

Geenitehnoloogia statistilised mudelid

(Statistical Models in Gene Technology)

Tanel Kaart

Tartu, 2001

Sissejuhatus

Matemaatika teaduskond
matemaatilise statistika õppekava

Aine kood: MTMS.01.067

Aine nimetus: Geenitehnoloogia statistilised mudelid
Statistical Models in Gene Technology

Õpetamise keel: eesti keel

Aine maht: 2 ap (32 tundi loenguid, 48 tundi iseseisvat tööd)

Õpetamise aeg: 2000/2001 õppeaasta kevadsemester

Hindamine: arvestus

Vastutav õppejõud: Tanel Kaart

Kuulajaskond: matemaatilise statistika eriala magistriõppe üliõpilased

Eeldused (soovituslikud):

Populatsioonigenetika matemaatilised alused (MTMS.02.007)

Biomeetria matemaatikutele (MTMS.01.031)

Mittetasakaaluline dispersioonanalüüs (MTMS.01.059)

Kirjandus:

M. Lynch, B. Walsh. 1998. Genetics and Analysis of Quantitative Traits.
Sinauer Associates, Inc.

(<http://nitro.biosci.arizona.edu/zbook/book.html>)

P. Sham. 1998. Statistics in human genetics. Arnold, London.

J. Ott. 1999. Analysis of human genetic linkage, 3rd edition. Johns
Hopkins University Press, Baltimore.

B. Kinghorn & J. van der Werf. QTL Course: June 2000 Belo Horizonte,
Brasil: Identifying and incorporating genetic markers and major genes in
animal breeding programs –

<http://www-personal.une.edu.au/~jvanderw/brazilcourse.html>

Annotatsioon:

Kursus annab ülevaate matemaatilise statistika meetoditest, millel
baseerub tänapäevane geenitehnoloogia. Põhipunktid: geenide
kaardistamine ja üksikute geenide mõjude hindamine, segregatsioon- ja
aheldusanalüüs, assotsiatsiooniuringud.

Loengute temaatiline kava:

Õppenädal	Teema
1.-2.	Sissejuhatus; mõisteid, fakte ja seadusi geneetikast
3.-4.	Segregatsioonanalüüs
5.	Geenide aheldus, ristsiirde sagedus, geneetiline kaart
6.-9.	Parameetriline aheldusanalüüs
10.-11.	Mitteparameetriline aheldusanalüüs
12.-14.	Assotsiatsiooniuringud
15.	Kokkuvõte + üldistused + pilk tulevikku
16.	Varus

Sissejuhatus

Kursus on mõeldud tutvustamaks statistikutele tänapäeva geneetikas esilekerkinud probleeme ja nende lahendamiseks kasutatavaid statistilise analüüsi meetodeid. Valdakonna ulatuslikkuse ja kursuse piiratud ajaressursside tõttu jääb esitus osaliselt vaid tutvustuse tasemele (erinevad meetodid esitatakse-tuletatakse mingi kindla juhu jaoks ja illustreeritakse näidetega). Loodetavalt õnnestub siiski tekitada kuulajates arusaam, et ega see “geenijahtimine” paljalt laboris toimu – ka statistika meetoditega on siin juba palju ära tehtud (ja palju on veel ka tegemata). Loodetavalt annab järgneva materjaliga tutvumine nii mõnelegi noorele matemaatiku haridusega inimesele julguse ja huvi geenide uurimisega lähemalt tutvust teha. Ka bioloogia, arsti või loomakasvatuse eriala kraadiõppureile ei tohiks järgnev ülejõu käiv olla (vajadusel tuleks omale selgeks teha põhilised statistika ja tõenäosusteooria seaduspärad).

Paralleelselt viimastel aastatel enimkõneldud inimgeneetikale saavad käsitletud ka loomagenetikas kasutatavad statistika meetodid. Ja ega seal tegelikult eriti vahet polegi, kui üldse, siis vaid selles, mida eksponeeritakse. Inimese puhul räägitakse tavaliselt mingit haigust põhjustavast geenist (leidmise korral saavad ravimifirmad töötada välja uusi arstirohte, võimalikuks saab geeniteraapia), loomade puhul aga majanduslikult olulist tunnust märkimisväärselt mõjutavast geenist (tunduvalt efektiivsemaks muutub loomade aretus). Samas on loomadegi puhul leitud mitmeid haigusttekitavaid geenimutatsioone, geeniteraapia asemel aga selliseid loomi järglas põlvkonnas saamisel lihtsalt ei kasutata või siis välditakse vanemate teadliku valiku abil retsessiivsete homosügootide saamist (näiteks stressi põhjustava HAL-geeni korral sigade aretuses). Ning enamus lugejaist on kindlasti kuulnud eugeenikast, kui millestki vägagi taunitavast, ja eks ta seda olegi, kuigi oma olemuselt ei ole eugeenika ju midagi muud kui aretusteooria rakendamine inimese peal. Ja kokkuvõtteks ei sõltu matemaatika sellest, kas teda rakendatakse inimesel või loomal või milleks saadud tulemusi üldse kasutatakse, otsustama mida, kuidas ja kas üldse uurida peab ikkagi inimene.

Kas võib seda kursust nimetada ka bioinformaatikakursuseks? Arvan, et mitte, sest bioinformaatika on selleks, et sellest kasvõi osagi taolise kursusega katta, liiga laialdane mõiste – bioinformaatikaga seondub igasugune bioloogilise informatsiooni töötlemisega seotud arvutialane töö. Siia mahuvad tööd geenidest ja nende produktidest kuni ökosüsteemideni. Siiski on, tulenevalt suurtest andmehulkadest, mida molekulaarbioloogia toodab, enamus bioinformaatika projekte keskendunud geenide ja valkude struktuursete ja funktsionaalsete aspektide uurimisele. Paljud neist projektidest on seotud inimese genoomi projektiga (vt. ka kursuse materjalile lisatud linke). Traditsioonilise mõtteviisi kohaselt kuuluvad kursusel käsitletavat meetodid pigem biostatistika alla. Samas, arvestades, et biostatistika erineb bioinformaatikast sama palju kui informaatika statistikast – nn. klassikalise statistika traditsiooniliseks rakendamiseks pole

arvutiteadusealaseid teadmisi eriti vaja, seevastu uusimate analüüsimeetodite osas on erinevus tihti vaid uuringuid teostavate institutsioonide nimetustes, võib järgnevat materjali vaadelda ka kui bioinformaatika ühe alamsuuna detailsemat tutvustust (viimastel aastatel mitmeis ülikoolides alustatud bioinformaatika magistriõppe kavades ongi käesolevale samased kursused kirjas valikainetena).

Tanel Kaart
jaan. 2001