

BIOINFORMATIKA

aneb jak se vypořádat s datovou explozí v biologii a biomedicině

Setkání Krajských inovačních platforem RIS3 strategie

Královéhradeckého kraje

25. 6. 2019, 9:30; Přírodovědecká fakulta UHK

Radka Symonová
Katedra biologie, PŘF UHK



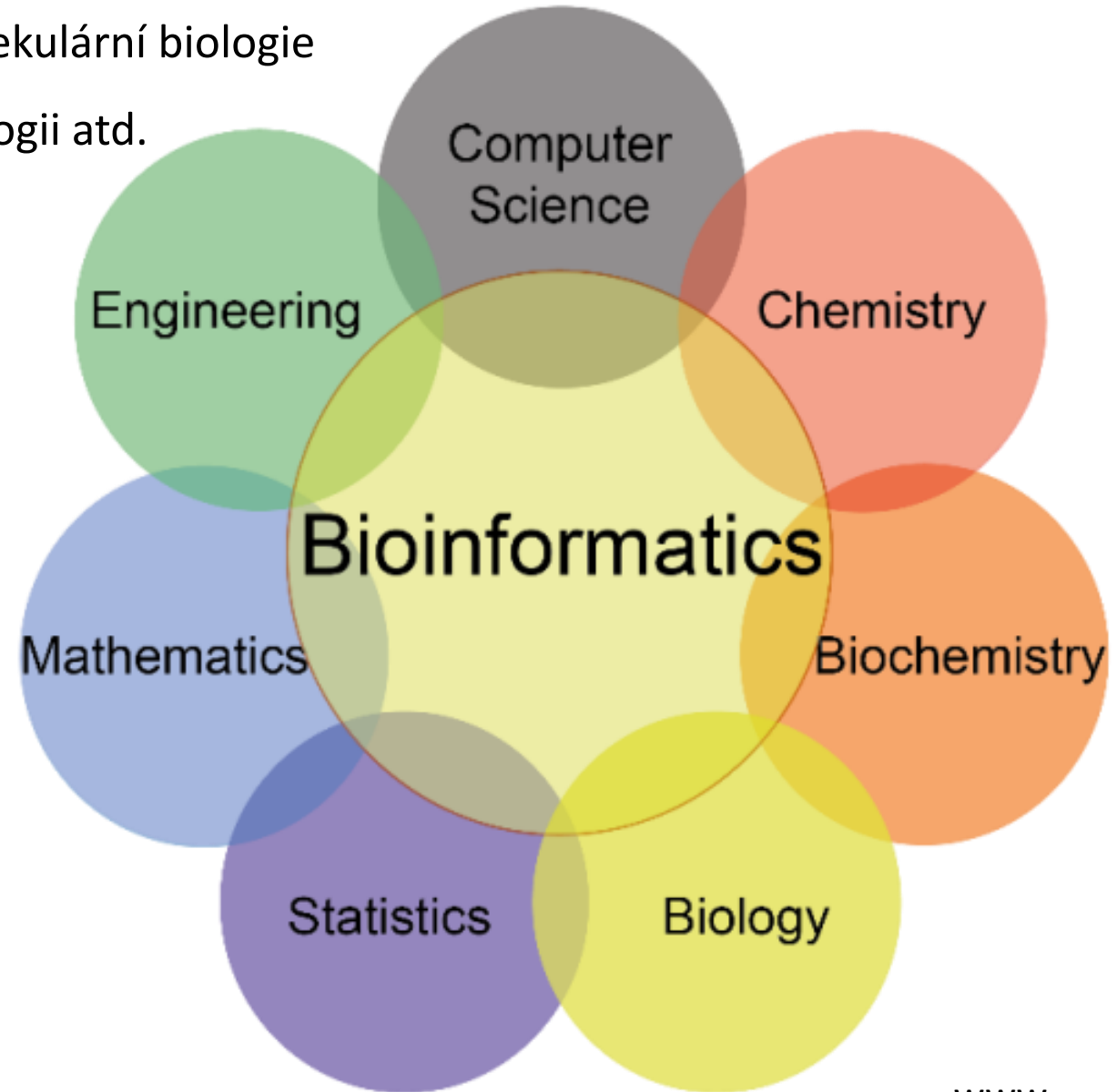
Univerzita
Hradec Králové
Přírodovědecká
fakulta

Co je bioinformatika?

= nové interdisciplinární odvětví, které se oddělilo od molekulární biologie

Zčásti jde o servisní obor pro biologii, medicínu, farmakologii atd.

Zčásti obsahuje specializované výzkumné směry z oblasti aplikované informatiky týkající se ***informační struktury biologických dat*** a algoritmů vhodných pro jejich prohledávání a interpretaci.

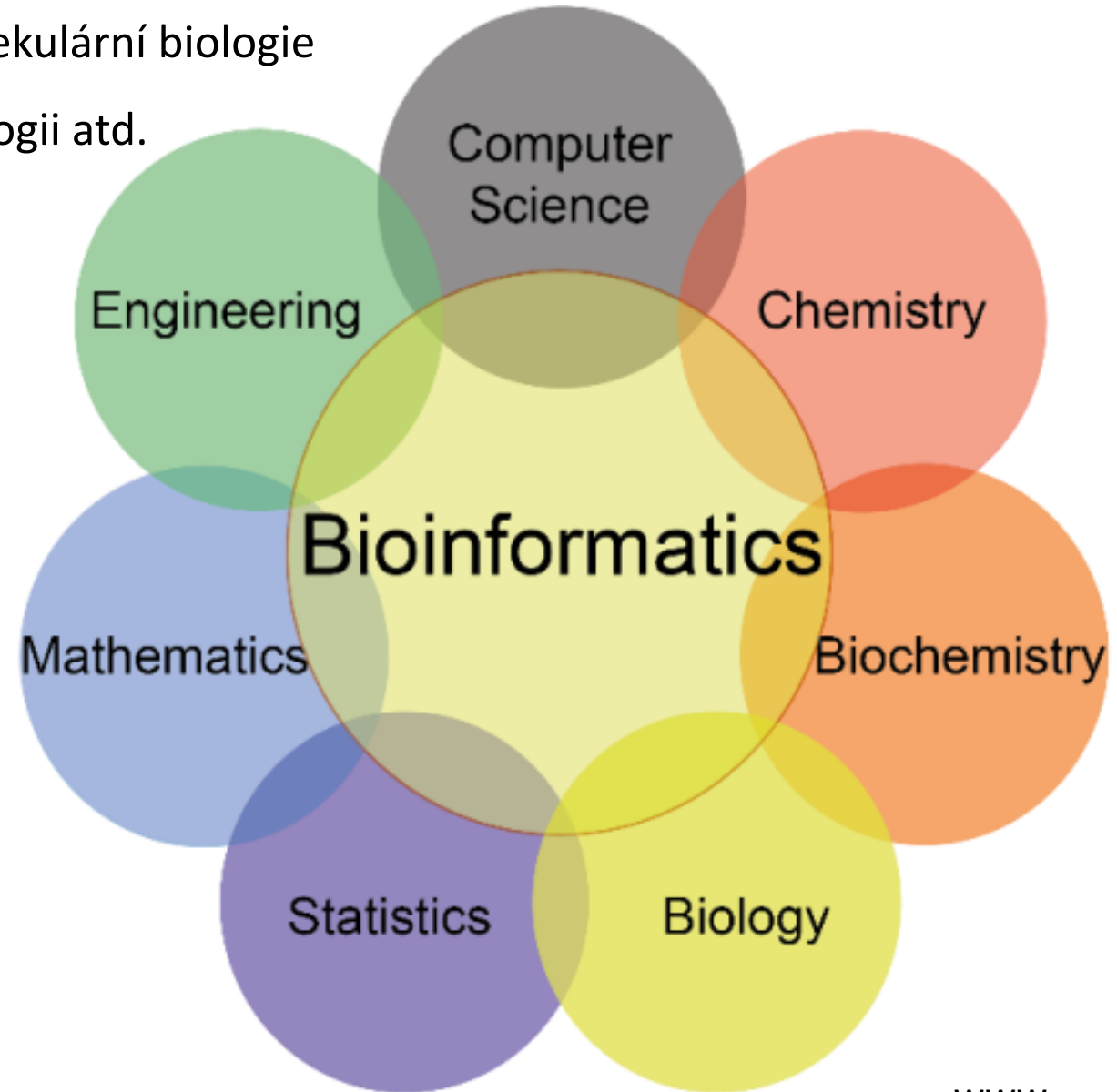
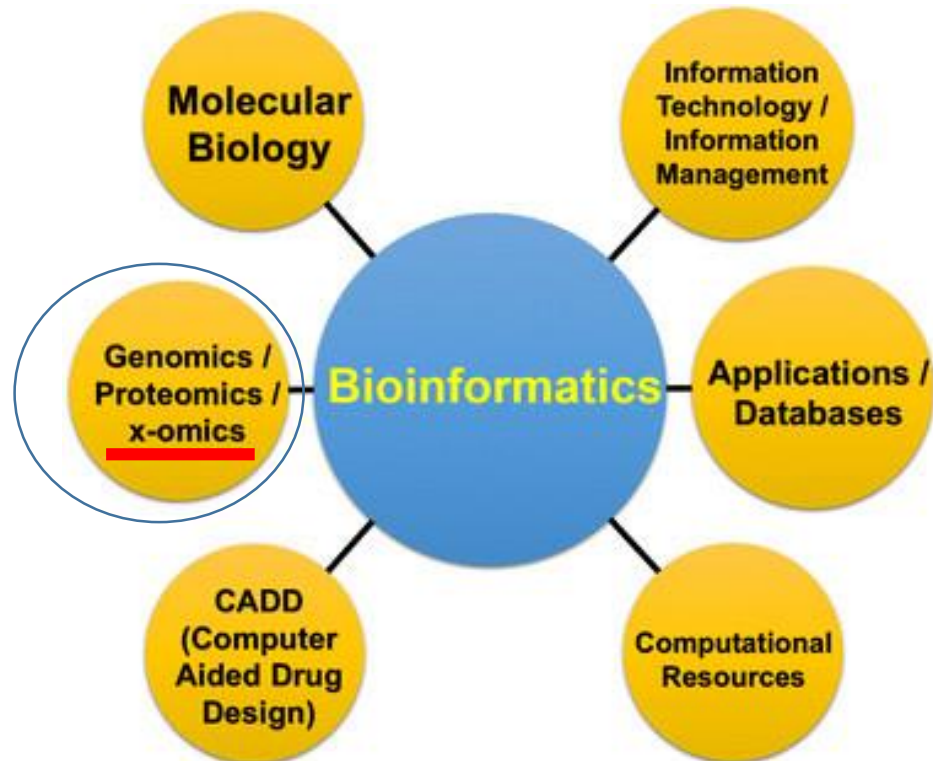


Co je bioinformatika?

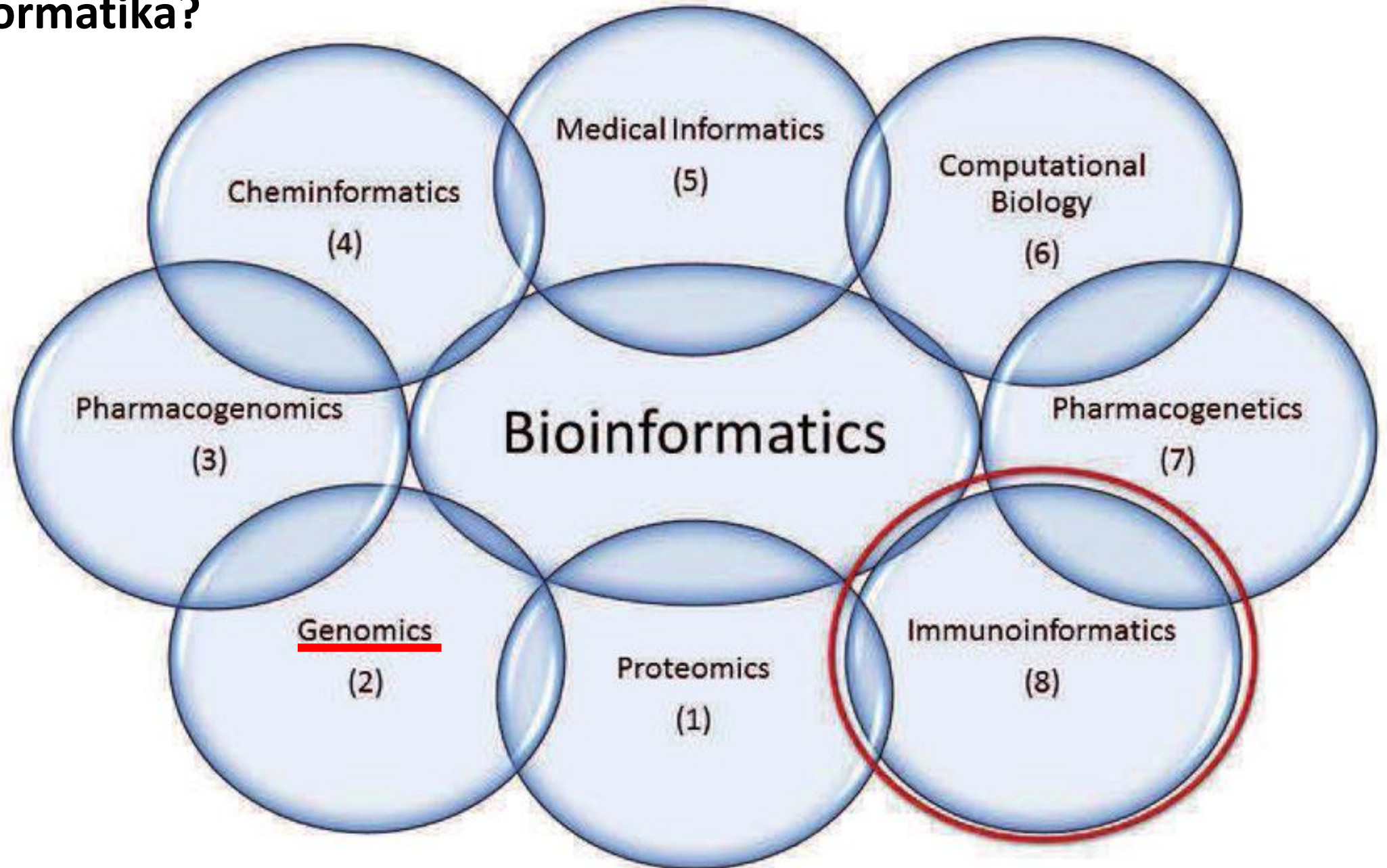
= nové interdisciplinární odvětví, které se oddělilo od molekulární biologie

Zčásti jde o servisní obor pro biologii, medicínu, farmakologii atd.

Zčásti obsahuje specializované výzkumné směry z oblasti aplikované informatiky týkající se ***informační struktury biologických dat*** a algoritmů vhodných pro jejich prohledávání a interpretaci.



Co je bioinformatika?



Definice bioinformatiky podle NIH vědecko-technologického konsorcia pro biomedicínské informace

Výzkum, vývoj nebo aplikace výpočetních nástrojů za účelem zvyšování rozvoje využití biologických a lékařských dat, včetně těch, které umožňují taková data získávat, ukládat, organizovat, archivovat, analyzovat nebo vizualizovat.



National Institutes of Health
Turning Discovery Into Health

Preamble

Bioinformatics and computational biology are rooted in life sciences as well as computer and information sciences and technologies. Both of these interdisciplinary approaches draw from specific disciplines such as mathematics, physics, computer science and engineering, biology, and behavioral science. Bioinformatics and computational biology each maintain close interactions with life sciences to realize their full potential. Bioinformatics applies principles of information sciences and technologies to make the vast, diverse, and complex life sciences data more understandable and useful. Computational biology uses mathematical and computational approaches to address theoretical and experimental questions in biology. Although bioinformatics and computational biology are distinct, there is also significant overlap and activity at their interface.

Definition

The NIH Biomedical Information Science and Technology Initiative Consortium agreed on the following definitions of bioinformatics and computational biology recognizing that no definition could completely eliminate overlap with other activities or preclude variations in interpretation by different individuals and organizations.

Bioinformatics: Research, development, or application of computational tools and approaches for expanding the use of biological, medical, behavioral or health data, including those to acquire, store, organize, archive, analyze, or visualize such data.

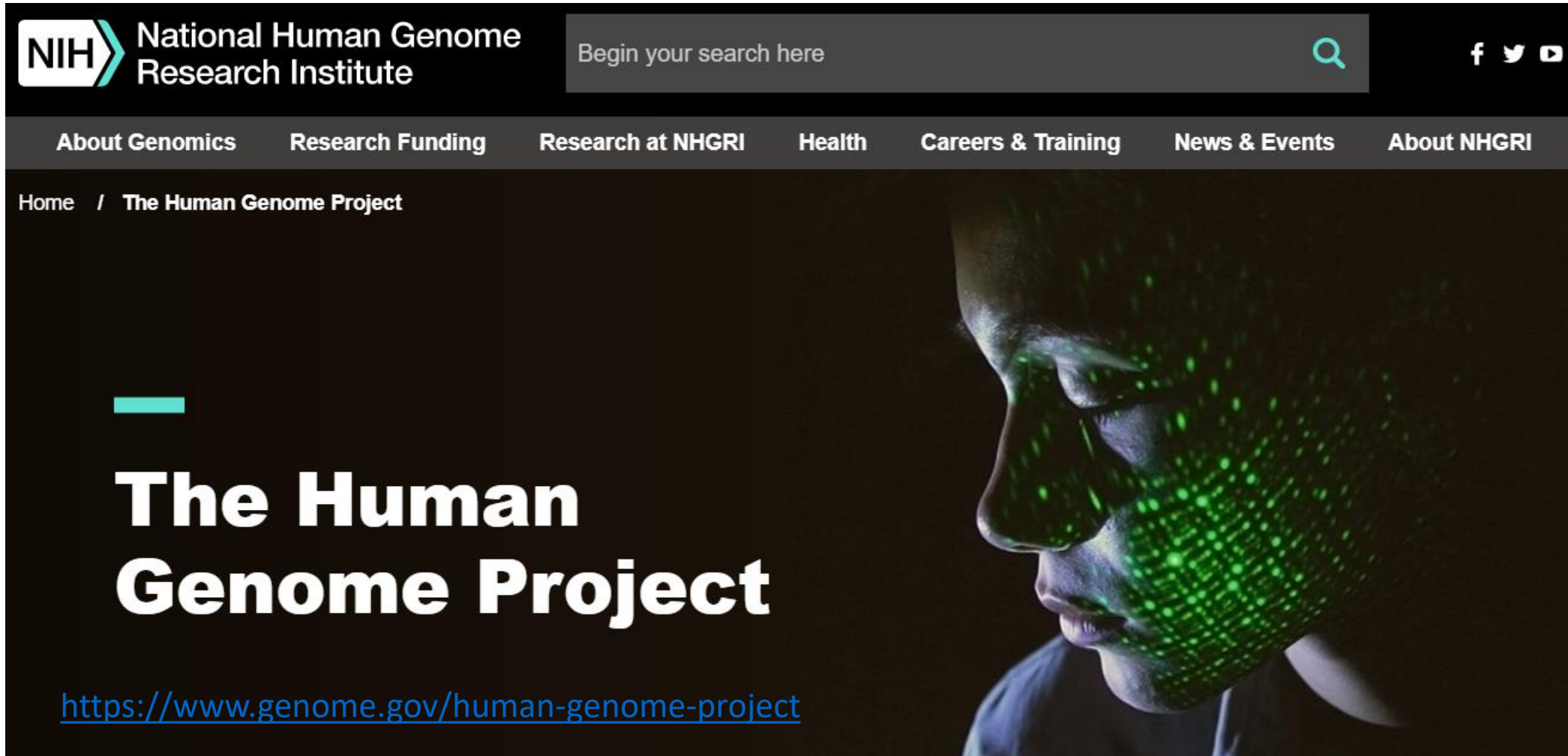
Computational Biology: The development and application of data-analytical and theoretical methods, mathematical modeling and computational simulation techniques to the study of biological, behavioral, and social systems.

Oblasti využití bioinformatiky

- ◆ **Microbial genome applications**
- ◆ **Molecular medicine**
- ◆ **Personalised medicine**
- ◆ **Preventative medicine**
- ◆ **Gene therapy**
- ◆ **Drug development**
- ◆ **Antibiotic resistance**
- ◆ **Evolutionary studies**
- ◆ **Waste cleanup**
- ◆ **Biotechnology**
- ◆ **Climate change Studies**
- ◆ **Alternative energy sources**
- ◆ **Crop improvement**
- ◆ **Forensic analysis**
- ◆ **Bio-weapon creation**
- ◆ **Insect resistance**
- ◆ **Improve nutritional quality**
- ◆ **Development of Drought resistant varieties**
- ◆ **Vetinary Science**

(Lapidus AL: Bioinformatics and its applications)

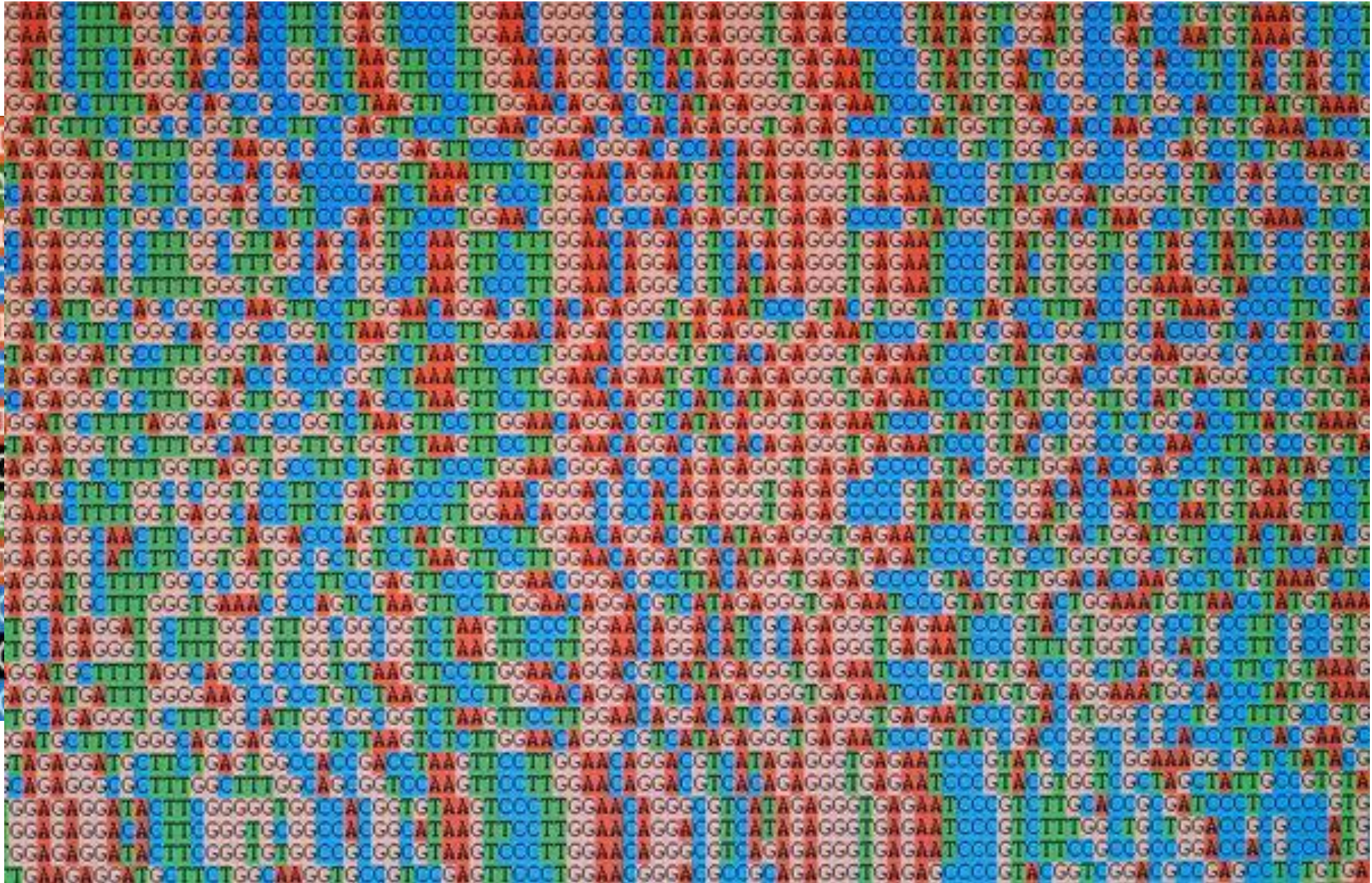
Počátky bioinformatiky.... HUGO 1990 – 2003



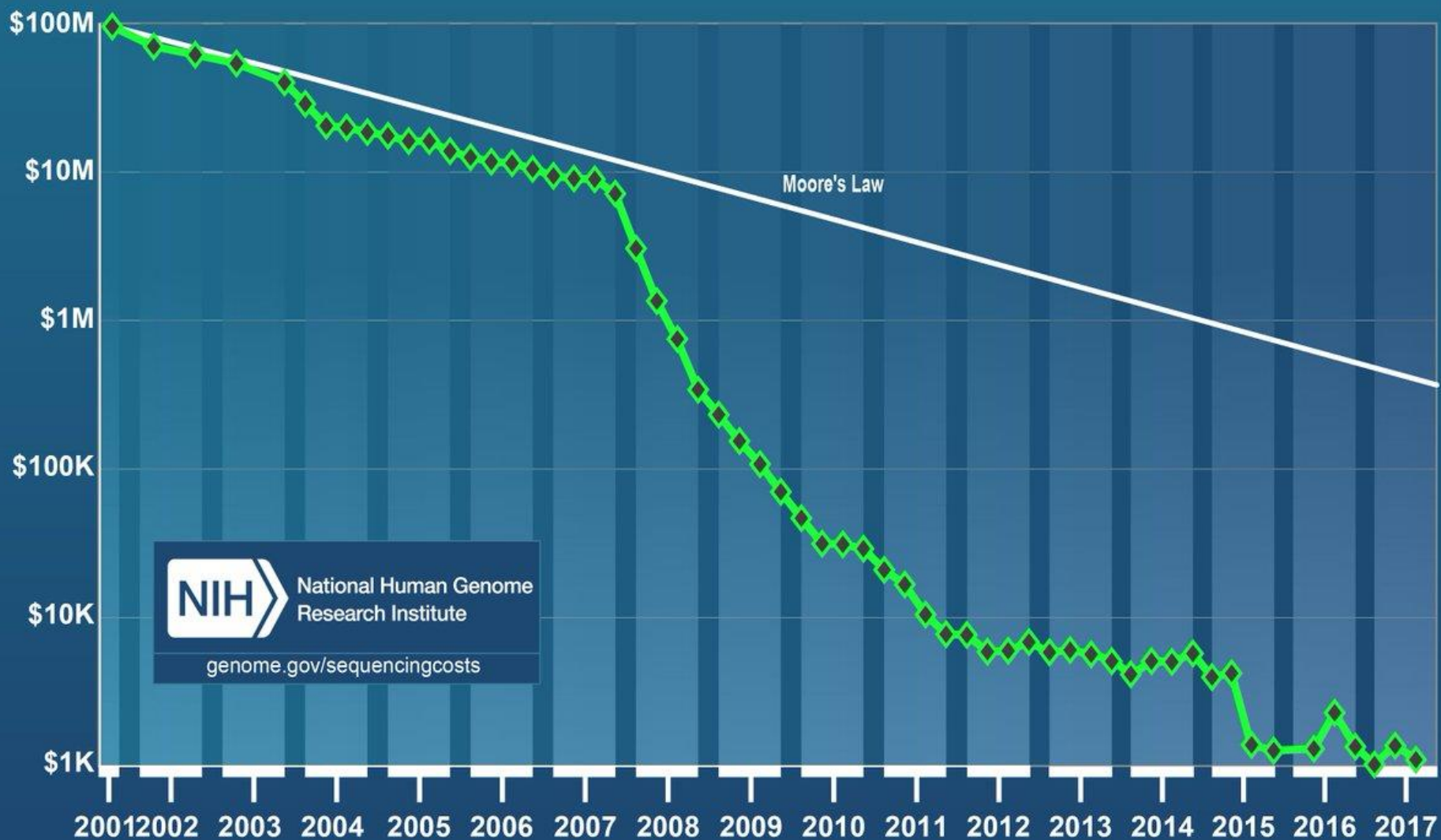
The Human Genome Project (HGP) was one of the great feats of exploration in history. Rather than an outward exploration of the planet or the cosmos, the HGP was an inward voyage of discovery led by an international team of researchers looking to sequence and map all of the genes -- together known as the genome -- of members of our species, *Homo sapiens*. Completed in April 2003, the HGP gave us the ability, for the first time, to read nature's complete genetic blueprint for building a human being.

Objektem bioinformatiky je genom...genomika

A, T, G, C



Cost per Genome

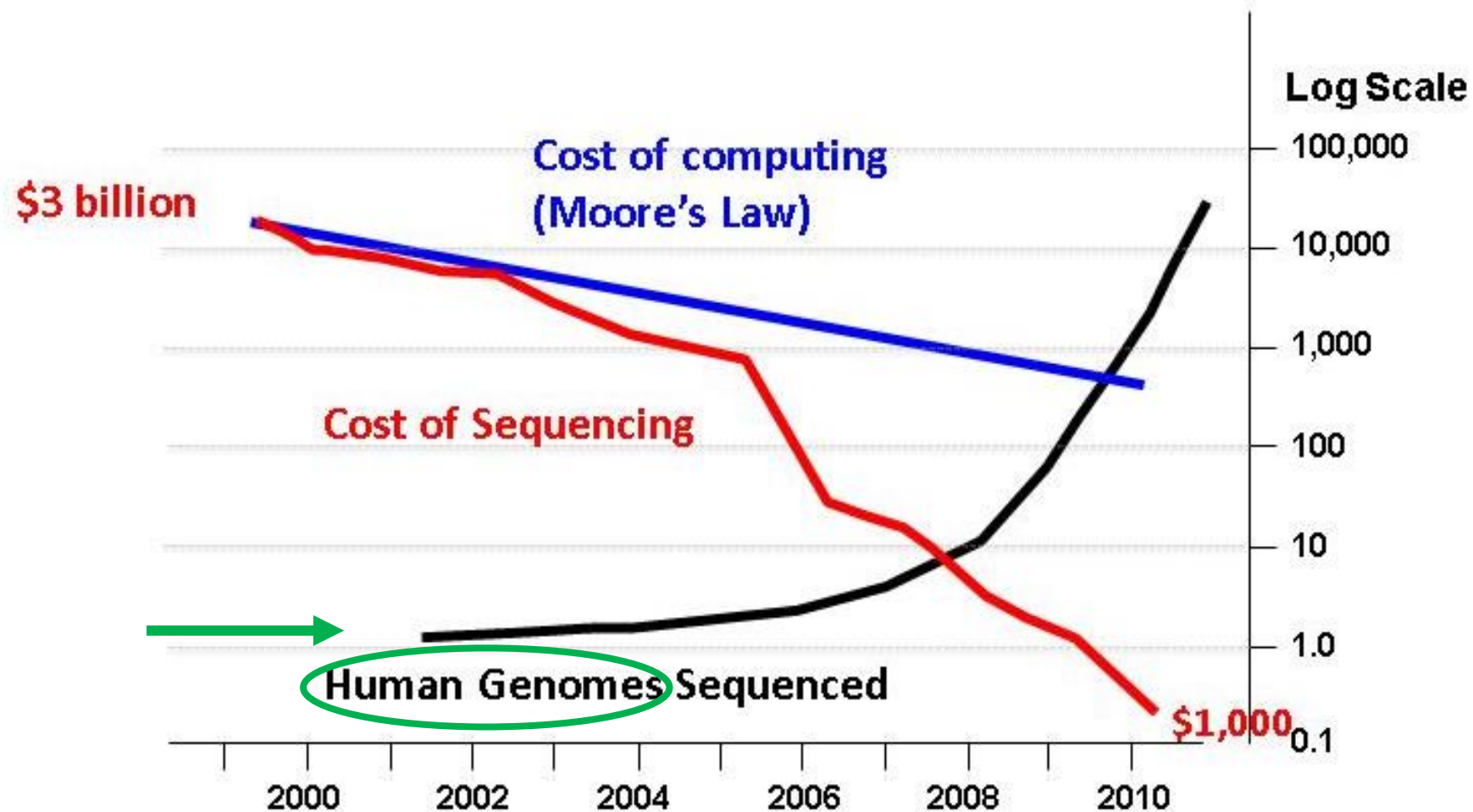


Evoluce nákladů
spojených
s informatikou
a bioinformatikou
a její důsledky

Evoluce nákladů spojených s informatikou a bioinformatikou a její důsledky

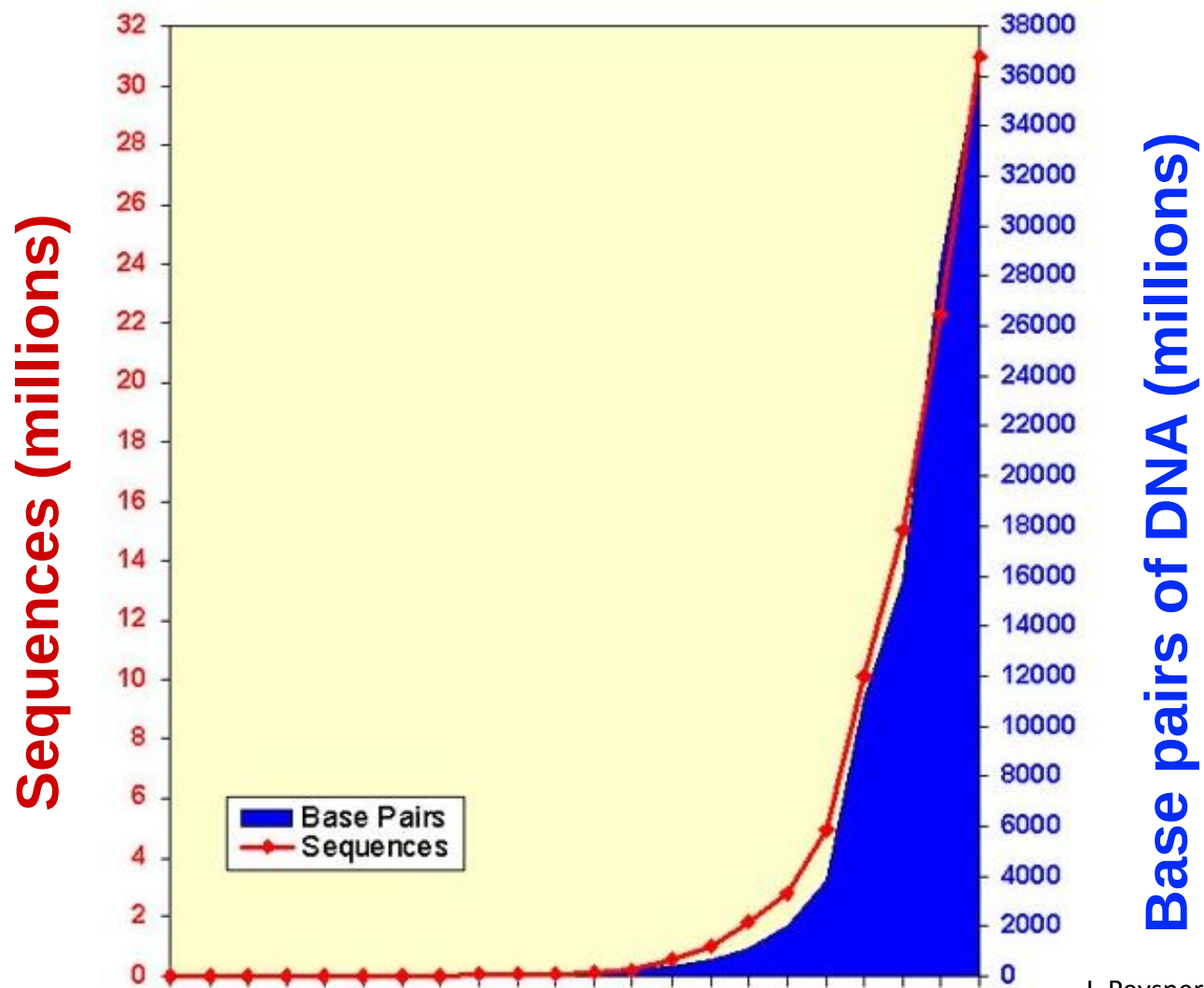


The Sequencing Explosion



Důsledky evoluce bioinformatiky...datové exploze všude v biologii

Growth of GenBank

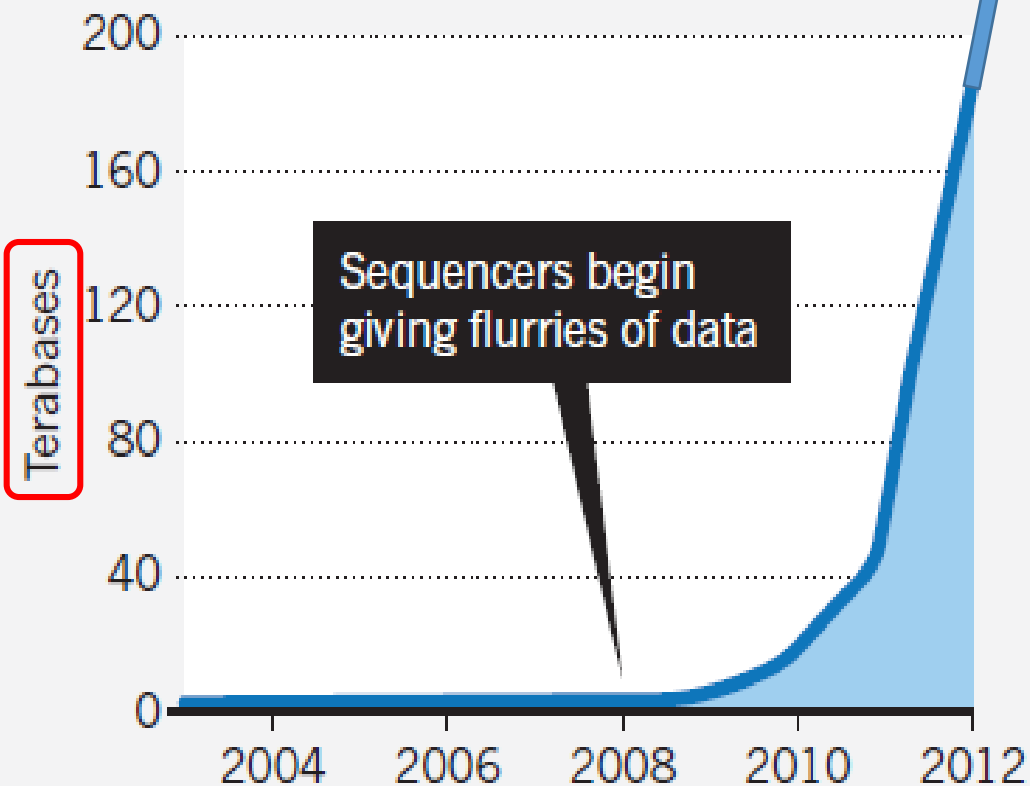


SOURCE: EMBL-EBI

J. Pevsner,
<http://www.bioinfbook.org/index.php>

DATA EXPLOSION

The amount of genetic sequencing data stored at the European Bioinformatics Institute takes less than a year to double in size.

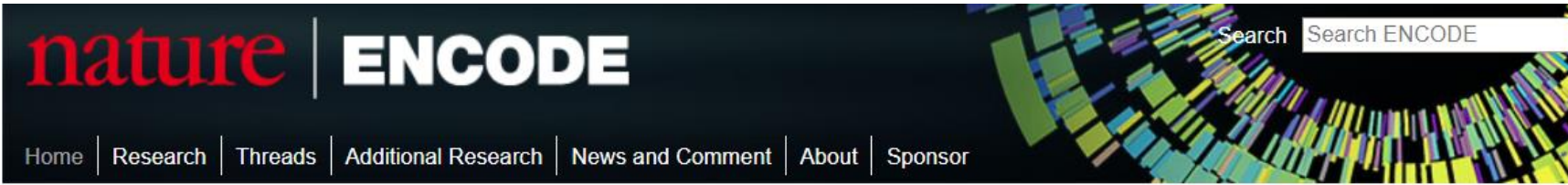


Důsledky evoluce bioinformatiky...datové exploze všude v biologii a medicíně

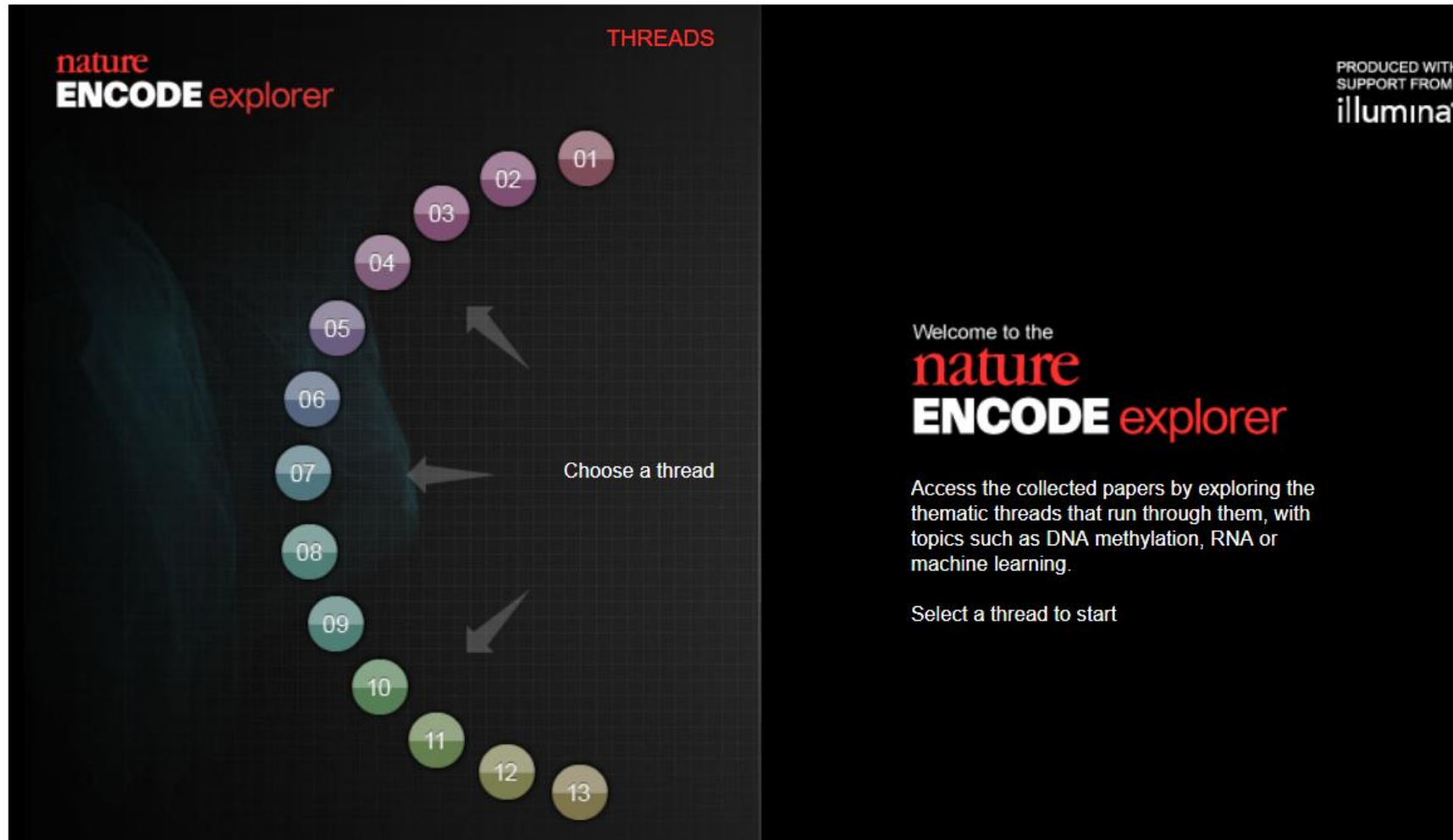
Již nyní jsou k dispozici tisíce osekvenovaných genomů virů, (arche)bakterií, prvoků, živočichů, rostlin, hub, hospodářsky významných rostlin a zvířat, patogenů, atd., atd.

Další tisíce jsou právě v této chvíli sekvenovány, asemblovány

Reakce na evoluci bioinformatiky a datové exploze?databáze



Ca.2,000 – 3,000
biomedicínských databází
→ bioinformatika =
(mj.) práce s databázemi

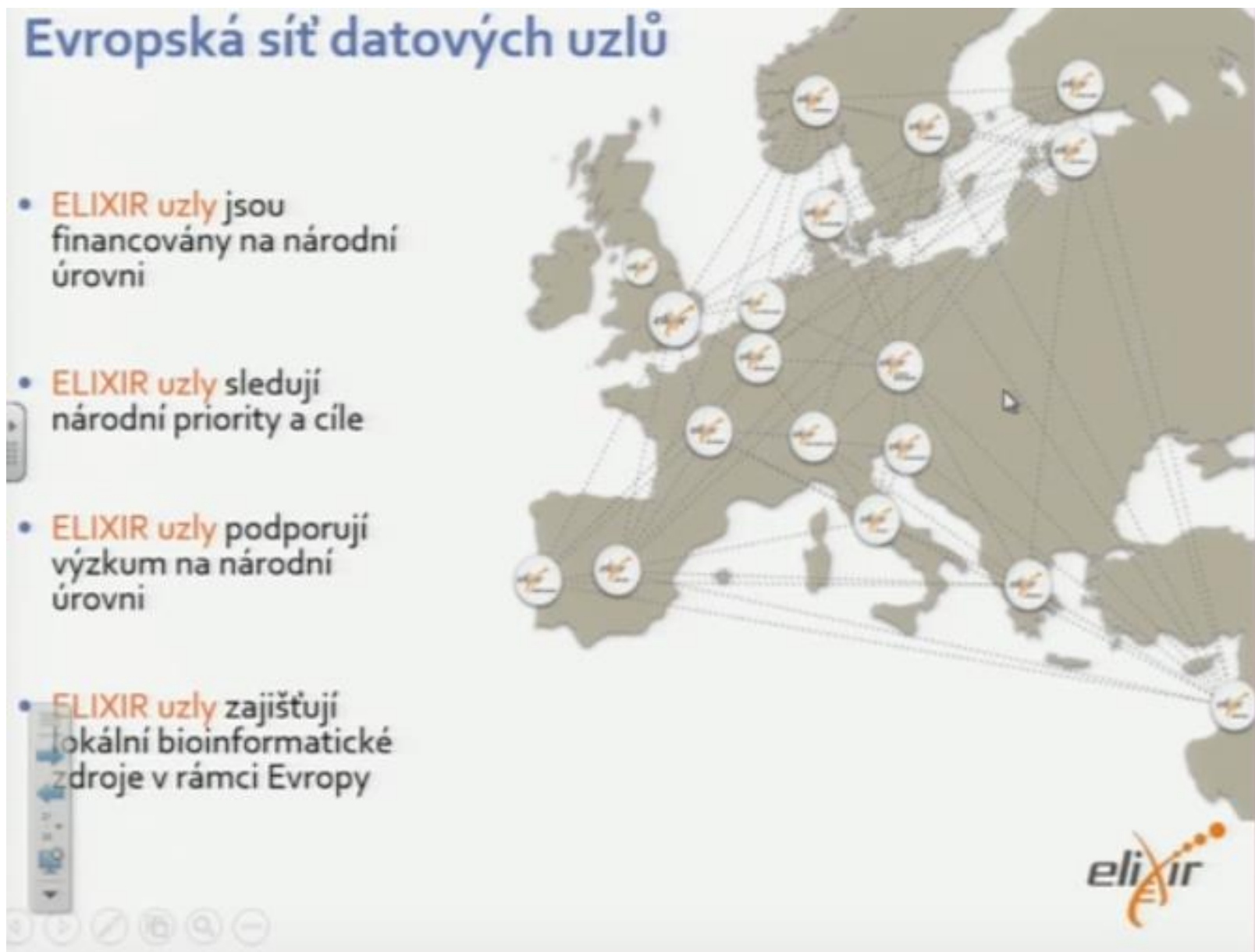


**Networking a
mezinárodní spojení
v úsilí vypořádat se s
biologickými a
medicínskými daty**

<http://www.nature.com/encode>

Networking a mezinárodní spojení v Evropě

<https://elixir-europe.org/>



*ELIXIR spojuje národní
bioinformatická centra a
EMBL-EBI do jedné udržitelné
evropské infrastruktury
pro biologická data*



Příklady využití bioinformatického boomu v medicíně



The 100,000 Genomes Project

The project has sequenced 100,000 genomes from around 85,000 people. Participants are NHS patients with a rare disease, plus their families, and patients with cancer.

2013-2018

The aim is to create a new genomic medicine service for the NHS – transforming the way people are cared for. Patients may be offered a diagnosis where there wasn't one before. In time, there is the potential of new and more effective treatments.

The project will also enable new medical research. Combining genomic sequence data with medical records is a ground-breaking resource. Researchers will study how best to use genomics in healthcare and how best to interpret the data to help patients. The causes, diagnosis and treatment of disease will also be investigated. We also aim to kick-start a UK genomics industry. This is currently the largest national sequencing project of its kind in the world.

Světový leader v bioinformatice a genomice - BGI



- BGI Research >
- China National Genebank >
- BGI College >
- GigaScience >
- BGI Genomics >
- BGI Manufacturing >
- BGI FGI >



Světový leader v bioinformatice a genomice - BGI



Bioinformatické centrum HK
Podpořený projekt na UHK



Erasmus+

Strategic Partnerships for higher education
(KA203)

Bioinformatics Centre Hradec Kralove

Total grant: 151264 EUR

Application Owner: SYMONOVA Radka

Form ID: KA203-7BA9D4F2

Last Submission: 26-03-2019

10:27:45 (Brussels Time)

Hradečtí partneři Bioinformatického centra HK



Univerzita Hradec Králové
Přírodovědecká fakulta

5



FACULTY OF PHARMACY
IN HRADEC KRÁLOVÉ
Charles University
1969–2019

Asociovaní hradečtí partneři Bioinformatického centra HK



Biskupské gymnázium
Hradec Králové

KRÁLOVÉHRADECKÝ
KRAJ



Alena Svátková

odbor školství | vedoucí oddělení projektů ve vzdělávání



Zahraniční strategičtí partneři Bioinformatického centra HK

-2 evropské firmy organizující bioinformatické kurzy a workshopy po EU



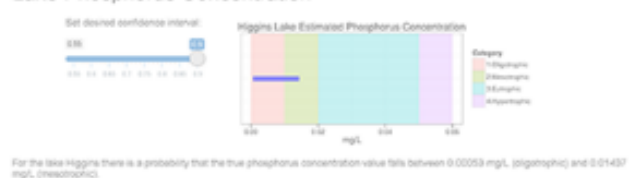
<https://www.transmittingscience.org/>

Explore Higgins Lake Model Coefficient Values

This Shiny application allows the user to interactively explore the sensitivity of Higgins Lake phosphorus concentration estimators to model coefficient values as defined in the Reichow and Simpson (2002) model. The default coefficient values represented below are the low, high, and most likely coefficients values published by Reichow and Simpson (2002) based on literature review and expert judgement. You can adjust the most likely values and observe subsequent changes in the expected phosphorus levels.



Lake Phosphorus Concentration

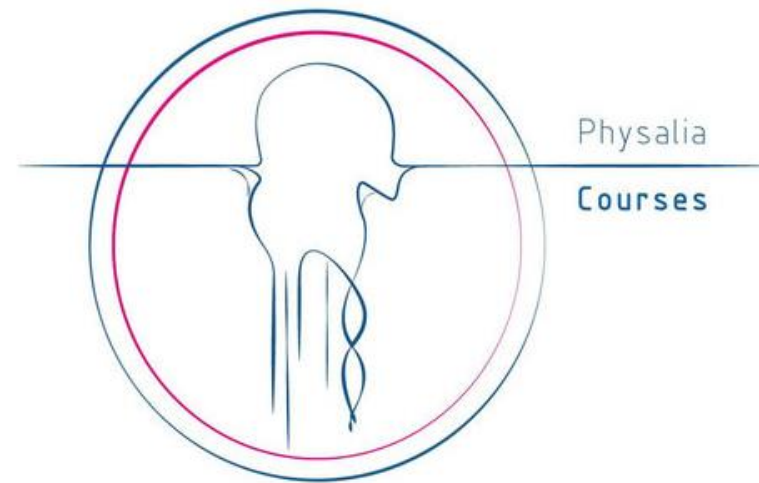


© Ashton Drew

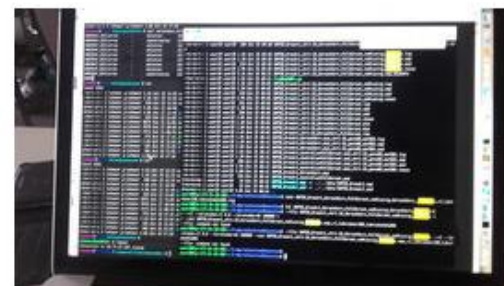
Interactive Data Analysis and Visualization
with R Shiny
Next Edition Coming Soon



Manipulation of NGS Data for Genomic and
Population Genetics Analyses
June 17-20, 2019 (Spain)



<https://www.physalia-courses.org/>



Introduction to Bash and
Python programming
Berlin, 14-18 October 2019

Instructor:
Dr. Martin Jones



Advanced R Programming
Berlin, 1-5 July 2019

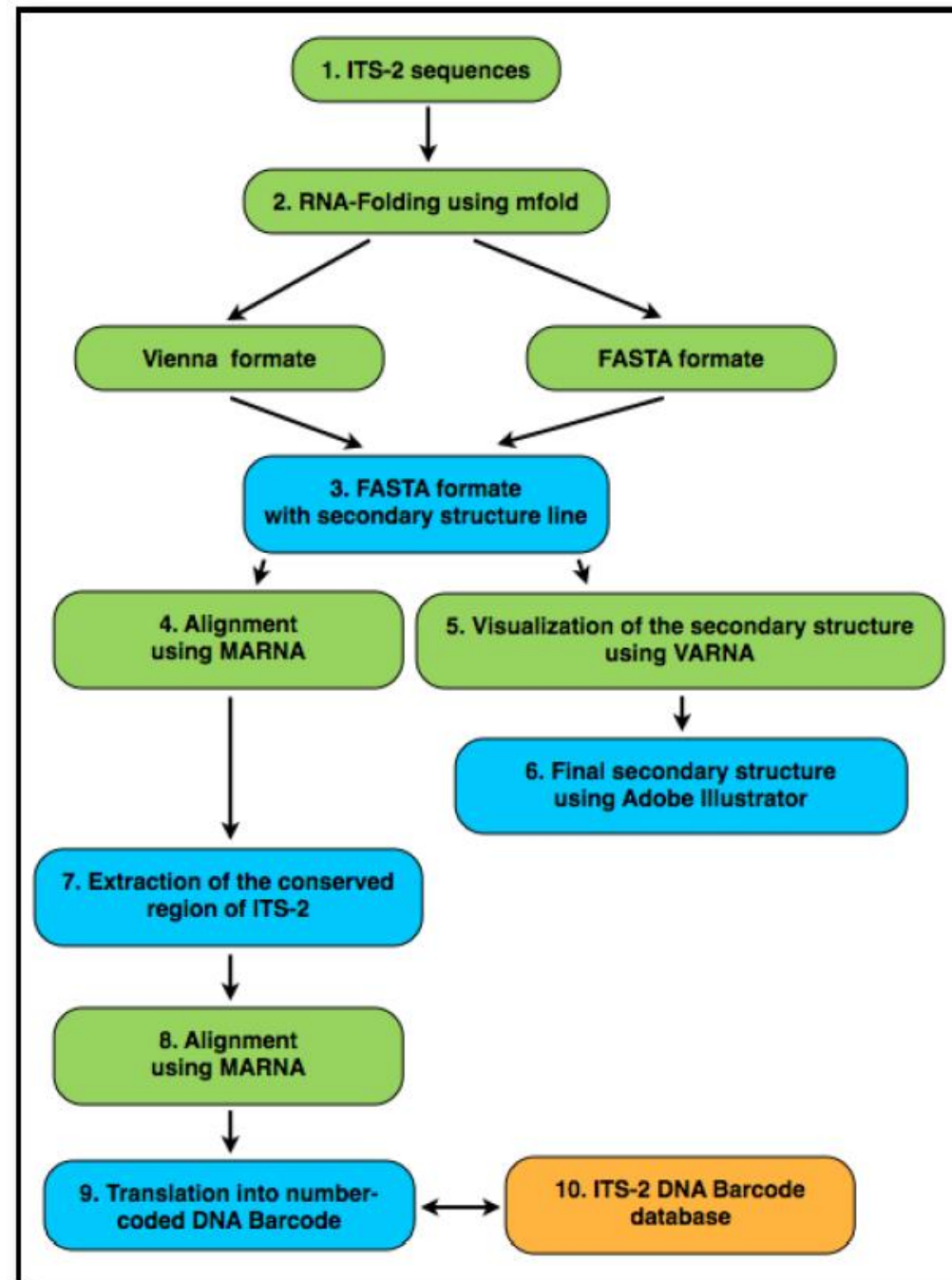
Instructor:
Dr. January Weiner

Zahraniční strategický a vědecký partner
Bioinformatického centra HK



Faculty of Biology

Bioinformatics in rRNA secondary structure



Potenciál pro etablování bioinformatiky na UHK



V Česku studuje informatiku jen 15 % dívek. Nejvíce jich je v Řecku a Bulharsku, nejméně v Belgii

Podle zjištění analytického projektu Evropa v datech činí podíl vysokoškolských studentek informačních technologií v evropských státech v průměru 18 %. V České republice tento podíl dosahuje jen 15 %. Nejlépe z celé Evropy jsou na tom v tomto směru Řecko a Bulharsko, kde množství studentek informatiky přesáhlo 30 %. Nejnižší podíl má potom Belgie s pouhými 8 %. Co do počtu ženy naopak převládají ve studiu přírodních věd.

autor -rd | Události | 25.03.2019

Biologie PřF UHK:

1:8 až **1:10**

Dívky

Chlapci

Co do počtu ženy naopak převládají ve studiu přírodních věd.

Evoluční genomika

- *evoluce architektury genomu obratlovců
- *rDNAom(ika) - diverzifikované role nejstarší a nejuniverzálnější frakce genomu eukaryot
- *mobilom(ika) - mobilní genetické elementy (transpozony) v evoluci genomu

pyladies



***rDNAom(ika)** - role nejstarší a nejuniverzálnější frakce genomu eukaryot
-etablovaná mezinárodní spolupráce Innsbruck, Alabama, Barcelona, Gdaňsk



Symonová et al. *BMC Genomics* (2017) 18:391
DOI 10.1186/s12864-017-3774-7

BMC Genomics



RESEARCH ARTICLE

Open Access

Higher-order organisation of extremely amplified, potentially functional and massively methylated 5S rDNA in European pikes (*Esox* sp.)

Radka Symonová^{1*}, Konrad Ocalewicz^{2,3}, Lech Kirtiklis⁴, Giovanni Battista Delmastro⁵, Šárka Pelikánová⁶, Sonia Garcia⁷ and Aleš Kovařík^{8*}



Review

Integrative rDNAomics—Importance of the Oldest Repetitive Fraction of the Eukaryote Genome

Radka Symonová

Faculty of Science, Department of Biology, University of Hradec Králové,
500 03 Hradec Králové, Czech Republic; radka.symonova@gmail.com



Review

Vertebrate Genome Evolution in the Light of Fish Cytogenomics and rDNAomics

Radka Symonová^{1,*} and W. Mike Howell²

Chromosoma
<https://doi.org/10.1007/s00412-017-0651-8>

ORIGINAL ARTICLE



Evolutionary trends in animal ribosomal DNA loci: introduction to a new online database

Jana Sochorová¹ · Sònia Garcia² · Francisco Gálvez³ · Radka Symonová⁴ · Aleš Kovařík¹

Received: 24 February 2017 / Revised: 6 November 2017 / Accepted: 7 November 2017
© The Author(s) 2017. This article is an open access publication

Bioinformatika na katedře biologie PŘF UHK

Evoluční genomika

*evoluce architektury genomu obratlovců

***mobilom(ika)** - mobilní genetické elementy (transpozony) v evoluci genomu



Univerzita
Hradec Králové
Přírodovědecká
fakulta



UPPSALA
UNIVERSITET

Evolutionary Biology Centre

Bioinformatika na katedře biologie PŘF UHK

Code Week 2019 (říjen 2019)



Univerzita
Hradec Králové
Přírodovědecká
fakulta

CodeWeek. 

Poděkování

Královéhradeckému kraji
za finanční podporu pro rozvoj
bioinformatiky na PŘF UHK

