

**Korekta**

Marta Szkop

**Skład**

Rafał Szkotak

**Wydawca**

Koło Bioinformatyki BIT UJ

**Kontakt**

[bioinfoworld.redakcja@gmail.com](mailto:bioinfoworld.redakcja@gmail.com)

**Sfinansowane przez**



RADA KÓŁ NAUKOWYCH  
UNIwersytetu Jagiellońskiego

# Spis treści

Tytułem wstępu (Dr hab. Wiesław Babik) .....	3
Cichy Kącik Biojęzyków .....	5
Filogenetyka z trzech perspektyw (Rafał Szkotak) .....	5
Cichy Kącik Modelarza .....	11
Mechanizmy sprzężenia zwrotnego – posiłki (Paulina Kania) .....	11
Model oscylatora molekularnego białka zegara biologicznego - PER u D. Melanogaster (Karolina Cibor) .....	16
Artykuł Gościa .....	22
Tworzenie drzew filogenetycznych metodą redukcjonistyczną (parsimony) (Kamil Antończak) .....	22
Cyfrowa komórka nerwowa (Judyta Jabłońska) .....	26
Sprawozdanie z działań Koła .....	35
Słowniczek dla biologów .....	36
Słowniczek dla informatyków .....	36

# Tytułem wstępu

**Dr hab. Wiesław Babik**

Wydział Biologii i Nauk o Ziemi

Instytut Nauk o Środowisku

Znaczenia metod rekonstrukcji filogenezy we współczesnej biologii i medycynie nie sposób przecenić. Od dawna analizy filogenetyczne wykorzystuje się, aby poznać pokrewieństwa między gatunkami oraz innymi jednostkami taksonomicznymi. Jednak współcześnie służą one także rozmaitym innym celom. Na przykład metody rekonstrukcji filogenezy pozwalają określić w jakim stopniu spokrewnione są różne geny w obrębie genomu, przyczyniając się do zrozumienia jak powstają i ewoluują rodziny genów. Metody filogenetyczne mają wielkie znaczenie praktyczne w medycynie i ochronie zdrowia. Pozwalają one na określenie pochodzenia szczepów patogenów, śledzenie obecnych i historycznych dróg ich transmisji między regionami geograficznymi czy nosicielami oraz identyfikację źródeł pochodzenia epidemii. W przypadku szybko mutujących wirusów metody filogenetyczne umożliwiają śledzenie zmian zachodzących w czasie choroby u jednego pacjenta. Wyniki analiz filogenetycznych mogą nawet być wykorzystywane jako materiał dowodowy w sądach.

Dzięki błyskawicznemu postępowi technologicznemu ogromne ilości sekwencje DNA produkuje się obecnie tanio i

szybko. To z kolei powoduje coraz większe zapotrzebowanie na wydajne i szybkie metody rekonstrukcji drzew filogenetycznych oraz na integrację ich w potoki pozwalające na automatyczne wykonywanie złożonych, wielostopniowych analiz. O ile jeszcze niedawno drzewa filogenetyczne budowane z kilkuset sekwencji uchodziły za duże, teraz nierzadko konstruuje się drzewa na podstawie setek tysięcy sekwencji. Wymusza to projektowanie coraz bardziej wydajnych algorytmów heurystycznych o jak najlepszych własnościach obliczeniowych. Algorytmy te ma charakteryzować szybkość działania na dużych zbiorach danych, ale też wystarczająca precyzja. Zapotrzebowanie na takie algorytmy wśród biologów i medyków jest ogromne, co stawia wielkie wyzwania przed bioinformatykami, ale i daje szansę na ciekawą, stymulującą intelektualnie i użyteczną dla innych pracę.

# Cichy Kącik Biojęzyków

## Filogenetyka z trzech perspektyw

Rafał Szkotak

Po przerwie wakacyjnej wznawiamy wydawanie naszego czasopisma. Numer trzeci poświęcony zostanie filogenetyce. W dalszym ciągu będziemy eksplorować biblioteki BioJavy[1] oraz Biopythona[2], jednakże zagłębimy się również w Bio++[3], będącą podzbiorem bibliotek języka C++.

Zacznijmy więc od drzew - zagadnienia znanego zarówno biologom, jak i informatykom. Drzewem filogenetycznym nazywamy poddrzewo drzewa życia, ukazujące nam wspólnych przodków danych osobników (genealogię gatunków), znajdujących się w rozgałęzieniach lub pochodzenie pewnych wariantów genów. Ciekawe pytanie zadaje Futuyma w swojej książce poświęconej ewolucji[4] „Czy kiedykolwiek pomyślałeś o sobie, jako o krewnym rozgwiezdy, motyla lub grzyba?”. Ostatnim pojęciem związanym z drzewami jest **klad**, czyli grupa organizmów (genów), mająca wspólnego przodka. Przejdźmy zatem do stworzenia prostego drzewa filogenetycznego.

W celu stworzenia drzewa pobrano sekwencje białkowe receptora Toll-podobnego u czterech ssaków: szczura (gi|287324526|), myszy (gi|13507603|), byka (gi|114052617|) oraz człowieka (gi|49065550|). Uliniowienia sekwencji

dokonano w programie ClustalW (<http://www.genome.jp/tools/clustalw/>), przy użyciu ustawień domyślnych. Plik wyjściowy zapisano w formacie *phylip*. Stwórzmy więc pierwsze drzewo. W tym celu importujemy:

```
from Bio import AlignIO
```

```
from Bio import Phylo
```

Stwórzmy obiekt zawierający uliniowanie (składnia funkcji *read* jest ujednolicona we wszystkich modułach):

```
alignment = AlignIO.read('ścieżka', 'format')
```

Kolejnym krokiem jest obliczenie dystansów pomiędzy sekwencjami. Moduł *Phylo* daje kilka możliwości:

1. Metody dystansowe:
  - **metodę średnich połączeń (UPGMA)**
  - **przyłączania sąsiada (NJ)**
2. Metodę **parsymonii**.

Moduł *Phylo* pozwala również na przeprowadzenie procedury ***bootstrap***.

W pierwszym przypadku należy dostać się do kalkulatora, przechodząc przez kilka podmodułów. Gdy to się uda, jako argument podajemy nazwę metody obliczającej dystans (w naszym przypadku będzie to *blosum62*), przy jego pomocy stworzymy macierz odległości ewolucyjnych i w końcu drzewo filogenetyczne:

```

from Bio.Phylo.TreeConstruction import DistanceCalculator,
from Bio.Phylo.TreeConstruction import
DistanceTreeConstructor

calculator = DistanceCalculator(nazwa_metody)

distance_matrix = calculator.get_distance(alignment)

constructor = DistanceTreeConstructor(calculator, 'upgma')
# lub 'nj'

tree = constructor.upgma(distance_matrix)

```

Procedura tworzenia drzewa metodą parsymonii prezentuje się podobnie: linie zawierające *calculator*, *distance\_matrix* oraz *constructor* należy zastąpić:

```

scorer = ParsimonyScorer()

searcher = NNITreeSearcher(scorer)

constructor = ParsimonyTreeConstructor(searcher,
starting_tree)

```

Obiekt *tree* przechowuje nasze drzewo. Można je narysować przy pomocy zwykłej funkcji wypisania: *print tree* (rysunek 1), stworzyć pseudografikę: *Phylo.draw\_ascii(tree)* (rysunek 2) lub jeśli mamy zainstalowaną bibliotekę *matplotlib* lub *pylab* stworzyć grafikę: *Phylo.draw(tree)* (rysunek 3). Drzewa można ponadto w różny sposób ukorzeniać (funkcja *root\_with\_outgroup*), szukać wspólnych przodków (metoda *common\_ancestor*) oraz kolorować czy zmieniać typ.



BioJava posiada trzy moduły służące do analiz filogenetycznych. Są to:

1. *org.biojavax.bio.phylo*
2. *org.biojavax.bio.phylo.io.nexus*
3. *org.biojavax.bio.phylo.io.phylip*

Pierwszy z nich zawiera, podobnie jak w Biopythonie, moduły z funkcjami do obliczeń macierzy dystansów oraz minimalnych parsymonii, a także modele dwuparametrowy Kimury i Jukes – Cantora pozwalające na uniknięcie **problemu wielokrotnych substytucji** (klasa *MultipleHitCorrection*). Dwa kolejne, jak wskazuje nazwa zawierają metody wczytujące i parsujące pliki zawierające uliniowania sekwencji w formatach *nexus* i *phylip*.

Z racji tego, że Bio++ (3) jest omawiany tutaj po raz pierwszy napiszę o nim parę słów. Składa się on z kilku modułów, podzielonych na dwie grupy: trzy biblioteki typowo biologiczne (*SeqLib*, *PhyLib* i *PopGenLib*) oraz dwie zawierające różne przydatne metody (*Utils* oraz *NumCalc*). Twórcom Bio++ przyświecały trzy cele: przyspieszenie implementacji, ułatwienie dystrybucji rozbudowanych aplikacji oraz udostępnienie **języka wysokiego poziomu** biologom. Biblioteką, która nas najbardziej interesuje jest *PhyLib* (*libcov*, *bpp-seq*) [4]. Funkcjonalność Bio++ jest zbliżona do Biopythona oraz BioJava. Spis funkcji (wraz z artykułami naukowymi) odnaleźć można pod adresem znajdującym się w bibliografii na pozycji [4]. Z ciekawostek można zauważyć, że spis funkcji nie zawiera metody parsymonii,

w zamian za którą pojawia się **BioNJ**, czyli ulepszona wersja modelu przyłączania sąsiada.

Jak zawsze zachęcam do własnych poszukiwań, lektury dokumentacji oraz eksperymentów na kodzie.

```
Tree(rooted=True)
  Clade(branch_length=0, name='Inner3')
    Clade(branch_length=0.0280577793368, name='Inner2')
      Clade(branch_length=0.101111699312, name='Byk')
      Clade(branch_length=0.101111699312, name='Człowiek')
    Clade(branch_length=0.0724088134281, name='Inner1')
      Clade(branch_length=0.0567606652205, name='Szczur')
      Clade(branch_length=0.0567606652205, name='Mysz')
```

RYSUNEK 1



RYSUNEK 2

## Bibliografia:

[1] Dokumentacja API BioJavy:

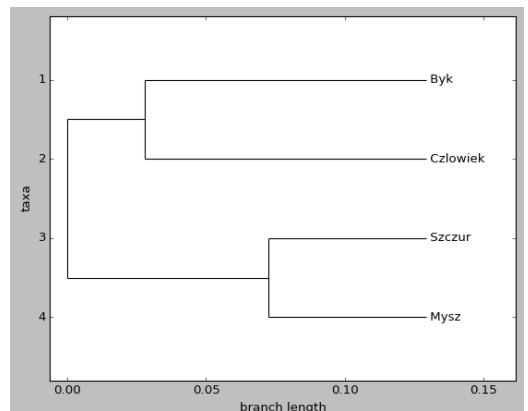
<http://www.biojava.org/docs/api/>

[dostęp 6.09.2015].

[2] Tutorial Biopythona:

<http://biopython.org/DIST/docs/tutorial/Tutorial.html>, rozdział 12, [dostęp 4.09.2015].

[3] Duthel i in., „Bio++: a set of C++ libraries for sequence analysis, phylogenetics, molecular evolution and population genetics.”, *BMC Bioinformatics*, 2006, 7-18.



RYSUNEK 3

[4] [http://biopp.univ-montp2.fr/wiki/index.php/List\\_of\\_methods](http://biopp.univ-montp2.fr/wiki/index.php/List_of_methods) [dostęp 4.09.2015].

# Cichy Kącik Modelarza

## Mechanizmy sprzężenia zwrotnego – posiłki

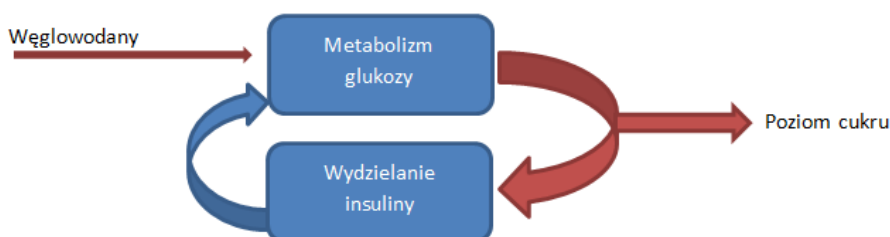
**Paulina Kania**

Enzymy są cząsteczkami, które obniżają energię aktywacji reakcji chemicznej, przyspieszając w ten sposób przebieg reakcji. Aktywność wielu enzymów może być hamowana przez różne typy inhibitorów.

W wielu reakcjach inhibitory biorą udział w mechanizmie regulacji aktywności enzymatycznej na drodze sprzężenia zwrotnego. Jeśli enzym produkuje jedną substancję ponad potrzeby komórki, to ta substancja może stać się inhibitorem dla tego enzymu, co zmniejsza lub całkowicie hamuje aktywność enzymu, a to z kolei zmniejsza stężenie produktu. Taka regulacja jest formą ujemnego sprzężenia zwrotnego.

Podobna sytuacja zachodzi w metabolizmie węglowodanów, podczas regulacji poziomu cukru we krwi.

Glukoza jest podstawowym związkiem energetycznym, przenikającym do poszczególnych komórek ciała. Gdy poziom cukru we krwi wzrasta (np. po posiłku) trzustka zwiększa wydzielanie insuliny do krwiobiegu. Przy braku insuliny glukoza nie jest w stanie przeniknąć z krwi do komórek.



*Schemat 1. Uproszczony schemat metabolizmu węglowodanów.*

Procesy zachodzące pomiędzy glukozą i insuliną można opisać następującymi układami równań (przedstawiona jest sytuacja przy braku zewnętrznego źródła glukozy i insuliny oraz odpowiedniej ilości materiałów zapasowych):

$$\begin{array}{ll} \left. \begin{array}{l} \frac{dg}{dt} = -vgi + \omega(\psi - g) \\ \frac{di}{dt} = -\alpha i \\ \frac{dg}{dt} = -vgi \end{array} \right\} & \begin{array}{l} \text{gdy modelowany} \\ \text{organizm jest na czczo } (g(t) < \psi) \end{array} \\ \left. \begin{array}{l} \frac{di}{dt} = -\alpha i + \beta(g - \psi) \end{array} \right\} & \begin{array}{l} \text{gdy modelowany organizm} \\ \text{jest po posiłku } (g(t) \geq \psi) \end{array} \end{array}$$

Oznaczenia:

$g$  – stężenie glukozy w plazmie krwi

$i$  – stężenie insuliny w plazmie krwi

$\psi$  - progowa wartość stężenia glukozy we krwi powyżej której uwalniana jest insulina

$\alpha, \beta, v, \omega$  – parametry modelu

Wykonano dwie symulacje: pierwsza dla człowieka zdrowego, druga dla człowieka chorego na cukrzycę typu I. W tym typie cukrzycy trzustka produkuje zbyt mało insuliny lub nie produkuje jej wcale. Jest to choroba wrodzona. Sytuacja na czczo jest przedstawiona gdy wartość glukozy przyrównuje się do 80 mg/dl.

Wartości dla człowieka zdrowego:

$$\alpha = 2.0 \text{ [1/h]}$$

$$\beta = 0.5 \text{ [1/h]}$$

$$v = 0.05 \text{ [cm}^3\text{/hmg]}$$

$$\omega = 1.0 \text{ [1/h]}$$

$$\psi = 80 \text{ [mg/dl]}$$

Wartości dla człowieka chorego:

$$\alpha = 2.0 \text{ [1/h]}$$

$$\beta = 0.05 \text{ [1/h]}$$

$$v = 0.05 \text{ [cm}^3\text{/hmg]}$$

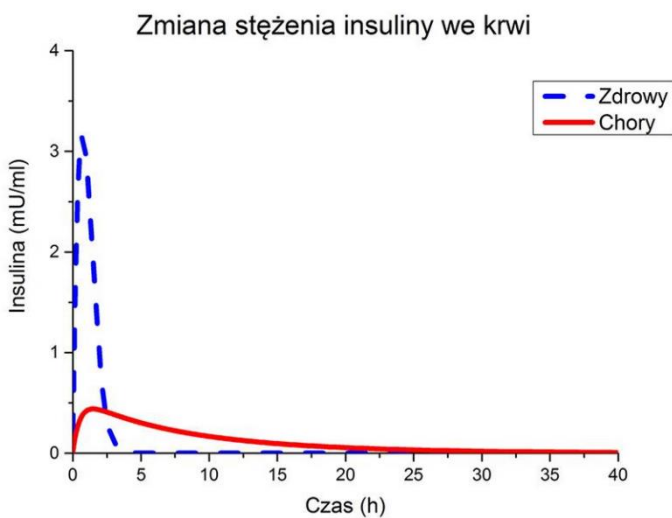
$$\omega = 1.0 \text{ [1/h]}$$

$$\psi = 80 \text{ [mg/dl]}$$

Obie symulacje wykonano w czasie 40 godzin, przy kroku czasowym równym 0.02 h. Wyniki przedstawiono na poniższych wykresach.



Wykres 1. Wykres przedstawiający zmianę stężenia glukozy we krwi dla człowieka zdrowego i chorego.



Wykres 2. Wykres przedstawiający zmianę stężenia insuliny we krwi dla człowieka zdrowego i chorego.

## Wnioski

Na powyższych wykresach można zauważyć porównanie poziomu cukru we krwi dla człowieka zdrowego i chorego, zarówno na czczo jak i po posiłku. Symulacja jest

przeprowadzona przy założeniu wchłaniania spożytego cukru w jelicie cienkim po posiłku w czasie równym 0. Organizm wtedy zaczyna wchłaniać cukry we krwi, dlatego można zauważyć gwałtowny spadek glukozy na wykresie nr 1. Następny wzrost wynika z tego, że glukoza zostaje uwalniana z wątroby.

Warto zauważyć jak zachowuje się insulina w obu przypadkach. Dla człowieka zdrowego w momencie, gdy cukier zaczyna spadać, produkcja insuliny wzrasta. W sytuacji człowieka chorego na cukrzycę typu I widoczny jest znaczny spadek w jej ilości oraz brak odpowiedzi na spadek glukozy.

Normalny poziom glukozy we krwi to 60 – 100 mg/dl, co jest zachowane w tej symulacji. Stężenie insuliny wynosi tutaj 0 mg/ml na czczo (jest to więc wartość nieprawidłowa fizjologicznie, ale zgodna z publikacją).

## Referencje

[1] Osowski S., „Modelowanie zmian zawartości cukru i insuliny we krwi.” *Modelowanie układów dynamicznych z zastosowaniem języka SIMULINK*, Oficyna Wydawnicza Politechniki Warszawskiej, Warszawa, 2007, 191–196.



# **Model oscylatora molekularnego białka zegara biologicznego - PER u *D. Melanogaster*.**

**Karolina Cibor**

## **Wstęp:**

Jedną z cech organizmów jest zdolność do pomiaru upływu czasu, dzięki endogennemu mechanizmowi określanemu mianem zegara biologicznego. Pozwala on realizować funkcje życiowe w najbardziej korzystnym okresie, odmierzając czas w oparciu o analizę różnic w natężeniu bodźców, głównie światła. Jego natężenie zmienia się w sposób przewidywalny i jest podstawowym dawcą czasu (niem. Zeitgeber) dla zegara, który dzięki odpowiednim receptorom odbiera informacje o fotoperiodzie, przetwarza je i generuje rytmy biologiczne.

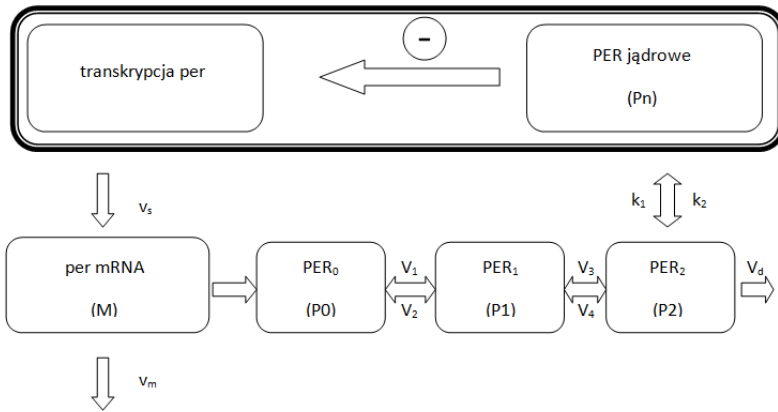
Zegar biologiczny generuje rytmy o cechach, odróżniających go od innych oscylacji w ustroju są to m.in.: endogenny charakter zegara przejawiający się zachowaniem generowanego rytmu w stałych warunkach (np. po umieszczeniu organizmu w stałej ciemności), okresowość rytmu, który wynosi w przybliżeniu 24 godziny, przez co określany jest jako cirkadianny czyli okołodobowy (łac. circa = zbliżony, dies = dzień)

W latach 70 odkryto, że za zmiany rytmów aktywności lokomotorycznej oraz linień poczwarek do form dorosłych u *D. Melanogaster* odpowiada mutacja pojedynczego genu, który nazwano period (ang. period = okres). Produkty jego ekspresji, czyli białka zegara biologicznego, oddziałują ze sobą tworząc mechanizm oscylatora molekularnego. Jego działanie objawia się w postaci rytmicznej ekspresji wspomnianego genu i jest realizowane w szlakach, gdzie białka zegara regulują ekspresję własnych genów, a także innych genów zegarowych.

W ciągu dnia (faza jasna cyklu) obserwujemy wzmożoną aktywność transkrypcyjną genu period (per). Gromadzenie się transkryptu

per w cytoplazmie skutkuje wzrostem ilości białka PER na początku nocy (fazy ciemnej cyklu). W połowie fazy ciemnej białko PER jest translokowane do jądra komórkowego, gdzie hamuje ekspresję własnych genów. Cykl ekspresji białek zegara trwa około 24 godziny ze względu na modyfikacje posttranslacyjne, które są czasochłonnym procesem. Obejmują one przede wszystkim fosforylację PER. Białko PER jest fosforylowane przez kinazę kazeinową I  $\epsilon$  zwaną Doubletime (DBT) oraz kinazę kazeinową 2 (CK2) lub defosforylowane przez fosfatazy PP1 i PP2A. Podwyższa to jego aktywność, jako czynnika hamującego transkrypcję.

## 1. Schemat modelu oscylacji PER



## 2. Kinetyczny opis systemu

$$\begin{aligned} \frac{dM}{dt} &= V_s \frac{K_I^n}{K_I^n + P_n^n} - v_m \frac{M}{K_m + M} \\ \frac{dP_0}{dt} &= k_s M - V_1 \frac{P_0}{K_I + P_0} + V_2 \frac{P_1}{K_2 + P_1} \\ \frac{dP_1}{dt} &= V_1 \frac{P_0}{K_I + P_0} - V_2 \frac{P_1}{K_2 + P_1} - V_3 \frac{P_1}{K_3 + P_1} + V_4 \frac{P_2}{K_4 + P_2} \end{aligned}$$

$$\frac{dP_2}{dt} = V_3 \frac{P_1}{K_3 + P_1} - V_4 \frac{P_2}{K_4 + P_2} - k_1 P_2 + k_2 P_n - V_1 \frac{P_0}{K_I + P_0} - v_d \frac{P_2}{K_d + P_2}$$

$$\frac{dP_n}{dt} = k_1 P_2 - k_2 P_n$$

$$P_t = P_0 + P_1 + P_2 + P_n$$

### Wartości Parametrów:

$$K_1 = K_2 = K_3 = K_4 = 2 \mu\text{M}$$

$$K_I = 1 \mu\text{M}$$

$$K_d = 0,2 \mu\text{M}$$

$$k_s = 0,38 \text{ h}^{-1}$$

$$k_1 = 1,9 \text{ h}^{-1}$$

$$k_2 = 1,3 \text{ h}^{-1}$$

$$n = 4$$

$$v_s = 0,76 \mu\text{M h}^{-1}$$

$$v_m = 0,65 \mu\text{M h}^{-1}$$

$$V_1 = 3,2 \mu\text{M h}^{-1}$$

$$V_2 = 1,58 \mu\text{M h}^{-1}$$

$$V_3 = 5 \mu\text{M h}^{-1}$$

$$V_4 = 2,5 \mu\text{M h}^{-1}$$

na zewnątrz cyklu granicznego:

$$P_0 = P_1 = P_2 = P_n = 0,8 \mu\text{M}$$

$$M = 1,9 \mu\text{M}$$

wewnątrz cyklu granicznego:

$$P_0 = P_1 = P_2 = P_n = 0,25 \mu\text{M}$$

$$M = 0,1 \mu\text{M}$$

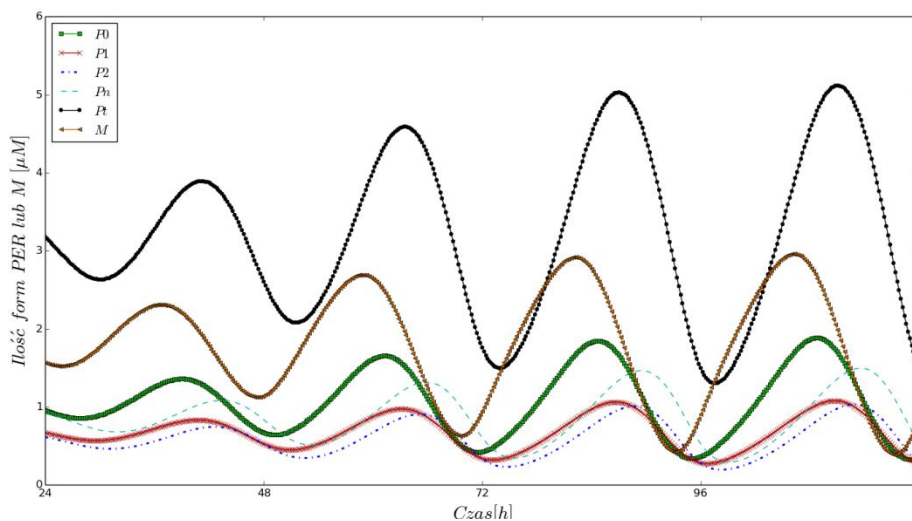
### 3. Opis symulacji

Model opisuje mechanizm oscylacji białka period z uwzględnieniem jego kilku fosforylacji oraz pętli ujemnego sprzężenia zwrotnego genu period (*per*).

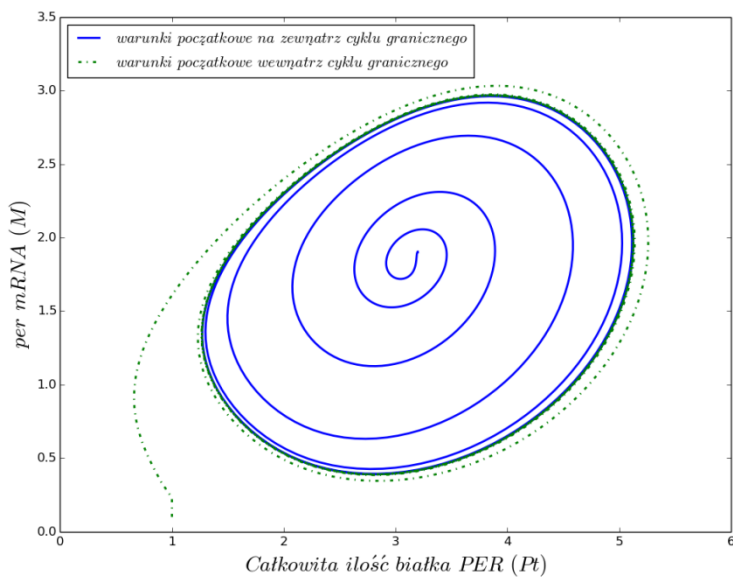
Założono, że *per* mRNA, którego stężenie w cytozolu oznaczono jako M, jest syntezowane w jądrze, a następnie transportowane do cytozolu, gdzie ulega degradacji. Tempo produkcji PER jest proporcjonalne do wielkości M. Przyjęto, że białko PER posiada 3 stany: nieufosforylowany  $P_0$ , mono- ( $P_1$ ) oraz difosforylowany ( $P_2$ ). Poszerzenie modelu o kolejne stany nie poprawiło jego dynamiki, zrezygnowano więc ze zwiększenia ich liczby.

Założono również, że jedynie forma jądrowa  $P_n$  może ulegać degradacji,  
a także pełni ona rolę represora w procesie ujemnego sprzężenia zwrotnego.

#### 4. Wykresy z wybranymi symulacjami.



WYKRES 1 WYKRES PRZEDSTAWIAJĄCY ZMIANĘ STĘŻENIA PER I M.



WYKRES 2 WYKRES ILUSTRUJĄCY EWOLUCJĘ UKŁADU W KIERUNKU GRANICY CYKLU W PRZESTRZENI  $(M, P_T)$ .

## 5. Wnioski.

Dla obranych wartości parametrów modelu wystąpiły okresowe zmiany stężeń PER i M z okresem około jednej doby.

Przyjmując dwa różne zestawy wartości początkowych system osiąga kształt krzywej zamkniętej, o charakterystycznej stałej amplitudzie i okresie.

Model poprawnie odwzorowuje działanie oscylatora molekularnego zaobserwowane w badaniach empirycznych.

### Referencje:

[1] Goldbeter A., „A Model for Circadian Oscillations in the Drosophila Period Protein (PER)”, *Proceedings of the Royal Society. B, Biological sciences*.

[2] Liu X1 i in., „Spatial and temporal expression of the period gene in Drosophila melanogaster”, *Genes Dev*.

# Artykuł Gościa

## Tworzenie drzew filogenetycznych metodą redukcjonistyczną (parsimony)

**Kamil Antończak**

Drogi czytelniku, gdybym spytał Ciebie, co jest bliżej Krakowa: Warszawa czy Zakopane, jaka byłaby Twoja odpowiedź? Wydawałoby się, że Zakopane, wszak to około 100 km, podczas gdy do Warszawy mamy 280 km. Jednak gdzie można się szybciej dostać? Samochodem do stolicy dojedziemy w 4 godziny, a do Zakopanego w 2.

A co z pociągami? Zapewne wszyscy słyszeli o słynnych Pendolino, dzięki którym dojedziemy do Warszawy o 50 min szybciej niż do drugiego miasta. No i na koniec pozostaje samolot, który leci tylko do jednego z tych miejsc i to w 50 minut. Tak więc czasem nie wszystko wydaje się takie proste, a odpowiedzi na niektóre pytania nie będą jednoznaczne. Tak samo jest w filogenetyce. Mnogość istniejących metod pozwalających stworzenie drzewa pokrewieństw powoduje, że nie tylko nie jesteśmy w stanie jednoznacznie ustalić jak dwa gatunki są od siebie odległe, to jeszcze możemy otrzymać różne drzewa pokrewieństw. W niniejszym artykule opiszę jedną z popularniejszych metod tworzenia drzew filogenetycznych.

Jednym z głównych problemów tworzenia w pełni poprawnych drzew jest szybko rosnąca złożoność tego zadania. Dla drzew ukorzenionych (posiadających grupę zewnętrzną) ilość wszystkich drzew, które można stworzyć jest dana wzorem:  $\frac{(2n-3)!}{(n-2)! \cdot 2^{n-2}}$ , co dla 10 taksonów daje nam 34459425 możliwości, a dla 135 otrzymujemy  $2,113 \cdot 10^{267}$ , co przekracza liczbę cząsteczek w znanym Wszechświecie, a także – jak łatwo się domyślić – przerasta możliwości większości komputerów. W artykule Michała Foryty, znajdującego się w drugim numerze BioInfo World można było przeczytać o sposobie kodowania informacji. W przypadku algorytmów filogenetycznych jest to jeden z większych problemów –

stworzenie systemu punktacji zgodnego z ewolucją, która jest mniej lub bardziej losowa.

Jedną z najprostszych technik jest metoda redukcjonistyczna. Stosuje ona brzytwę Ockhama, która nakazuje przyjąć najprostsze rozwiązanie. W tym przypadku drzewo układamy tak, aby liczba zmian była jak najmniejsza. Weźmy na przykład przypadek:

<i>Cecha</i>	Osobnik A	Osobnik B	Osobnik C
1	1	1	1
2	1	1	1
3	1	0	0
4	0	1	1
5	0	1	0
6	0	1	1
7	1	1	0

Niech Osobnik D będzie grupą zewnętrzną. Weźmy wszystkie możliwe drzewa, zapisane metodą Newick: (D, (A, (B, C))), (D, (B, (A, C))) oraz (D, (C, (A, B))). W pierwszym przypadku dokonujemy tylko 8 zmian, podczas, gdy w pozostałych dwóch odpowiednio 10 oraz 9 zmian. Nasze pierwsze drzewo więc niejako jest najlepsze zgodnie z tą techniką. W ten sposób, nasze optymalne rozwiązanie mówi nam, że cecha 7 uległa homoplazji (czyli nie jest homologiczna). Niestety, niezauważenie homoplazji może często prowadzić do błędnych wyników, które trudno będzie skorygować tą metodą. To dość poważne ograniczenie.

Oczywiście dochodzimy do problemu, czy wszystko w ewolucji jest możliwe? Z jednej strony trudno przewidzieć co się stanie, z drugiej jednak wiemy co jest mniej lub bardziej prawdopodobne. Dlatego istnieje kilka metod optymalizacji, które istotnie wpływają na



wynik. Pierwszą jest optymalizacja Camina-Sokala, która zakłada, że transformacja jest nieodwracalna. To oznacza, że nie możemy przejść do stanu, który już był. Jej nieco łagodniejszą i bardziej życiową wersją jest optymalizacja Dollo, która mówi, że nie można powrócić tylko do stanu zaawansowanego, jednak ta definicja jest dość elastyczna (ponieważ trudno powiedzieć co już jest stanem zaawansowanym, a co jeszcze nim nie jest). Następnie, mamy jeszcze optymalizację Wagnera, która zakłada, że transformacje są odwracalne, ale w zadanej kolejności. To oznacza, że cecha A może przejść w B a potem w C, lecz nie może cofnąć się do A nie przechodząc znów przez stan B. Ostatnia to optymalizacja Fitcha, która pozwala na odwracalną i nieuporządkowaną transformację.

Kolejną zaletą tej metody jest możliwość przypisania konkretnym mutacjom różnych wag i w ten oto sposób można zaznaczyć, że tranzycja występuje dużo częściej niż transwersje. Tak samo, jak można wyłączyć punktacje dla mutacji w tzw. hotspotach, które zbędnie zaburzają obraz filogenetyczny uzyskany w metodach odległościowych biorących pod uwagę tylko pełen genom.

Przed podsumowaniem warto jednak powiedzieć jeszcze kilka słów o wadach tej metody - niestety, nie daje nam ona informacji na temat czasu dywergencji. Na kolejny problem natkniemy się, jeśli otrzymamy kilka najlepszych wyników. Wtedy musimy zastosować inne metody, które pomogą nam wybrać najlepsze drzewo. Warto też zaznaczyć fakt, że jest to problem klasy NP i dla większych zestawów danych stosuje się metody heurystyczne (warto przypomnieć sobie (drugi akapit) ile mamy drzew do przejrzania). Niemniej, mimo tych minusów metoda ta jest bardzo popularna wśród biologów ze względu na wykorzystanie bardzo prostego, a jednak elastycznego algorytmu prowadzącego do otrzymania wyników zadowolających. Prawie nigdy nie otrzymamy drzewa poprawnego w 100%, jednak nie zawsze to jest najważniejsze.

## **Bibliografia**

[1] <http://konkursy.byd.pl/userfiles/files/Rogalla.pdf>

[2]

[https://en.wikipedia.org/wiki/Maximum\\_parsimony\\_\(phylogenetics\)](https://en.wikipedia.org/wiki/Maximum_parsimony_(phylogenetics))

[3] Falniowski A., Wykłady *Metody numeryczne w taksonomii*.

# Cyfrowa komórka nerwowa

## Judyta Jabłońska

Magistrantka w zakładzie Neuroanatomii UJ,

Członek Koła Neuronaukowców UJ "Neuronus"

Temat prezentowany podczas konferencji studenckiej "Liczby-Komputery-Życie" 15-17.05.2015 w Krakowie

Mózg ludzki budzi wielkie zainteresowanie nie tylko wśród naukowców. Jeden z najbardziej złożonych narządów naszego ciała pomimo ponad stu lat badań nadal skrywa wiele tajemnic. Jak to się dzieje, że myślimy, śnimy, jesteśmy głodni albo decydujemy się na kolejną i kolejną drzemkę w poniedziałkowy poranek? Żeby odpowiedzieć na te pytania warto zajrzeć bardzo głęboko, aż na poziom komórkowy.

W laboratoriach neurobiologicznych rzecz wygląda całkiem interesująco. Badania prowadzone są najczęściej na materiale biologicznym pochodzącym od zwierząt (w większości przypadków gryzoni). Choć nie tylko. Współpraca naukowców z inżynierami i twórcami oprogramowania zaowocowała eksperymentami z wykorzystaniem wirtualnych modeli komórek. Właśnie takie rekonstrukcje zostały wprowadzone w życie niemal jak monstrum Frankensteina. Brzmi nieprawdopodobnie? Przykładu nie trzeba szukać daleko. W Lozannie (Szwajcaria) w 2005 roku Henry Markram wraz z zespołem rozpoczął ambitny *Blue Brain Project*. Rozwój tego pomysłu pozwolił na odtworzenie kolumny korowej rozumianej jako podstawowa jednostka budulcowa mózgu, a także na symulację jej aktywności. Taka kolumna składa się nawet z 30.000 cyfrowych rekonstrukcji odzwierciedlających kształt komórek nerwowych, czyli neuronów. Odwzorowania tego typu tworzone są z kolei na bazie wcześniej uzyskanych danych doświadczalnych. Cele projektu na następne lata są jeszcze ambitniejsze i zawierają symulacje nawet 200 milionów komórek jednocześnie. A w bardziej odległej przyszłości nawet całego mózgu.

Pamiętam, że gdy podczas pierwszych zajęć na studiach zobaczyłam zdjęcia tkanki nerwowej to nie wywarły na mnie szczególnego wrażenia. Dlatego, że neuron wygląda trochę jak splątane ze sobą włókienka albo gałęzie drzewa gdyby spojrzeć na nie od dołu pnia. Dopiero po dokładniejszym poznaniu roli i zróżnicowania tych niezwykle pracowitych komórek naprawdę je doceniłam. Dowiedziałam się, że są bardzo zróżnicowane pod względem kształtu. Okazuje się, że w całym morzu (a raczej galaretkie) naszego mózgu znajdują się komórki o prostej morfologii takie jak neurony Von Economo (które kiedyś były podejrzewane o wytwarzanie świadomości) aż do skrajnie rozczapierzonych komórek Purkiniego w mózdku. Ba! Także ich aktywność elektryczna może charakteryzować się odmiennością. Rzeczą, której nie wolno pominąć jest fakt, że aktywność elektryczna komórki jest zależna od jej morfologii. Krótko mówiąc - jest co odkrywać i pracy nie zabraknie. Ale do rzeczy! W jaki sposób analizuje się morfologię komórki? Jak zapisuje się i udostępnia takie informacje? I do czego to jest w ogóle przydatne?

## **Pierwszy krok**

Istnieje kilka sposobów badania kształtu komórek nerwowych. Pierwszym krokiem zawsze jest uzyskanie materiału do badań. Mózg kroi się na cienkie skrawki o grubości mniejszej niż milimetr. Komórki nerwowe z natury są bezbarwne dlatego, aby je zobaczyć w takim preparacie należy tkankę w odpowiedni sposób wybarwić. Dzięki barwnikowi część widma światła przechodzącego przez preparat zostanie pochłonięta. Można także zastosować znacznik fluorescencyjny, co spowoduje emisję promieniowania o określonej długości fali. W praktyce naświetla się preparat światłem o jednym kolorze, a obserwuje obraz w innym, zależnie od podanego czynnika. Dzięki temu otrzymujemy obraz, na którym można dostrzec interesujące nas struktury, czyli wypustki komórki nazywane akson(y) i dendryty. O ile tylko nie zostawiliśmy na za długo preparatu w roztworze inkubacyjnym, żeby zrobić sobie kolejną kawę... i przy okazji oddać się lekturze jakiegoś artykułu.

## **"Ciemność widzę, widzę ciemność"**

Dla zbadania kształtu komórki niezbędne jest wysoce kontrastowe zdjęcie z dokładnie wybarwionymi wszystkimi wypustkami neuronu. Taki obraz można uzyskać różnymi metodami. Pierwszym sprawdzonym przez dziesięciolecia sposobem jest impregnowanie tkanki solami metali ciężkich (np. srebra) w barwieniu metodą Golgiego. Minusem tej metody jest to, że no cóż... Trwa niemal w nieskończoność, a dokładniej co najmniej kilka tygodni. Za plus można by uznać fakt, że wybarwieniu ulega tylko niewielki odsetek (<5%) komórek. Stanowi to istotną zaletę, ponieważ w galaretkę naszego mózgu komórki znajdują się bardzo blisko siebie i posiadają liczne wypustki. Tak więc przy wybarwieniu wszystkich neuronów zobaczylibyśmy jednolitą ciemną masę.

Drugą, szybszą drogą jest nastrzykiwanie komórki znacznikiem poprzez ukłucie nieutrwalonego, żywego skrawka elektrodą i podanie z jej pomocą barwnika np. biocytyny lub znacznika fluorescencyjnego. Może mieć to miejsce przy rejestracji elektrofizjologicznej aktywności komórki. Umożliwia to połączenie badań morfologicznych z fizjologicznymi. Dane z tego typu doświadczeń są bardzo cenne, ponieważ pozwalają na wysuwanie hipotez dotyczących zależności pomiędzy morfologią, a aktywnością komórki!

Inną drogę stanowi oznaczanie metodami genetycznymi. Do genomu komórki wprowadza się gen dla fluoryzującego białka. Można to zrobić poprzez bezpośrednie podanie plazmidu lub nasłanie na komórkę wirusa. W praktyce nie jest to trudne, ponieważ w paczce przychodzi do nas gotowy zamówiony wirus, który wystarczy tylko wprowadzić do preparatu. W efekcie otrzymamy zwierzęta, u których niektóre komórki będą świecić pod mikroskopem fluorescencyjnym.

## Fotografowanie

Obraz kształtu komórki otrzymuje się najczęściej dzięki mikroskopii świetlnej lub fluorescencyjnej. Każda z metod posiada swoje wady i zalety.

Z mikroskopią świetlną, czyli bazującą na świetle przechodzącym przez preparat na pewno nieraz już się spotkałeś. Komórka dzięki zawartemu w niej barwnikowi pochłania przechodzące przez nią światło. W efekcie, po barwieniu metodą Golgiego lub nastrzyknięciu biocytyną zobaczymy ciemną sylwetkę komórki na jasnym tle.

Nieco inaczej sprawa wygląda w mikroskopie konfokalnym (odmiana fluorescencyjnego), gdzie neuron oznaczony barwnikiem fluorescencyjnym będzie odróżniał się od ciemnego tła. Z tego typu obrazami kojarzą mi się stare amerykańskie neonowe napisy żarzące się ciemną nocą. W przypadku komórek mechanizm jest oczywiście inny. Polega na absorpcji promieniowania o odpowiedniej długości fali przez barwnik, a następnie emisji światła o innym kolorze. Fotografując preparat w mikroskopie świetlnym otrzymamy tylko jeden obraz, co jest kłopotliwe, ponieważ przy grubych preparatach fragmenty i poniżej płaszczyzny ogniskowania będą nieostre. Mikroskop konfokalny natomiast może "podzielić" preparat na skrawki optyczne, fotografując go warstwa po warstwie. W efekcie zamiast jednego nieostrego miejscami obrazu uzyskamy serię (stos) ostrych zdjęć. To wielka zaleta, bo na podstawie takiego stosu łatwo jest stworzyć trójwymiarowy obraz komórki, którym można swobodnie manipulować w trzech wymiarach.

Warto wiedzieć o jeszcze jednej kwestii. Otóż często bywa tak, że otrzymujemy zdjęcia nie całej komórki tylko jej fragmentów. Dzieje się tak dlatego, że komórki nerwowe mają ogromne rozmiary (jeśli idzie

o zasięg wypustek), więc nie sposób obejrzeć czy sfotografować całą komórkę z rozdzielczością pozwalającą na uwidocznienie szczegółów

na jednym obrazie. Pojedyncze zdjęcia trzeba więc jeszcze ze sobą poskładać jak puzzle i to w trzech wymiarach.

### **Kurs rysunku i Matrix**

W każdym przypadku po godzinach lub tygodniach walki z preparatyką, barwieniem i fotografowaniem otrzymuje się trofeum w postaci pięknych zdjęć. Można by teraz włączyć program do pomierzenia podstawowych parametrów takich jak objętość wypustek czy liczba komórek. Natomiast my ambitnie chcemy utworzyć cyfrowy, trójwymiarowy model komórki rodem z filmu science-fiction. Faktycznie osiągnięcie tego może okazać się niełatwe, ale jest warte tego trudu. Taki model pozwala na dokładną analizę morfologiczną, w trzech wymiarach, a także modelowanie funkcjonowania komórek.

### **Programy do trejsingu**

W praktyce wygląda to w ten sposób, że badacz siada przez komputerem, włącza odpowiednio zmodyfikowany program graficzny

i przez kolejne kilka godzin (albo i dni) dokładnie odrysowuje (trejsuje), klikając, kształt wybarwionej wypustki przesuwając stos obrazów, żeby uzyskać trójwymiarową rekonstrukcję. Na pomoc przy tej mrówczej robocie nadciąga grupa pomocników w postaci stabilnych, samodzielnych aplikacji dostępnych komercyjnie lub bezpłatnie. Zdecydowanie najczęściej stosowanym programem jest dostępna komercyjnie i droga *Neurolucida* (MBF Bioscience) wykorzystana do stworzenia 40,6% modeli dostępnych aktualnie w bazie danych *NeuroMorpho.org*. Faktycznie, mówi się o niej, że jest narzędziem kompleksowym i wygodnym. Niestety nigdy nie widziałam jej na oczy, więc nie potwierdzę. Zaraz za nią nadciągają *Amira* (Visualization Sciences Group), *3DMA-Neuron* (Nihon Visual Science) oraz *Imaris* (Bitplane). Wszystkie z tych wymienionych programów są płatne.

W szranki z tymi programami mogą stanąć darmowe *NeuronStudio* czy *Farsight*. Chcąc rozpocząć badania z wykorzystaniem modeli

można pokusić się o inne rozwiązanie i efektywnie zastosować pluginy do *ImageJ* (popularnego programu do analizy obrazu), z których najpopularniejsze to *Simple Neurite Tracer* czy *NeuronJ*, a nawet zainstalować specjalnie zmodyfikowaną dystrybucję *ImageJ* o nazwie *Fiji*. Dodam od siebie, że nawet w przypadku finansowanych badań, bez obawy o jakość analizy można zdecydować się na narzędzie bezpłatne.

### **'Po co się tak męczyć. Przecież można by...'**

Taką opinię usłyszałam od bioinformatyków, gdy prezentowałam ten temat podczas konferencji. I w pełni się z nią zgadzam! Ten proces zdecydowanie wymaga automatyzacji, nie tylko po to żeby ułatwić życie badaczowi ale także, aby modele były jak najbardziej wystandaryzowane. Jest to istotne, ponieważ w przypadku ręcznego trejsingu wpływ na detale kształtu komórki może mieć nawet to, czy badacz wypił rano słabszą czy mocniejszą kawę. No może nieco przesadzam natomiast myślę, że rozumiesz już tę ideę.

Automatyzacja procesu trejsingu jest możliwa. Opracowywane algorytmy bazują na dwóch sposobach. Może to być iteracyjne pomniejszanie krawędzi fragmentu wypustki, aż do utworzenia możliwie jak najcieńszej linii. Inną metodę stanowi poszukiwanie punktów końcowych lub początkowych danej wypustki, a następnie dojście jak po sznurku do ciała komórkowego (lub na odwrót - od ciała komórki do końców wypustek).

Algorytm automatyzujący zaimplementowany został m.in. w rozszerzeniu *NeuroLucidy*, a kilka najlepszych rozwiązań wyłoniono na drodze ogólnościatowego konkursu *DIADEM Challenge*, który trwał rok i zakończył się we wrześniu 2010 roku. Rezultaty pracy pięciu zespołów zostały nagrodzone i zaimplementowane w dostępne na 'rynku neuroinformatycznym' narzędzia. Pomimo tych starań, rozwiązania automatyzujące są rzadko stosowane. W znacznej



większości przypadków faktycznie wykorzystywany w badaniach pozostaje nadal algorytm pół-automatyczny. Dlaczego?

### **Trudności w automatyzacji**

Doświadczenia w laboratorium często składają się z wielu etapów, w którym każdy jest istotny, ponieważ jego rezultat będzie wykorzystany w kolejnym kroku. Tak dzieje się i w tym przypadku. Na jakość obrazu wpływa sposób skrojenia, wybarwienia preparatu, zdolność rozdzielcza mikroskopu, a także rodzaj mikroskopii, wariant znaczników lub barwników. Nie bez znaczenia okazuje się także typ komórki.

Obrazy nie są doskonałe, wypustki nakładają się na siebie, zlewają się ze sobą (zwłaszcza z powodu pewnego rozmycia w wymiarze Z). Cyfrowe zdjęcia charakteryzują się skończoną rozdzielczością. Zbudowane są z niepodzielnych elementów - pikseli w obrazie 2D lub wokseli w 3D. Natomiast niektóre szczegóły komórki mogą mieć rozmiary mniejsze niż woksel naszego trójwymiarowego zdjęcia. Stąd niekiedy, nawet przy automatycznym trejsingu, ręczny retusz jest nieuniknionym końcowym etapem obróbki obrazu.

### **Budowa i zapis modelu**

Nie powiedzieliśmy jeszcze o tym jak wygląda model tak 'od kuchni'. Pomyślany jest w niezwykle sprytny sposób. Rekonstrukcję komórki nerwowej można wyobrazić sobie jako ciąg prostych rurek sklejonych ze sobą w miejscach nazywanych węzłami (nodes). Każdy węzeł zostaje opatrzony numerem, a następnie zapisuje się dla niego współrzędne X,Y i Z w kartezjańskim układzie współrzędnych, wartość promienia oraz jako ostatnie -indeks poprzedzającego węzła (czyli tego z poprzednich, z którym jest połączony jedną rurką). Proste, prawda? Właśnie te informacje dotyczące węzłów zapisywane są w tabeli, gdzie każdy wiersz odpowiada pojedynczemu węzłowi.

Taki zapis pozwala na zrekonstruowanie kształtu przestrzennego komórki bez potrzeby posiadania dodatkowych danych wejściowych. Dla mnie istotny jest jednak fakt, że uzyskana rekonstrukcja jest minimalnych rozmiarów, tak że można ją przechowywać w bazie danych i dzielić się raz uzyskanym modelem z całym gronem naukowców. Stanowi to miód na serce pragnące współpracy i zintegrowanego rozwoju różnych dziedzin neurobiologii i neuroinformatyki.

### **No ale jak to się aktywuje?!**

Na początku zostało wspomniane, że w *Blue Brain Project* tworzone są cyfrowe fragmenty kory mózgowej. Takie kolumny składające się z dziesiątek tysięcy komórek zdradzają sposób funkcjonowania ludzkiego mózgu. Przechodzenie impulsu elektrycznego przedstawione jest wizualnie jako rozświetlenie. Nie znaczy to jednak, że podczas myślenia świecą nam się fragmenty komórek! Aktywność komórek polega na zmianach potencjałów elektrycznych ich błon i na wydzielaniu różnych substancji chemicznych. Taką aktywność komórek można symulować na przykład przy pomocy programu o nietrudnej do zapamiętania nazwie - NEURON, rozwijanego przez panów Michaela Hinesa i Teda Canevale (Yale University).

## **Do czego takie badania mogą być przydatne?**

Wykorzystać je można w badaniu zmian morfologii pojedynczej komórki albo całego ich zbioru podczas rozwoju, w trakcie uczenia się, stresu, starzenia, chorób neurodegeneracyjnych, do których należą m.in. choroba Alzheimerera, Parkinsona czy stwardnienie zanikowe boczne. W taki sposób sprawdza się wpływ różnorodnych czynników chemicznych np. czynników wzrostowych, czy substancji podejrzewanych o działanie lecznicze, a także mechanicznych jak punktowy uraz. Dziś symuluje się aktywność całej kolumny korowej. Co będzie więc jutro? Symulacja pracy całego mózgu, rozszyfrowanie świadomości, odwrócenie biegu procesu starzenia się? Może to jutro będzie zależeć właśnie od Ciebie. Neurobiologia jest pasjonującą nauką, tajemniczą, dotykającą naszego codziennego funkcjonowania.

## **Warto poznać:**

[1] [bluebrain.epfl.ch](http://bluebrain.epfl.ch)

[2] [www.diademchallenge.org](http://www.diademchallenge.org)

[3] Donohue D.E. i Ascoli G.A., „Automated reconstruction of neuronal morphology: an overview”, *Brain research reviews* 2011, 67: 94–102.

[4] Meijering E., „Neuron tracing in perspective”, *Cytometry Part A the journal of the International Society for Analytical Cytology* 2010, 77: 693–704.

[5] Parekh R. i Ascoli G.A., „Neuronal Morphology Goes Digital: A Research Hub for Cellular and System Neuroscience”, *Neuron* 2013, 77: 1017–1038.

# Sprawozdanie z działań Koła

Kolejny rok akademicki przed nami. W Kole został wybrany nowy zarząd, pełen pomysłów na nadchodzące miesiące.

Wartym uwagi wydarzeniem było seminarium inauguracyjne, które staje się już tradycją dla Koła Naukowego Bioinformatyki BIT. Tym razem poprosiliśmy naszego opiekuna, dr. Jacka Śmietańskiego z Instytutu Informatyki i Matematyki Komputerowej o wystąpienie.

Seminarium pt. "Jajko czy kura? Czyli gdzie dwóch się bije, tam trzeci korzysta" dotyczyło RNA, a konkretniej metod komputerowego przewidywania oddziaływań oraz pełnej struktury 3D cząsteczek RNA. Odkrycia ostatnich lat niezbicie dowodzą, że RNA pełni nie tylko rolę przekaźnika informacji pomiędzy nicią DNA a białkiem. Poszczególne klasy tak zwanych "niekodujących RNA" pełnią rozmaite funkcje metaboliczne, czasem nawet bez udziału białek. Funkcjonalność takich RNA, podobnie jak w przypadku białek, determinowana jest przez ich strukturę przestrzenną. Podczas prezentacji dr Śmietański opowiedział o metodach komputerowego przewidywania oddziaływań oraz pełnej struktury 3D cząsteczek RNA.

Co do naszych planów na nadchodzące miesiące można wymienić kilka głównych składowych:

- Na początku grudnia pojawi się kolejny numer naszego czasopisma BioInfo World, który tym razem będzie dotyczył sztucznej inteligencji. W związku z tym nawiązaliśmy współpracę z Kołem Robotyki i Sztucznej Inteligencji. Zachęcamy wszystkich zainteresowanych publikowaniem o zgłaszanie swoich artykułów na adres [bioinfoworld.redakcja@gmail.com](mailto:bioinfoworld.redakcja@gmail.com)
- Pewnym i stałym elementem naszej działalności jest również cykl seminariów. Z pewnością zorganizujemy już niedługo również pierwsze warsztaty! Więcej informacji pojawi się wkrótce na naszym fanpage'u: [www.facebook.com/koloBITUJ](http://www.facebook.com/koloBITUJ)
- Będziemy również przerabiać naszą stronę internetową; jeśli ktoś chciałby podszkolić swoje umiejętności lub je nabyć to zapraszamy!

## Słowniczek dla biologów

**Brzytwa Ockhama** – dążenie do prostoty, dowodzenie przy jak najmniejszej liczbie założeń i pojęć.

**Język wysokiego poziomu** – pojęcie można wyjaśnić stosując uproszczenie, iż jest to język programowania zrozumiały dla ludzi.

**Problem klasy NP zupełny** – problem, którego nie da się rozwiązać w sensownym czasie, bez użycia dodatkowych założeń lub uproszczeń.

## Słowniczek dla informatyków

**Czas dywergencji** – czas, który upłynął od rozdzielenia się dwóch kładów.

**Enzym** – białko, przyspieszające reakcję, dzięki obniżeniu bariery energetycznej uniemożliwiającej jej zajście.

**Inhibitor** – związek chemiczny zatrzymujący lub spowalniający reakcję.

**Takson** – grupa organizmów spokrewnionych ze sobą, posiadających cechę odróżniającą ją od innych stworzeń.

**Kład** – grupa organizmów, najczęściej posiadających wspólnego przodka.

**NJ** – jedna z metod filogenetycznych (przyłączenie sąsiada). Polega na porównywaniu dystansów pomiędzy parami sekwencji utworzonymi wcześniej węzłami.

**Parsymonia** – podejście do filogenetyki zakładające minimum zmian pomiędzy najbliższymi bytami.

**Tranzycje i transwersje** – mutacje pojedynczych nukleotydów, odpowiednio purynowego w purynowy, pirymidynowego w pirymidynowy oraz purynowego w pirymidynowy i odwrotnie

**Uliniowanie** – dopasowanie sekwencji w celu odszukania podobnych regionów.

**UPGMA** – jedna z metod filogenetycznych. Ideą metody jest łączenie klastrow w większe grupy na podstawie średniej wszystkich odległości ewolucyjnych w zbiorze.

Kolo Naukowe Bioinformatyki



Dołącz  
do  
nas!

więcej informacji na:

[www.bit.ii.uj.edu.pl](http://www.bit.ii.uj.edu.pl)



/koloBITUJ

