The Biomedical Informatics Research Group (BMI)

16 Aug 2017

Torbjørn Rognes
Department of Informatics, UiO
torognes@ifi.uio.no



UiO : University of Oslo

Study program and options

Study program:

- Informatics: technical and scientific applications
  (Informatikk: tekniske og naturvitenskapelige anvendelser)

Program options:

- Computational science with applications
  (Anvendt informatikk)
- Imaging technology and biomedical informatics
  (Avbildningsteknologi og biomedisinsk informatikk)

Research groups:

- DSB: Digital Signal Processing and Image Analysis
  (Digital signalbehandling og bildeanalyse)
- BMI: Biomedical Informatics
  (Biomedisinsk informatikk)

BMI is located on floor 4

- Research Group for Biomedical Informatics (BMI)
- Forskningsgruppen for biomedisinsk informatikk (BMI)

- Located just south of the reception on floor 4.



Staff



Knut
Liestøl



Ole Christian
Lingjærde



Geir Kjetil
Sandve



Torbjørn
Rognes



Nils D.
Christophersen



Eivind
Hovig



Karin
Lagesen



Jon K.
Lærdahl



Lex
Nederbragt



Rolf
Skotheim

Staff



Ståle Nygård



Sveinung Gundersen



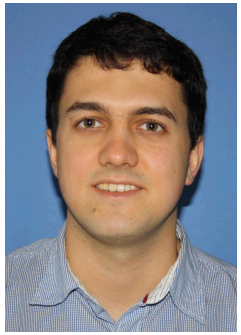
Stefania Salvatore



Hildur Sif Thorarensen



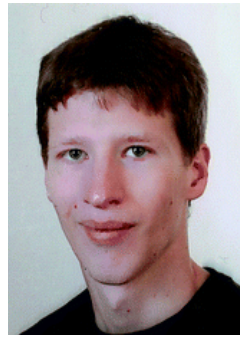
Sumana Kalyanasundaram



Oleg Agafonov



Tuva K. Thoresen



Florian Krull



Diana Domanska



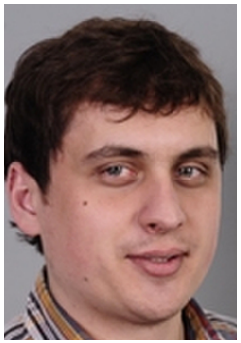
Chakravarthi Kanduri



Ksenia Khelik



Chloé Beate Steen



Boris Simovski



Merete M. Worren



Ghislain Fournous



Ivar Grytten

Associated staff at Simula

Simula Research Laboratory:

- Aslak Tveito
- Xing Cai
- Joakim Sundnes
- Are Magnus Bruaset



simula

Major research collaborations / centres

- Oslo University Hospital
- Big Insight (SFI)
- BigMed (ICT Lighthouse)
- Centre for Cancer Biomedicine (CCB) (SFF)
- Centre for Computational Inference in Evolutionary Life Science (CELS)
- Centre for Immune Regulation (CIR)
- KG Jebsen Centre for Breast Cancer Research
- KG Jebsen Coeliac Disease Research Centre

What is bioinformatics?

An interdisciplinary subject:

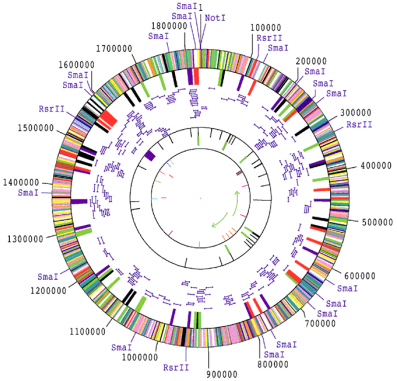
Computer science/maths/statistics + biology/medicine

Definition:

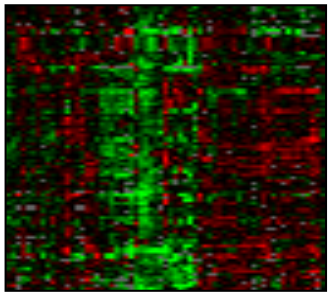
Bioinformatics is the development and use of computational and mathematical methods to gather, process and interpret molecular biological data.

Aim of research:

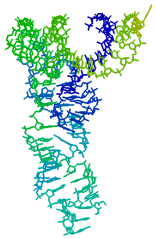
To increase our understanding of biological mechanisms while developing better theories and methods in computer science and statistics.

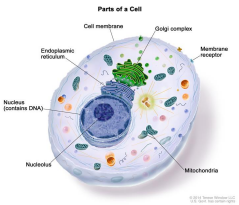






Bioinformatics at many levels








DNA	RNA	Protein	Cell	Organ	Individual	Population	Biosphere
Genomics	Transcript-omics	Proteomics	System biology	Neuro-informatics	Precision medicine	Population genetics	Metagenomics
Genome assembly	RNomics	Structural biology	Cell simulation	Organ modelling/ simulation	Variant detection	Epidemiology	Evolutionary biology
Genefinding	Microarrays	Drug design	Metabolism studies	Meta-genomics			Phylogenomics
Annotation	RNA-folding	MS analysis		Cancer genomics			
ChIP-Seq	RNA-seq	Binding site analysis					
	Structural biology	Interaction networks					

Genomes and DNA



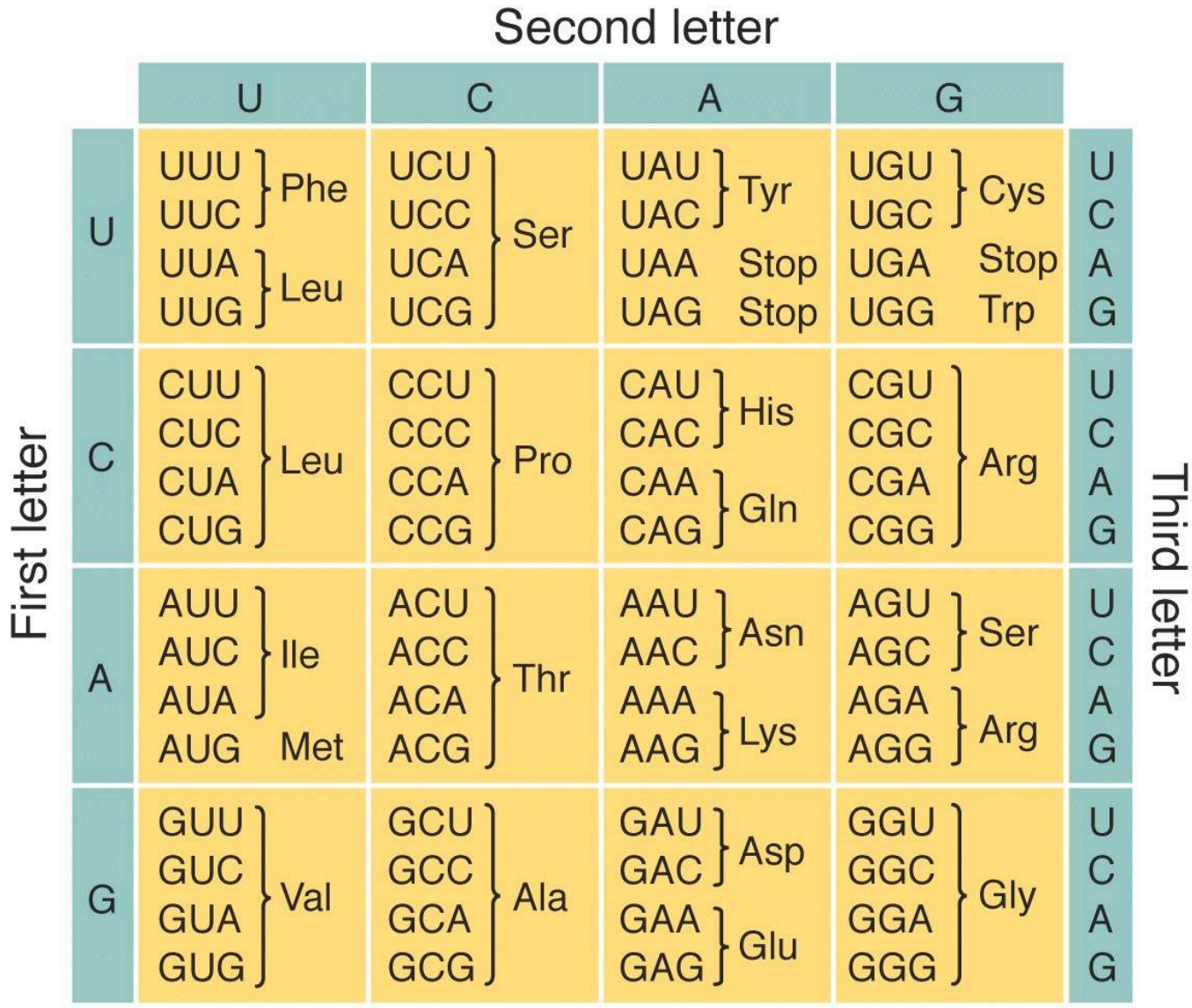
Human genome with 23 pairs of chromosomes (22 + XX or XY)

ca 3 000 000 000 bp

The universal genetic code

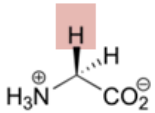
Start codon:
AUG

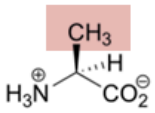
Stop codons:
UAA
UAG
UGA

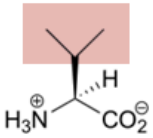


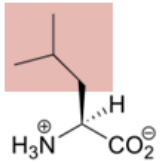
20 amino acids

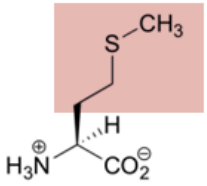
Nonpolar, aliphatic side groups

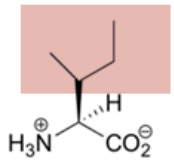

Glycine
Gly, G


Alanine
Ala, A

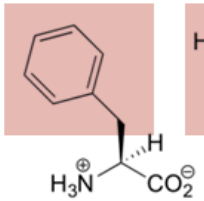

Valine
Val, V

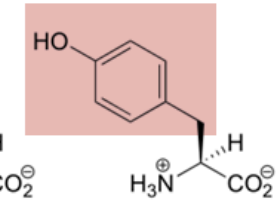

Leucine
Leu, L

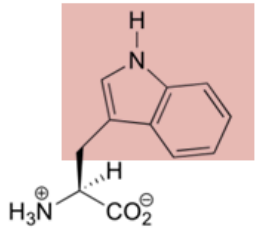

Methionine
Met, M


Isoleucine
Ile, I

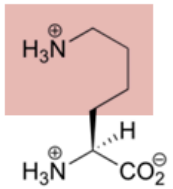
Aromatic side groups

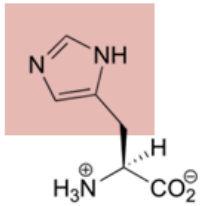

Phenylalanine
Phe, F

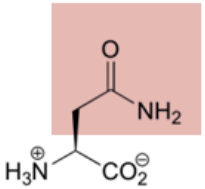

Tyrosine
Tyr, Y


Tryptophan
Trp, W

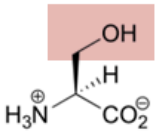
Positively charged side groups

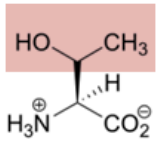

Lysine
Lys, K

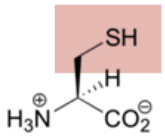

Histidine
His, H

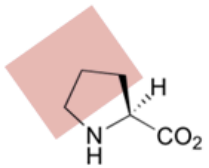

Asparagine
Asn, N

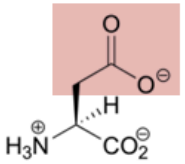
Polar, uncharged side groups

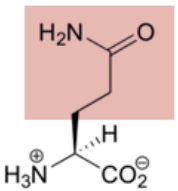

Serine
Ser, S


Threonine
Thr, T

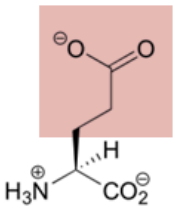

Cysteine
Cys, C

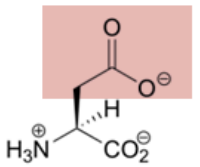

Proline
Pro, P


Aspartate
Asp, D


Glutamine
Gln, Q

Negatively charged side groups


Glutamate
Glu, E


Aspartate
Asp, D

Protein 3D structure and design

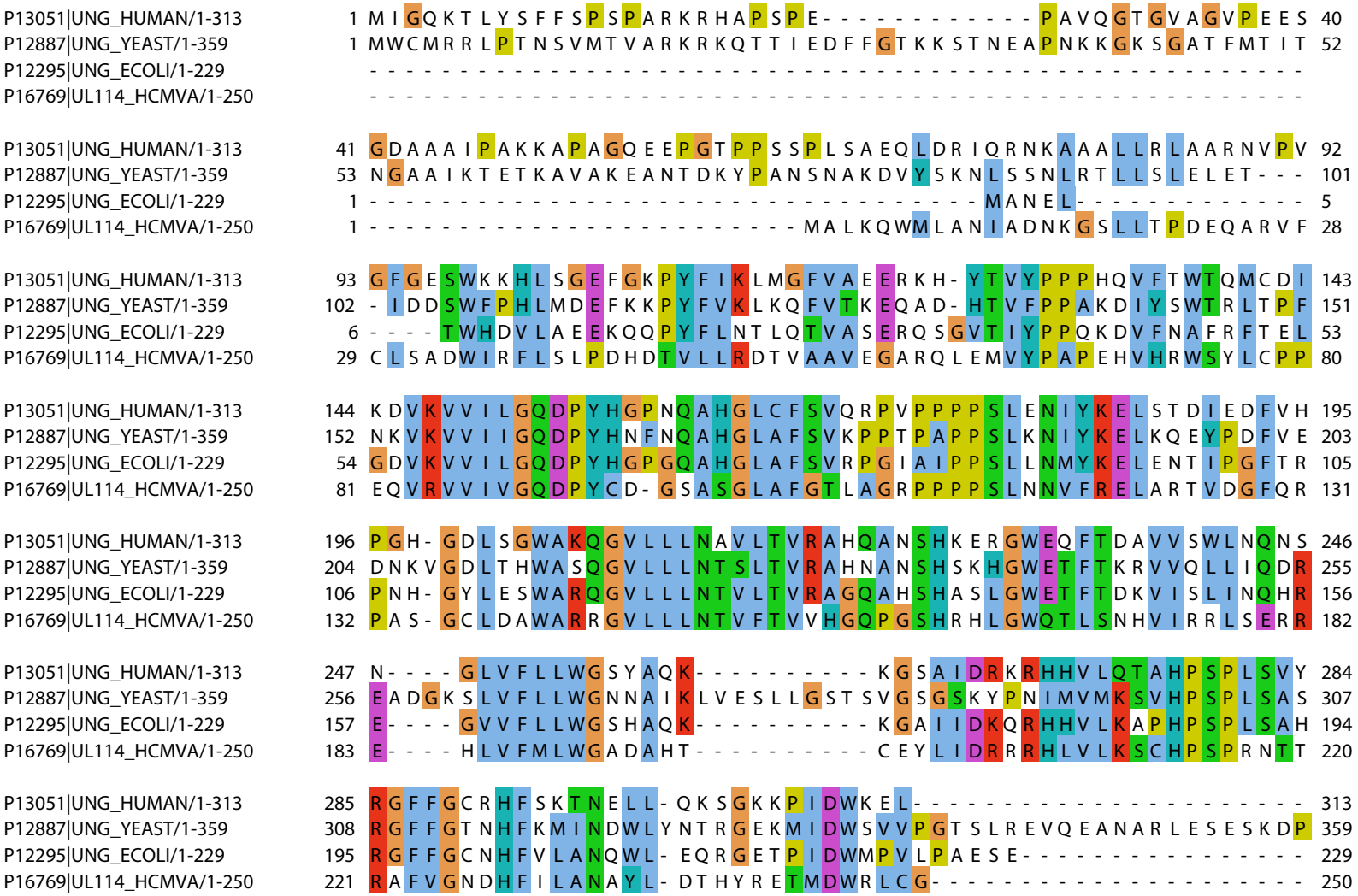
MPARALLPRRMGHRT
LASTPALWASIPCPR
SELRLDLVLPSGQSF
RWREQSPAHWSGVLA
DQVWTLTQTEEQLHC
TVYRGDKSQASRPTP
DELEAVRKYFQLDVT
LAQLYHHWGSVD...

Structure prediction →

← Protein design



Multiple sequence alignments



Genome sequencing

High-Throughput Sequencing (HTS), Deep sequencing, Next Generation Sequencing (NGS)



Obtain sample



DNA extraction, preparation



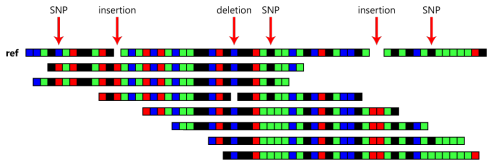
Sequencing



Raw data

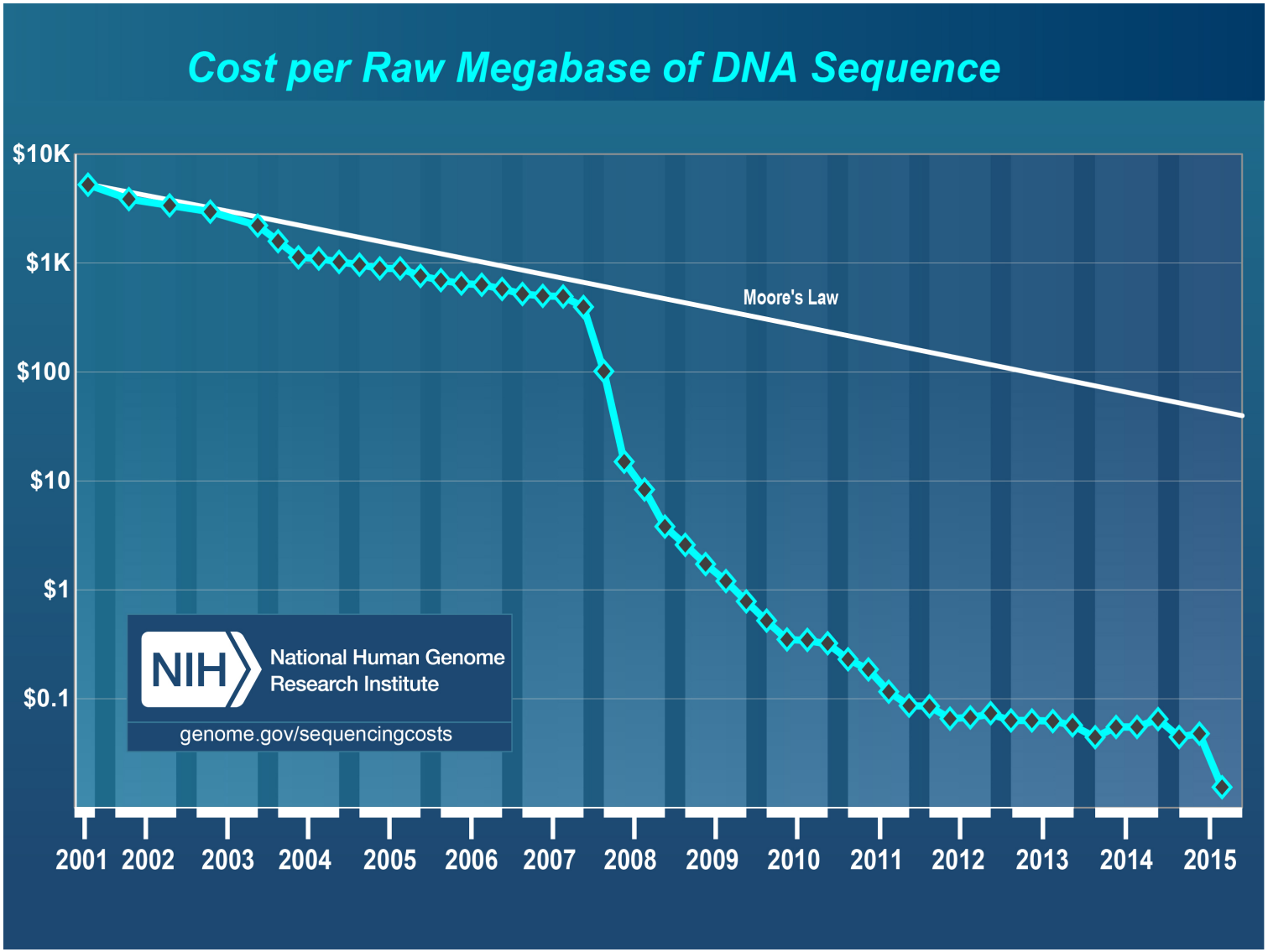


Sequence analysis

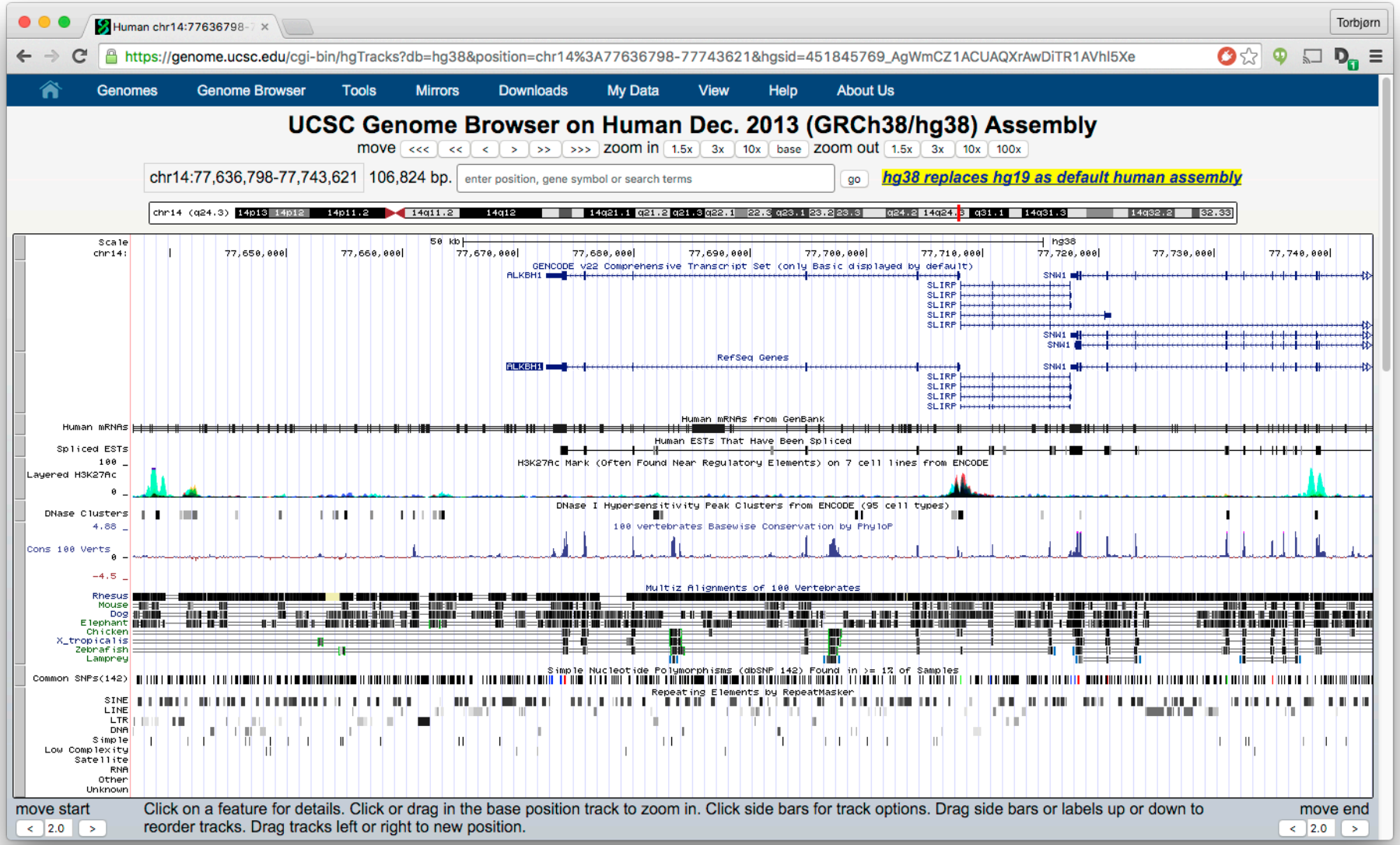


Interpretation

The cost of sequencing

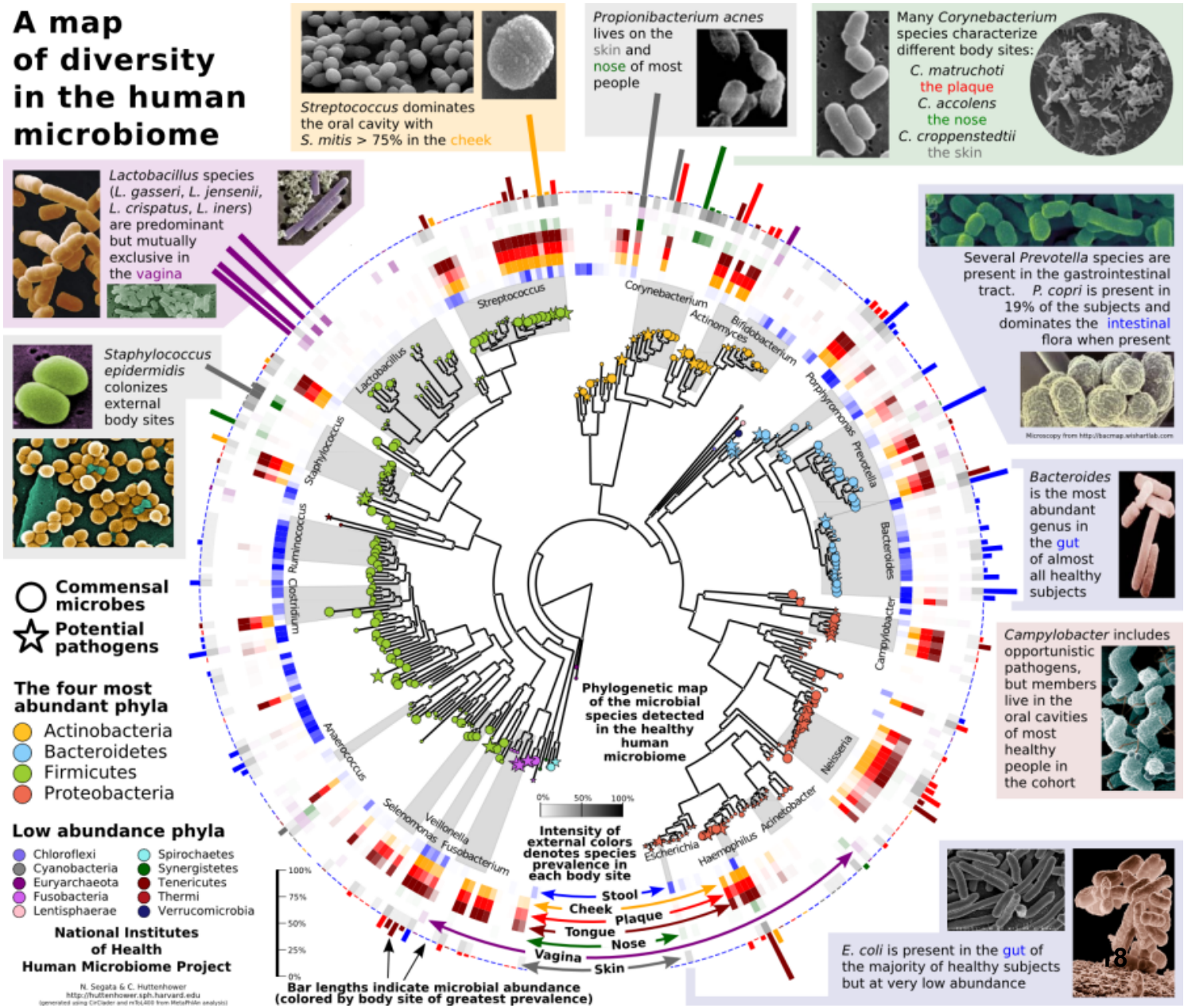


Genome browsers



Kilde: genome.ucsc.edu

Metagenomics



More info about the group

- About 20 master students currently associated with the group
- 9 master students finished in 2016

- Many (open source) computational tools developed by master students and staff

- Further details on our home page:

  http://www.mn.uio.no/ifi/english/research/groups/bmi/

Master theses submitted in 2016

- *Aligning reads against a graph based reference genome*
- *Graph Classification via Neural Networks*
- *Applying client-side state management on server-generated web pages*
- *Visualizing Chimeric RNA*
- *Computational analyses of transcriptome instability in cancer*
- *StrainTracerDB: Analysis and storage of whole genome bacterial data*
- *Tools for cluster analysis of data from genome-wide association studies*
- *Ranklust: An extension of the Cytoscape clusterMaker2 plugin and its application to prioritize network biomarkers in prostate cancer*
- *Bioinformatics approaches to deconvolute phenotype-specific immune cells from complex tissue*

Teaching

Courses taught by main BMI staff

- **IN1000** - Introduction to object-oriented programming (10c) (autumn)
- **IN1900** - Introduction to Programming with Scientific Applications (10c) (autumn)
- **INF4350** - Introductory Course in Bioinformatics (10c) (spring)
- **INF5380/9380** - High Performance Computing in Bioinformatics (5c) (spring, even years)
- **INF-BIO5121/9121** - High Throughput Sequencing technologies and bioinformatics analysis (10c) (autumn)
- **MBV-INF4410/9410** - Bioinformatics for Molecular Biology (10c) (autumn)

Courses taught by Simula staff:

- **INF3331/4331** - Problem solving with high level languages (10c) (autumn)
- **INF3380** - Parallell programming for scientific problems (10c) (spring)
- INF5063
- **INF5560** - Computational Physiology (10c) (summer)
- **INF5620/9620** - Numerical methods for partial differential equations (10c) (autumn)
- **INF5631/9631** - Numerical methods for partial differential equations; project assignment (5c) (spring and autumn)

Recommended courses

INF3380 - Parallell programming for scientific problems

INF4331 - Problem solving with high level languages

INF4350 - Introductory Course in Bioinformatics

INF4490 - Biologically Inspired Computing

INF4800 - Søketeknologi

INF4820 - Algorithms for artificial intelligence and natural language processing

INF5063 - Programming heterogeneous multi-core architectures

INF5380 - High Performance Computing in Bioinformatics

INF5750 - Open Source Development

INF5860 - Machine Learning for Image Analysis

INF-BIO5121 - High Throughput Sequencing technologies and bioinformatics analysis

STK4030 - Statistical Learning: Advanced Regression and Classification

UNIK4590 - Pattern Recognition


Thanks!