

Podstawy bioinformatyki dla biotechnologów

Wykład 2
Definicja bioinformatyki

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2

Od jakiego pułapu startujemy?

Zakładamy, że te pojęcia są w małym palcu:

- DNA, RNA – struktura, funkcje, rodzaje
- Genom
- Białka – struktury, funkcje, rodzaje, wiązania peptydowe
- Równanie, funkcja, algorytm

Synteza przedmiotów: chemia, fizyka, biologia, matematyka, informatyka,

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 3

informatyka

Wykład inauguracyjny
Prof. dr. hab. inż. Mariana Adamskiego

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 5

plan

- Sprawy organizacyjne
- Definicje bioinformatyki
- Miejsce i dziedziny bioinformatyki
- Projekty bioinformatyczne

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 2

matematyka

Matematyka (z łac. mathematicus, od gr. μαθηματικός mathēmatikós, od μαθηματ-, μαθημα mathēmat-, mathēma, „nauka, lekcja, poznanie”, od μαθηθῆναι maththēnain, „uczyć się, dowiedzieć”; prawd. spokr. z goc. mundon, „baczyć, uważać”) – nauka dostarczająca narzędzi do otrzymywania ścisłych wniosków z przyjętych założeń.

Encyklopedia PWN. [26 lutego 2011]

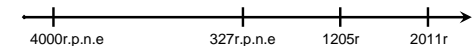
Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 4

Skale czasowe

- **matematyka**
- **biologia** (w tym szczególnie biologia Molekularna i genetyka)
- **cybernetyka / informatyka**
- **bioinformatyka**
- **nauka** (epokowe wydarzenia z każdej dziedziny nauki od początków jej historii, np. wynalezienie papieru, prochu strzelniczego, czcionka, lot balonem itp.)
- **religia** (np. Egipt i wrócenie z gwiazd, okres życia Jezusa, śmierć ostatniego apostoła, I sobór w Nicei, Luter itp.)
- **wydarzenia historyczne** (wojny, postaci historyczne, odkrycie Ameryki)



Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 6

multimedia

- Human genome sequencing-Animated tutorial
- DNA - The Human Race (i pozostałe 4 odcinki)
- DNA - The Search For Adam
- THE REAL EVE

Bioinformatyka

- Jest to dziedzina biologiczna wywodząca się z biotechnologii
- Aby zrozumieć bioinformatykę biolog/biotechnolog musi poznać kilka względnie prostych zasad matematyki i umieć obsługiwać komputer (nie koniecznie programować)
- Aby zrozumieć bioinformatykę informatyk/matematyk musi poznać i dogłębnie zrozumieć większość zasad i praw biologii (głównie molekularnej)

Bioinformatyka wg EBI

„Bioinformatyka stanowi multidyscyplinarne pole naukowe będące interfejsem pomiędzy biologią a informatyką.”

Cele:

- odkrycie bogactwa biologicznej informacji ukrytej w ogromnej ilości danych
- otrzymanie jaśniejszego wglądu w w fundamenty biologicznego organizmu

Przełożył: Adam Szlaski

Do poczytania

- http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml
- <http://www.sanger.ac.uk/about/history/hgp/>
- <http://www.ncbi.nlm.nih.gov/>
- <http://www.ebi.ac.uk/>
- <http://www.ddbj.nig.ac.jp/>
- <http://www.insdc.org/>
- <http://www.barcodeoflife.org/>
- <http://www.staff.amu.edu.pl/~ewas/>

Bioinformatyka wg NCBI

„Dziedzina nauki, w której biologia, informatyka i technologia informacyjna zostały scalone w jedną dyscyplinę.”

Cele:

- opracowanie nowych algorytmów i statystyki
- analiza i interpretacje różnych typów danych (m.in. nukleotydowych i aminokwasowych sekwencji, domen i struktur białkowych)
- opracowanie i wdrożenie narzędzi, które umożliwią efektywne zarządzanie dostępem i różnymi rodzajami informacji.

Przełożył: Adam Szlaski

Bioinformatyka wg BISTIC

Badania naukowe, rozwój lub aplikacje będące narzędziami obliczeniowymi umożliwiającymi poszerzenie możliwości wykorzystania danych biologicznych, medycznych, behawioralnych lub zdrowotnych w celu pozyskiwania, przechowywania, organizowania, archiwizacji lub wizualizacji tych danych

17 lipca 2000 – komitet powołany przez The Biomedical Information Science and Technology Initiative Consortium (www.bisti.nih.gov)

Przełożył: Adam Szlaski

Bioinformatyka

"The mathematical, statistical and computing methods that aim to solve biological problems using DNA and amino acid sequences and related information."

Fredj Tekaia (UMSC Paryż)

Bioinformatyka jest to dyscyplina nauk biologicznych zajmująca się stosowaniem narzędzi matematycznych i informatycznych do rozwiązywania problemów biologii (głównie biologii molekularnej) i zagadnień biotechnologicznych.

JP Jastrzębski

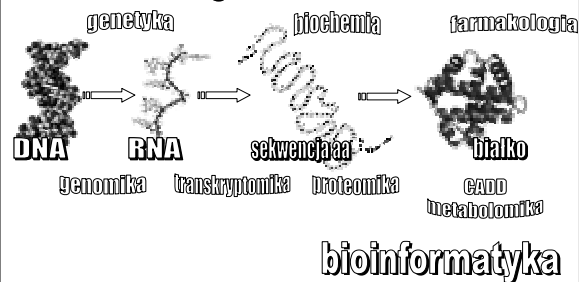
Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 13

Biotechnologia a bioinformatyka

biotechnologia



Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 15

Historyczne podłoże bioinformatyki

Lata 80' – United States Department of Energy tworzy GenBank

1.X.1990 – Human Genom Project (plan ukończenia 2005)

(United States Department of Energy, National Institutes of Health) - Dr. Francis Collins

1996 – zsekwencjonowanie genomu drożdży (13 milionów par zasad i 6 275 genów)

1997 – zsekwencjonowanie genomu *Caenorhabditis elegans* (13500 genów)

IV-V.1998 – debaty publiczne w Europie; dr Craig Venter i NIH w USA

II 2001 – publikacje w Nature i Science



<http://www.genom.gov/>



Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 17

Bioinformatyka

dyscyplina nauk biologicznych wywodząca się z biotechnologii (genetyki), zajmująca się stosowaniem narzędzi matematycznych i informatycznych do rozwiązywania problemów biologii (głównie biologii molekularnej) i zagadnień biotechnologicznych. Podstawowymi poddziedzinami bioinformatyki są: genomika, proteomika, transkryptomika i metabolomika.

in vivo – badania przyżyciowe; mało możliwości manipulacji

in situ – w tkance; ograniczone możliwości manipulacji

in vitro – w szkiełku; największe „naturalne” możliwości manipulacji

in silico – w komputerze; możliwość analizowania wszelkich, nawet pozornie niemożliwych układów

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 14

Dygresja - Obalenie dogmatów!!!

1 sekwencja DNA

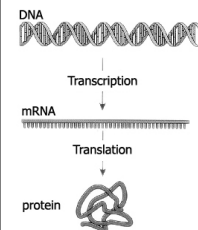
→ 1 sekwencja cDNA

→ 1 sekwencja mRNA

→ 1 sekwencja aa (str. I-rz.)

→ 1 struktura II-rzędowa

→ 1 struktura 3D



Źródło:
<http://encephalon.ca/>

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 16

genomika

dziedzina biologii molekularnej i biologii teoretycznej (pokrewna genetyce i ściśle związana z bioinformatyką) zajmująca się analizą genomu organizmów. Głównym celem genomiki jest poznanie sekwencji oraz mapowanie genomu ale również określenie wszelkich zależności i interakcji wewnątrz genomu.

W odróżnieniu od genetyki genomika obejmuje ogół zjawisk genetycznych całościowo i przy pomocy biologii teoretycznej (głównie bioinformatyki) stara się określić i opisać wszystkie zależności tych zjawisk oraz wpisać je w ogół procesów metabolicznych żywego organizmu.

- genomika funkcjonalna (poznanie funkcji wszystkich genów w genomie)
- genomika strukturalna (poznanie sekwencji i jej wstępny opis)
- genomika teoretyczna (ogólne prawa rządzące genomami)
- genomika porównawcza (ewolucja genomów)
- genomika indywidualnych różnic (zmienność międzypodgatunkowa genomów tego samego gatunku)

Dr Jan Paweł Jastrzębski

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 18

proteomika

gałąź nauki zajmująca się badaniem białek - ich struktury, sprawowanych przez nie funkcji i zależności między nimi.

Proteomika obejmuje analizę całych **proteomów** (zestaw wszystkich białek w komórce, liniach komórkowych, tkankach lub całych organizmach).

Proteomika jest dziedziną znacznie szerszą i bardziej złożoną niż genomika, ponieważ liczba genów kodujących białka jest znacznie mniejsza niż liczba białek w komórce (genów w komórce człowieka jest około 22 tysiące, natomiast białek mniej więcej 400 tysięcy) .
(Wikipedia)

Białka są polimerami 20 typów monomerów (aminoakwasy), DNA – 4 (ATCG)

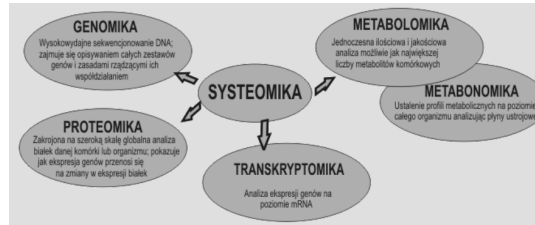
- proteomika funkcjonalna (*analiza funkcji wszystkich białek w proteomie*)
- proteomika strukturalna (*poznanie struktury przestrzennej białek*)
- proteomika teoretyczna (*ogólne prawa rządzące proteomem*)
- proteomika porównawcza (*ewolucja białek i analiza miejsc zmienności genetycznej*)
- proteomika indywidualnych różnic (*zmienność międzysobnicza proteomów i poszczególnych białek tego samego gatunku*)

Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 19

Wszystko...omika



Koło Naukowe przy Katedrze i Zakładzie Biofarmacji i Farmakodynamiki
Akademii Medycznej w Gdańsku
<http://www.sbs.farmacja.amg.gda.pl/biofarmacja/>

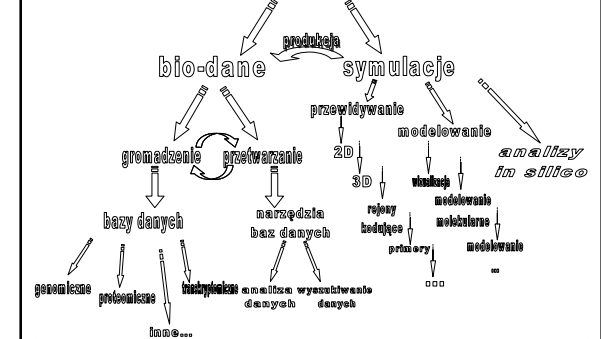
Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 21

Bioinformatyka

(dla biotechnologów)



Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 23

Transkryptomika i metabolomika

Transkryptomika - jest to dziedzina, za pomocą której określane jest miejsce i czas aktywności genów poprzez badanie transkryptomu, czyli ogółu cząsteczek mRNA znajdujących się w danym momencie w komórce.

Metabolomika - dziedzina nauki zajmująca się badaniem zestawu wszystkich metabolitów obecnych w organizmie, tkance czy komórce - metabolomu. Jest zaliczana obok genomiki, transkryptomiki i proteomiki do biologii systemowej.

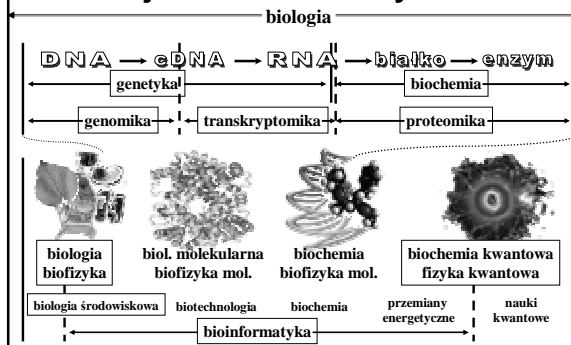
Biologia systemowa jest dziedziną nauki zajmującą się badaniem złożonych oddziaływań występujących w systemach biologicznych. Biologia systemowa łączy informacje zdobywane przez dziedziny nauki takie jak: genomika, transkryptomika, proteomika i metabolomika.

Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 20

Miejsce bioinformatyki w nn

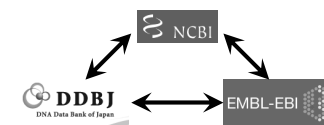


Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 22

NCBI - EBI - DDBJ



NCBI – GenBank
EBI – EMBL
DDBJ – DDBJ

(<http://www.ncbi.nlm.nih.gov/Genbank/index.html>)
(<http://www.ebi.ac.uk/emb1/>)
(<http://www.ddbj.nig.ac.jp/index-e.html>)

DDBJ/EBI/NCBI – International Nucleotide Sequence Database Collaboration

- GenBank → There are approximately 85,759,586,764 bases in 82,853,685 sequence records in the traditional GenBank divisions and 108,635,736,141 bases in 27,439,206 sequence records in the WGS division as of February 2008.

Dr Jan Paweł Jastrzębki

KFIBR, Wydział Biologii
UWM

Podstawy bioinformatyki
Wykład 2 / 24

Serwis NCBI

NCBI
National Center for Biotechnology Information

Established in 1988 as a national resource for molecular biology information, NCBI creates public databases, conducts research in computational biology, develops software tools for analyzing genome data, and disseminates biomedical information - all for the better understanding of molecular processes affecting human health and disease. [More...](#)

What does NCBI do?

- Assembly Archive
- Clusters of orthologous groups
- Coffee Break, Genes & Disease, NCBI Handbook
- Electronic PCR
- Entrez Home
- Entrez Tools
- Gene expression omnibus (GEO)

Whole Genome Association

The NCBI Whole Genome Association (WGA) resource provides researchers with access to genotype and associated phenotype information that will help elucidate the link between genes and disease. For more information, click here to see the [WGA](#) resource page and click here to read the [press release](#).

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 25

Bazy danych NCBI (od sitemap)

NCBI Site Map

Bazy danych NCBI (od sitemap)

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 27

Inne zadania bioinformatyki

- molekularne

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 29

NCBI - mapa strony

NCBI - mapa strony

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 26

Bazy danych NCBI - Entrez

Bazy danych NCBI - Entrez

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 28

Modelowanie dynamiki molekularnej

Generating constant water flow in a molecular dynamics simulation

1. Constrain terminal atoms
2. Push water molecules near the left edge of the water box to generate flow.
3. Flow induces conformation change

Dr Jan Paweł Jastrzębki KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 30

wizualizacja

Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 31

Modelowanie homologiczne

Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 33

Inne zadania bioinformatyki

- Globalne / środowiskowe

Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 35

Modelowanie *ab initio*

Protein Folding

Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 32

Projektowanie leków - CADD

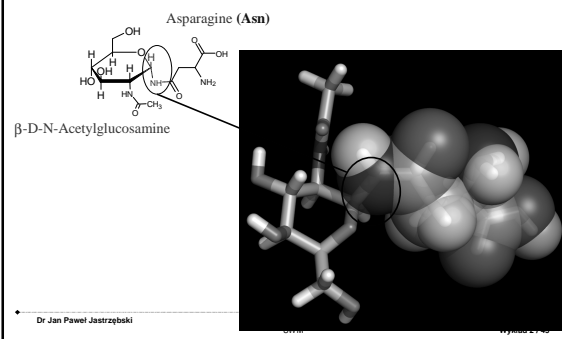
Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 34

mitochondrialna Ewa

Dr Jan Paweł Jastrzębski KFIBR, Wydział Biologii UWM Podstawy bioinformatyki Wykład 2 / 36

EXPANSION TIMES (YEARS AGO)	
Africa	120,000 - 150,000
Out of Africa	55,000 - 75,000
Asia	40,000 - 70,000
Australia/PNG	40,000 - 60,000
Europe	35,000 - 50,000
Americas	15,000 - 35,000
Na-Dene/Esk/Aleuts	8,000 - 10,000

Asn—GlcNAc N-linked glycosylation site



Model bakterii