

Bioinformatiikka tieteenalana ja ammattina



Sampsa Hautaniemi, TkT
Laskennallisen systeemibiologian laboratorio
Biolääketieteen laitos &
Genomibiologian tutkimusohjelma
Biomedicum Helsinki
Helsingin yliopisto

Sisällysluettelo

- ρ Miksi bioinformatiikkaa tarvitaan?
- ρ Mistä tullaan, missä ollaan ja minne ollaan menossa?
- ρ Millaista bioinformatiikan tutkimus on?
- ρ Mitä bioinformatiikan opiskelijalta ja tutkijalta odotetaan?
- ρ Mitä bioinformatiikka tieteenalana ja ammattina tarjoaa?
- ρ Minne bioinformatiikasta valmistuneet menevät töihin?

Miksi bioinformatiikkaa tarvitaan?

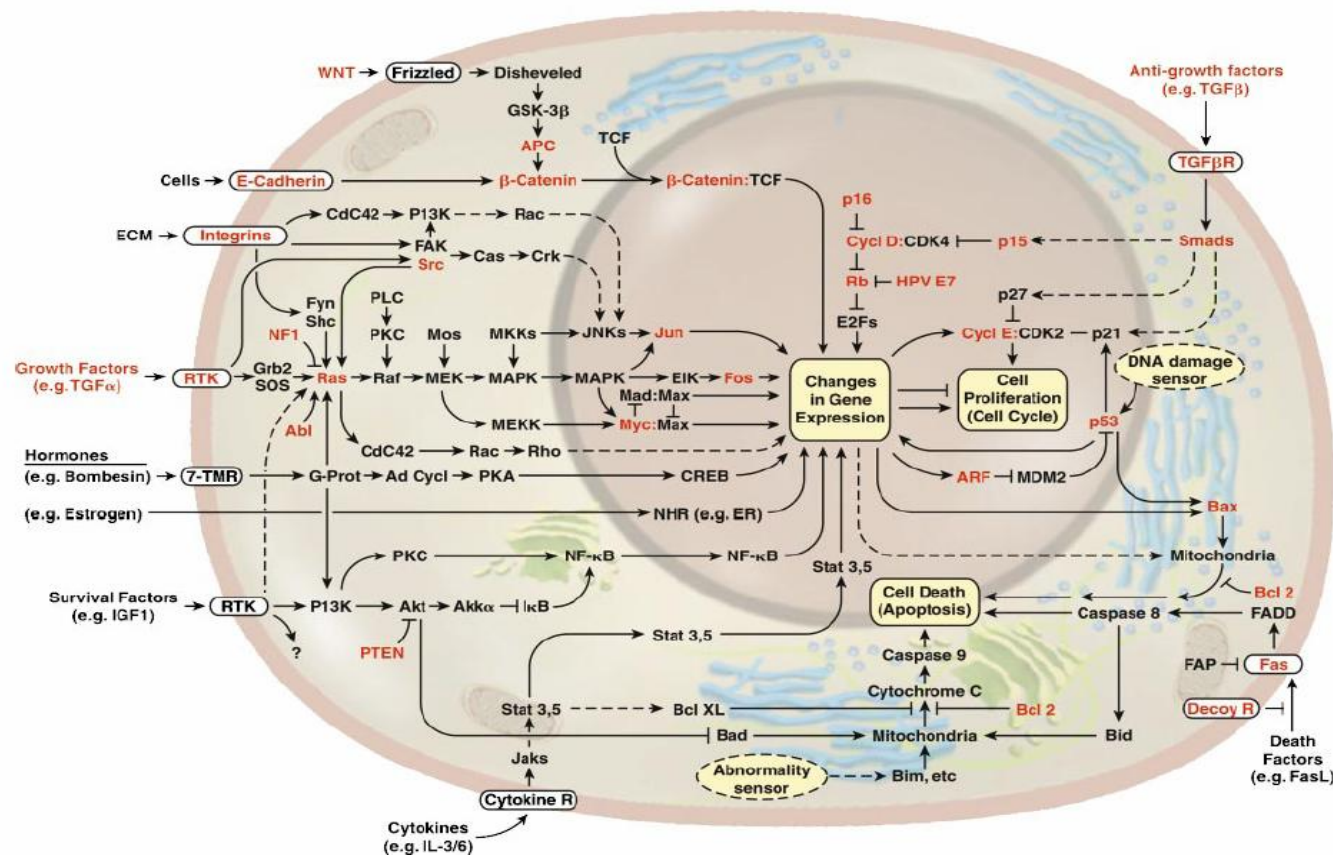


Figure 2. The Emergent Integrated Circuit of the Cell

Bioinformatiikka on olennainen osa modernia lääketieteellistä tutkimusta.

Missä ollaan?

p Palaset alkavat olla kasassa.

What is a gene, post-ENCODE? History and updated definition

Mark B. Gerstein,^{1,2,3,9} Can Bruce,^{2,4} Joel S. Rozowsky,² Deyou Zheng,² Jiang Du,³ Jan O. Korbel,^{2,5} Olof Emanuelsson,⁶ Zhengdong D. Zhang,² Sherman Weissman,⁷ and Michael Snyder^{2,8}

¹Program in Computational Biology & Bioinformatics, Yale University, New Haven, Connecticut 06511, USA; ²Molecular Biophysics & Biochemistry Department, Yale University, New Haven, Connecticut 06511, USA; ³Computer Science Department, Yale University, New Haven, Connecticut 06511, USA; ⁴Center for Medical Informatics, Yale University, New Haven, Connecticut 06511, USA; ⁵European Molecular Biology Laboratory, 69117 Heidelberg, Germany; ⁶Stockholm Bioinformatics Center, Albanova University Center, Stockholm University, SE-10691 Stockholm, Sweden; ⁷Genetics Department, Yale University, New Haven, Connecticut 06511, USA; ⁸Molecular, Cellular, & Developmental Biology Department, Yale University, New Haven, Connecticut 06511, USA

While sequencing of the human genome surprised us with how many protein-coding genes there are, it did not fundamentally change our perspective on what a gene is. In contrast, the complex patterns of dispersed regulation and pervasive transcription uncovered by the ENCODE project, together with non-genic conservation and the abundance of noncoding RNA genes, have challenged the notion of the gene. To illustrate this, we review the evolution of operational definitions of a gene over the past century—from the abstract elements of heredity of Mendel and Morgan to the present-day ORFs enumerated in the sequence databanks. We then summarize the current ENCODE findings and provide a computational metaphor for the complexity. Finally, we propose a tentative update to the definition of a gene: **A gene is a union of genomic sequences encoding a coherent set of potentially overlapping functional products.** Our definition sidesteps the complexities of regulation and transcription by removing the former altogether from the definition and arguing that final, functional gene products (rather than intermediate transcripts) should be used to group together entities associated with a single gene. It also manifests how integral the concept of biological function is in defining genes.

Mistä ollaan tultu?

- p Vuonna 1973 24:n nukleotidin sekvenointi oli Iso Juttu.
- p Geenin määritelmä on muuttunut kovasti vuodesta 1860 (diskreetti perinnöllisyyden yksikkö) vuoden 1940 (proteiinin teko-ohje) kautta nykyiseen (aliohjelma genomisessa käyttöjärjestelmässä).

Mihin ollaan menossa?

- p Sekvenssidataa tulee kohta lisää paljon ja nopeasti:
 - n X Prize Foundation: \$10 miljoonaa sille/niille jotka kehittävät laitteen jolla voi
 1. sekvenoida 100 ihmisen genomia maksimissaan kymmenessä päivässä,
 2. tarkkuudella 1 virhe/100,000 nukleotidia,
 3. käsittäen vähintään 98% genomista
 4. alle \$10,000 per genomi.
- p Sama pätee useisiin muihin datalähteisiin.
- p Dataa pitäisi pystyä käsittelemään ja yhdistelemään järkevästi, tehokkaasti ja luotettavasti.
- p Geneettinen testaus, kohdennettu hoito.

Millaista bioinformatiikan tutkimus on?

p Poikkitieteellistä.

- n Vahva menetelmätieteellinen osaaminen on olennaista.
- n Biologiasta ja lääketieteestä on hyvä olla kiinnostunut.

p Kansainvälistä.

- n Liikkuvuus on olennainen osa bioinformatiikan tutkimusta.

p Haastavaa.

- n Tutkimusongelmat ovat oikeasti monimutkaisia.

p Mielenkiintoista.

- n Projektit ovat erilaisia, tuloksilla on lääketieteellistä merkitystä.

Mitä bioinformatiikan opiskelijalta ja tutkijalta odotetaan?

- p Menestystä opinnoissa.
 - n Yleisarvosana vähintään 3.0 ja pääaineen keskiarvo mielellään lähellä 4:ä.
- p Hyvää kommunikaatiokykyä.
 - n Suurin ongelma bioinformatiikan tutkimuksessa on kuvata lääketieteellinen kysymys matemaattisesti. Tätä on hankala tehdä ilman keskusteluja biotieteilijöiden kanssa.
- p Epävarmuuden sietämistä.
 - n Useissa tutkimuskysymyksissä reunaehdot ei pystytä määrittelemään. Myös tällöin on tärkeää pystyä tekemään työtä ja tuloksia.
- p Ahkeruutta.
 - n Aidosti poikkitieteellisessä tutkimuksessa tulee syventyä omaan ja sovellusalaan. Biotieteiden opiskelu ottaa aikaa.

Mitä bioinformatiikka tieteenalana ja ammattina tarjoaa?

- p Mielenkiintoisia projekteja.
 - n Projektit voivat liittyä HIVin, kantasoluihin, luuytimen siirtoon, erilaisiin syöpiin, reumoihin, jne jne.
- p Mahdollisuuden soveltaa osaamistaan maailman huippujen kanssa.
 - n Toimiva ja innovatiivinen menetelmä otetaan nopeasti käyttöön ympäri maailmaa. Yhteistyömahdollisuuksista ei ole pulaa.
- p Oppia siitä, miten monimutkaisia ongelmia ratkotaan tehokkaasti.
 - n Biologisen datan käsittelyyn sopivat menetelmät ovat käyttökelpoisia muillakin sovellusaloilla.

Minne bioinformatiikasta valmistuneet menevät töihin?

- p Maisterit aloittavat yleensä jatko-opinnot.
 - n Maisterius on hyvä alku. Tohtorintutkinto antaa eväät käytännön bioinformatiikan toteuttamiseen.
- p Tohtorit saavat hyvin töitä bioaloihin liittyvistä yrityksistä ja tutkimuslaitoksista.
 - n Yliopistot, lääketeollisuus, bioteknologiayritykset, pankkiiriliikkeet, konsultointiyritykset, järjestelmätoimittajat.