

Hiiriä, hiivoja ja kärpäsiä – mitä malliorganismien geenit kertovat elämästä ja sen evoluutiosta

Hannu Sariola, Irma Thesleff ja Marja Makarow

Malliorganismeiksi kutsutaan lajeja, joita tutkijat käyttävät biologisessa tai lääketieteellisessä tutkimuksessa ja joista on sen takia kertynyt enemmän biologista ja fysiologista tietoa kuin muista lajeista. Nykyään eniten käytetty mallieläin on hiiri. Yksinkertaisemmillä lajeilla, kuten leivinhiivalla, sukkulamadolla, seeprakalalla ja banaanikärpäsellä, on kuitenkin suuri merkitys tutkijoille, koska lyhyen sukupolvikierron takia niistä saadaan tietoa nopeammin kuin hiirestä. Banaanikärpäsen, sukkulamadon ja leivinhiivan genomit on sekvensoitu kokonaan, ja hiiren genomien sekvenssi on valmistumassa. Eri lajien perimän systemaattinen kartoitus tuottaa uutta tietoa biologisista lainalaisuuksista ja elämän geneettisen säätelyn perustasta. Miten geenit ja niiden perheet ovat syntyneet evoluutiossa? Mitkä geenit ovat elämälle välttämättömiä? Miten monta geeniä tarvitsee yksinkertaisin itsenäisesti elävä eliö?

Sveitsiläinen Friedrich Miescher löysi vuonna 1869 Reinjoen lohesta DNA:n (hän kutsui deoksiribonukleiinihappoa nukleiiniksi) (Lewin 2000). Löydöksen merkitystä ei kuitenkaan ymmärretty yli 70 vuoteen. Vasta vuonna 1944 yhdysvaltalaisen Oswald Aweryn ryhmä osoitti, että perimäaines on DNA:ta eivätkä ominaisuudet periydy valkuaisessa. Vajaat kymmenen vuotta myöhemmin, vuonna 1953, James Watson ja Francis Crick esittivät mallin DNA-kaksoiskierteen rakenteesta (taulukko 1). Vaikka geenitutkimuksen peruspalaset olivat koossa, näiden havaintojen jälkeen kului melkein puoli vuosisataa, ennen kuin ensimmäisten lajien genomit oli kartoitettu, geenit tunnistettu ja niiden emäsjärjestys määritetty.

Genomihankkeiden valmistumisen merkityksestä ihmiskunnalle ei liene epäilystä, mutta on mielenkiintoista tarkastella sitä logiikkaa, jolla

Taulukko 1. Geenitutkimuksen lyhyt historia (Lewin 2000).

1869	Nukleiinihappo löydetään
1903	Perimä sijaitsee kromosomeissa
1910	Geenit sijaitsevat kromosomeissa
1913	Geenit ovat tiettyssä järjestyksessä
1927	Mutaatiot johtuvat geenien muutoksista
1931	Rekombinaatio johtuu tekijävaihdosta
1944	Geenit muodostuvat DNA:sta
1945	Geenit koodaavat proteiineja
1951	Ensimmäinen proteiinin aminohappojärjestys
1953	DNA-kaksoiskierre
1961	Geneettinen koodi muodostuu nukleiinihappotripletistä
1977	Eukaryoteissa geenit ovat rakenteeltaan epäjatkuvia
1977	DNA:ta voidaan sekvensoida
1995	Ensimmäisen lajin genomi sekvensoidaan kokonaisuudessaan

eri lajeja on valittu genomihankkeisiin. Ensiksi sekvensoitiin tietysti yksinkertaisia lajeja, kuten viruksia ja bakteereita. Ainoastaan niiden pienet genomit oli mahdollista kartoittaa alkuvaiheen rajatuilla resursseilla ja alkeellisella tekniikalla.

Usein taustalta löytyi kuitenkin myös lääketieteellinen motivaatio ja sekvensoitavaksi valittiin keskeisiä patogeneja. DNA-tekniikan kehittyessä kohteeksi otettiin biologisten malliorganismien ja ihmisen kookkaammat genomit. Hiivan geneettinen koodi saatiin purettua vuonna 1996 ja banaanikärpäsen vuonna 2000. Ihmisen genomikartta on valmis; geenien pätkät on tallennettu tietopankkeihin vielä suurelta osin tuntemattomina pätkinä, ns. EST-klooneina, joten kulnee vielä joitain vuosia, ennen kuin tämä valtava informaatiokaos saadaan järjestettyä loogiseksi ja mahdollisimman virheetömäksi tiedoksi kaikkien geenien emäsjärjestyksestä.

Hiiret, hiivat, kärpäset ja madot

Biologisten malliorganismien valinta geenikartoituksen ensisijaisiksi kohteiksi on perusteltua. Näistä lajeista on kertynyt runsaasti geneettistä ja muuta biologista tietoa jo ennen genomihankkeita, joten niiden sekvenssitietoa voidaan tulevaisuudessa hyödyntää tehokkaasti erilaisten biologisten kysymysten selvittämiseen. Malliorganismeja ovat mm. leivinihiiva (*Saccharomyces cerevisiae*), sukkulamato (*Caenorhabditis elegans*), lituruoho (*Arabidopsis thaliana*), tupakka, *Xenopus laevis* -sammakko, banaanikärpänen, hiiri ja uusimpana seeprakala (kuva 1). Ne ovat joutuneet tutkijoiden mielenkiinnon kohteeksi eri syistä, eikä lajin kehittyminen malliorganismiksi ole juuri perustunut rationaaliin ajatteluun eikä biologisiin tai lääketieteellisiin syihin. Pikemmin taustalta löytyy sattuma tai yksittäisen tutkijan into tai käytännöllisiä syitä, kuten lajin pieni koko, nopea lisääntyminen, suuret poikueet, halpuus ja kasvattamisen helppous laboratoriossa. Tältä pohjalta on myös ymmärrettävissä, että muutamat keskeiset lajit ovat »unohtuneet» tutkijoilta. Vaikka ihminen on jalostanut koiria eri tarkoituksiin tuhansia vuosia, koirasta ei ole tullut tutkijoiden malliorganismia, ja sen seurauksena koiran genetiikka on vielä lapsenkengissä. Samaan aikaan kun kärpäsestä tunnetaan 13 600 geeniä, koirasta niitä tunnetaan vasta 200.

Banaanikärpäsen ominaisuuksien periytymistä on tutkittu lähes sata vuotta. Usein banaanikärpästutkimuksen aluksi mainitaan vuosi 1910, kun Thomas Hunt Morgan löysi banaanikärpäsviljelmästä ensimmäisen mutantin, valkosilmäisen kärpäsen, ja ryhtyi tutkimaan banaanikärpäsen perimää mutaatioiden avulla (Rubin ja Lewis 2000). Vaikka perimän kemiallista luonnetta ei tuolloin vielä ymmärretty, Morgan onnistui mutaatioita tutkimalla muotoilemaan vallankumouksellisen teorian kromosomeista perimän sisältäjinä, mistä hänelle myönnettiin Nobelin palkinto vuonna 1933.

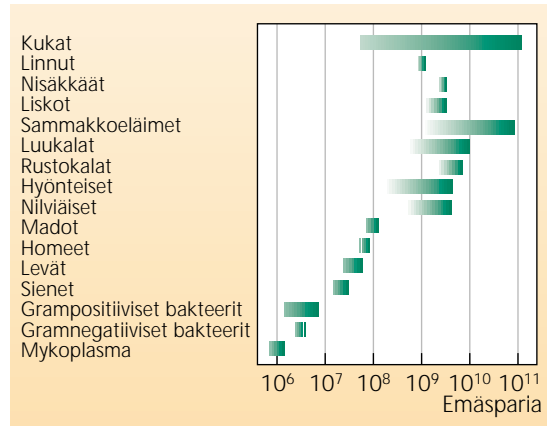
Banaanikärpäsen genomien erityisominaisuus on ns. jätti- eli polyteenikromosomisto, jossa geenien paikka on nähtävissä yksinkertaisesti mikroskoipimalla. Polyteenikromosomisto syntyy, kun kromosomit kahdentuvat sivusuunnassa mutta eivät irtoakaan toisistaan vaan muodostavat mikroskoopissa näkyvän rakenteen, jossa geenien paikat voidaan tunnistaa poikkijuovien avulla. 1930-luvun tekniikalla kuvatut sylkirauhasen polyteenikromosomiston juovakartat ovat osoittautuneet niin tarkoin, että geenit voidaan paikantaa niissä parhaimmillaan 100 000 emäsparin tarkkuudella.

Seeprakala ja sukkulamato ovat niin ikään muodostuneet suosituiksi perinnöllisyyden tutkimuskohteiksi, koska niissä on helppo aikaansaada mutaatioita. *C. elegans* -sukkulamato on ollut keskeinen mm. ohjelmoituneen solukuoleman eli apoptoosin ymmärtämisessä. Seeprakalasta on selkärankaisena lajina tulossa tärkeä ihmisen tautimalli (The *C. elegans* sequencing consortium 1998, Hengartner 1999, Plowman ym. 1999). Vähän aikaa sitten käynnistyneen koko genomien kattavan mutaatioseulonnan avulla yritetään löytää kaikkien seeprakalan geenien toiminnot. Koska seeprakalat ovat läpinäkyviä, sisäelintenkin mutaatiot ovat helposti tunnistettavissa pelkän stereomikroskoopin avulla.

Hiiri. Vaikka hiiren valta-asema kokeellisena eläinmallina näyttää vankkumattomalta, se ei ole ollut tässä asemassa kovinkaan pitkään. Keskeinen ongelma on tietysti ollut hiiren geneettisen muokkaamisen vaikeus, joka poistui oikeastaan vasta geenien siirto- ja poistotekniikan myötä 1980-luvun alussa. Siirtogeenitekniikan nopea yleistyminen ja kehittyminen viime vuo-

sikymmenen aikana on suurin syy siihen, että hiiri on nyt ehdottomasti tärkein malliorganismi ihmisen fysiologian ja patologian tutkijoille. Tämän vuoden alussa arvioitiin, että maailmassa tutkitaan vuodessa 25 miljoonaa hiirtä ja kaikista eläinkokeista 90 % tehdään hiirillä (Malkoff 2000). Pitkälle kehitetyillä tekniikoilla voidaan paitsi vaihtaa hiiren oma geeni ihmisen tautigeeniksi myös kohdistaa geenien mutaatioita haluttuihin kudoksiin. Toisaalta hiiren spontaaneja mutaatioita on kuvattu useiden vuosikymmenten ajan, ja niitä tunnetaan yli tuhat. Näissä esiintyvä geenivirhe on viime vuosina selvitetty noin 150 tapauksessa, ja puolet on osoittautunut ihmisen tautimalleiksi. Menossa on myös laaja mutageneesitutkimus, jonka on luvattu tuottavan tutkijoiden käyttöön tuhansia uusia hiirimutantteja jo vuonna 2001. Hiiren genomien sekvensointi on meneillään monikanalisena projektina, ja sen on arvioitu valmistuvan viimeistään vuonna 2005. Voi kuitenkin käydä niin, että yksityinen Celera Genomics julkaisee ainakin osittaisen sekvenssin aikaisemmin, kuten on käymässä ihmisen genomien suhteen (Weinberg 1999, Peltonen-Palotie 2000, Service 2000).

Leivinhiiwa. Vaikka ihminen kehittyi miljardi vuotta hiivan jälkeen, on kolmasosalla hiivan geneeistä vastine ihmisen perimässä (Makarow 1996). Niinpä hiiwa onkin erinomainen malli myös nisäkässolun perustoimintojen tutkimuksessa (Goffeau ym. 1996, Botstein ym. 1997). Minkä tahansa hiivan geenin poisto on tehtävissä muutamassa viikossa tai kuukaudessa. Poistogeenisen hiiren tuottaminen on paljon pitempi, vaikeampi ja kalliimpi urakka. Hiivan geeni voidaan korvata ihmisen vastaavalla geenillä ja tutkia normaalin ihmisproteiinin toimintaa tai tautigeenin aiheuttamaa patologiaa solu- ja molekyyllitasolla. Hiivamallissa on selvitetty esimerkiksi ataksia-telangiektasian (AT) molekulaarinen mekanismi. Taudin oireet, katkokävely, laajentuneet verisuonet, puutteet immuunivasteessa ja satakertainen leukemian ja lymfooman riski johtuvat siitä, että soluihin jatkuvasti ilmaantuvat DNA:n vauriot eivät ehdi korjautua ennen kuin solunjakautuminen vie vauriot tyrsoluihin. AT-tautigeenin hiivavastineet huo-



Kuva 1. Genomin DNA-määrä korreloi lajin kehitystasoon ainoastaan alkeellisissa lajeissa.

lehtivat solusyklin pysäytyksestä DNA-vaurioiden korjaamisen ajaksi, jolloin mutaatiot eivät siirry seuraavaan solusukupolveen (Morrow ym. 1995). Jopa prionitautien mekanisme voidaan tutkia leivinhiiwassa. Hiiwalla on kaksi omaa proteiinia, jotka tietyissä olosuhteissa omaksuvat nisäkkäitten prioniproteiinille ominaisen saostuneen muodon, joka periytyy jälkeläisille ilman geneettisiä muutoksia (Patino ym. 1996).

Sammakko. Biologit käyttivät vuosisadan alkupuolella kokeellisena mallina runsaasti salamantaria ja sammakkoa. Salamantarella on nykyään ainoastaan marginaalinen merkitys kokeellisena mallina. *Xenopus laevis* -sammakko on tulossa uudelleen kehitysbiologien tutkimuskohteeksi ennen kaikkea geeninsiirtotekniikan parantumisen ansiosta. Sammakon avulla onkin viime vuosina selvitetty lukuisten geenien kehityksellisiä tehtäviä (Brändli 1999).

Banaanikärpäsen geenit

Banaanikärpäsen on tähän mennessä monimutkaisin laji, jonka genomi on käytännössä kokonaan sekvensoitu. Science-lehti juhlisti teemanumerolla *Drosophila*-genomihankkeen valmistumista maaliskuussa 2000. Hanke oli kansainvälinen yhteisprojekti, johon osallistui yhdysvaltalaisia ja eurooppalaisia tutkimuslaboratorioita sekä genomiprojekteihin erikoistunut Celera Genomics (Adams ym. 2000). Valtaosa geenien

sekvenssitiedosta tuotettiin vuoden 1999 aikana! Ratkaisevaa hankkeen nopealle onnistumiselle oli genomien sekvenssointiin valittu ns. haulikkomenetelmä (Myers ym. 2000). Siinä genomi pilkotaan ensin sattumanvaraisesti 2 000–10 000 emäsparin pätkiin, jotka siirretään vektoreihin sekvenssointia varten. DNA-pätkien emäsjärjestys luetaan automaattisilla laitteilla, minkä jälkeen tämä täysin satunnaisista pätkistä kerätty sekvenssitieto tallennetaan tietokoneen muistiin. Koska eri pätkien sisältämä tieto on osittain päällekkäistä, tietokone pystyy organisoimaan geenitiedon vähitellen oikeaan järjestykseen liittämällä pätkien päällekkäin meneviä jaksoja virtuaalisesti toisiinsa.

Banaanikärpäsän genomia ei sekvenssoitu täydellisesti, vaan hanke keskittyi ns. eukromatiiniseen alueeseen, joka sisältää lähes kaikki geenit (Myers ym. 2000). Nyt selvitettiin emäsjärjestys 120 000 000 emäsparin alueelta. Tämän alueen ulkopuolelle jää noin 30 % genomista (heterokromatiinialue). Sitä ei tutkittu, koska sieltä ei ole löytynyt kuin yksi tai kaksi geeniä. Vaikka päätyö on tehty, hankkeeseen osallistuvat laboratoriot jatkavat edelleen yhteistyötä paikataksaan genomissa vielä olevia aukkoja ja korjataksaan virheitä. Geenitiedon järjestämisen vaikeutta kuvattiin Science-lehdessä seuraavasti: »Kuvittele rakentavasi auton moottoria osista, joista sinulla ei ole muuta tietoa kuin satunnainen numerokoodi eikä mitään käsitystä niistä tai rakenteesta.»

Banaanikärpäsän genomi oli täynnä yllätyksiä. Sadan vuoden tutkimuksen tuloksena kärpäsestä oli löydetty 2 500 geeniä. Kun koko genomi kartoitettiin, geenejä löytyi yhteensä noin 13 600. Niistä puolet oli tunnistettavissa nisäkkäiden geenien sukulaisiksi ja osa oli lähes identtisiä ihmisen geenien kanssa. Vastine löytyi mm. p53-kasvunestäjä-, tau- ja MEN1-tautigeeneille (Kornberg ja Krasnow 2000).

Banaanikärpäsessä on ainoastaan kaksi kertaa enemmän geenejä kuin hiivassa. Kun banaanikärpäsän geenejä verrattiin leivinihiivan ja *C. elegansin* geeneihin, löydettiin mielenkiintoisia samankaltaisuuksia ja eroja (O'Brian ym. 1999, Rubin ym. 2000). Noin 20 % proteiineista on samankaltaisia kaikissa näissä lajeissa, jo-

ten näiden geenien proteiinituotteet ovat todennäköisesti samoja muissakin eukaryoottisissa lajeissa ja muodostavat solun elämän perustarpeet. Lajien yhteiset geenit koodittavat proteiineja, jotka huolehtivat aineenvaihdunnasta, DNA- ja RNA-synteesistä sekä proteiinien valmistamisesta, laskostamisesta, kohdentamisesta ja hajotuksesta.

Merkittävä ero banaanikärpäsän ja leivinihiivan geenien välillä oli se, että banaanikärpäsän samoin kuin *C. elegans* -madon geeneistä suhteellisen suuri osa koodaa solujen väliseen signaalointiin osallistuvia proteiineja (Chervitz ym. 1998). Nämä ovat siis evoluution myötä ilmenyneitä geenejä, jotka ovat välttämättömiä monisoluisten eliöiden kehittymiselle ja toiminnalle. Banaanikärpäsän proteiineista 20 % huolehtii geenien ilmentymisestä, 20 % on entsyymejä ja vajaat 10 % säätelee solunjakautumista tai signaalinvälitystä. Kolmasosa geeneistä on rakenteeltaan aikaisemmin tuntemattomia. Eräät proteiiniperheet osoittautuivat erityisen suuriksi. Banaanikärpäsestä löytyi 352 sinkkisormiproteiinia (madosta 138), jotka säätelevät muiden geenien ilmentymistä, ja noin 300 proteiini-kinaasia, jotka osallistuvat proteiinien valmistamiseen ja muokkaamiseen. Madossa proteiini-kinaaseja on noin 500, ja ihmisestä niitä arvelaan löytyvän noin 1 100.

Elämäni geenit

Geenillä tarkoitetaan DNA-ketjun osaa tai osia, joita tarvitaan yhden proteiinin valmistamiseen. Tämän määritelmän perusteella geeniin kuuluu sekä proteiinia koodaavia että geenin ilmentymistä sääteleviä DNA-jaksoja. Lajien genomien koko ja geenien määrä vaihtelevat valtavasti (taulukko 2). Nyt kun useiden erilaisten malliorganismien genomit on selvitetty tai ovat selviämässä, on lajien geenejä verrattu keskenään pyrkimyksenä löytää niitä geneettisiä ominaisuuksia, jotka olisivat yhteisiä eri lajeille ja jotka siten voisivat muodostaa elämän välttämättömän perintömateriaalin. Bakteereiden genomeja vertaamalla on päädytty geenien minimimäärästä arvioon, jonka mukaan alkeellisimmat organismit tulivat toimeen 128 geenillä ja ny-

Taulukko 2. Genomien koko ja geenien määrä eri lajeissa (Lewin 2000).

Laji	Genomi, milj. emäsparia	Geenien määrä	Tappavia mutatioita
<i>Bakteerit ja hiivat</i>			
Mycoplasma genitalium	0.58	470	300
Rickettsia prowazekii	1.11	834	
Haemophilus influenzae	1.83	1 743	
Methanococcus jannaschi	1.66	1 738	
Bacillus subtilis	4.2	4 100	
Escherichia coli	4.6	4 288	1 800
Saccharomyces cerevisiae	13.5	6 034	3 600
<i>Kasvit</i>			
Arabidopsis thaliana	135	25 000	
<i>Eläimet</i>			
Drosophila melanogaster	180	13 600	3 100
Caenorhabditis elegans	97	19 099	
Homo sapiens	3 300	40 000–100 000?	

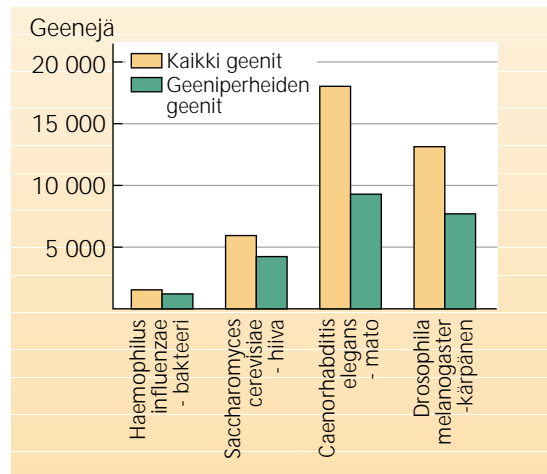
kyiset kehittyneemmät pieneliöt tarvitsevat vain noin 250 geeniä (Sariola ja Thesleff 1997). Toinen tapa tarkastella elämälle välttämättömiä geneejiä on tehdä koko genomien kattava mutaatioanalyysi. Tässä analyysissä E. colin geneeistä puolet osoittautui välttämättömiksi ja mutaatio 12 %:ssa hiivan geneejiä johti kuolemaan, eli kummassakin lajissa ainoastaan osa geneeistä on elämälle välttämättömiä.

Malliorganismeista on myös etsitty ihmisen tautigeenien sukulaisia. Kun karpäsen geneejiä verrattiin 289:een ihmisen tautigeeniin, samankaltaisuutta löytyi 177 geenistä, mikä tietysti tarjoaa aivan uusia mahdollisuuksia tutkia näiden tautien patogeneesiä (Rubin ym. 2000). On arvioitu, että hiiren ja ihmisen geenit muistuttavat toisiaan 97-prosenttisesti ja että vielä hiivan ja ihmisen geneeistä kolmasosa on sukulaisia. Keskeisten molekyylien hämmästyttävää rakenteellista ja toiminnallista samanlaisuutta kuvaa esimerkiksi se, että yksi karpäsen proteiini indusoi luun muodostusta jyräjoiden kudoksessa samaan tapaan kuin sen sukulainen rotassa, vaikka karpäsellä ei luita olekaan. Rotan vastaavan geenin siirto karpäseen korjaa karpäsegeenin mutaatiosta johtuvan kehityspoikkeaman (Sariola ja Thesleff 1997).

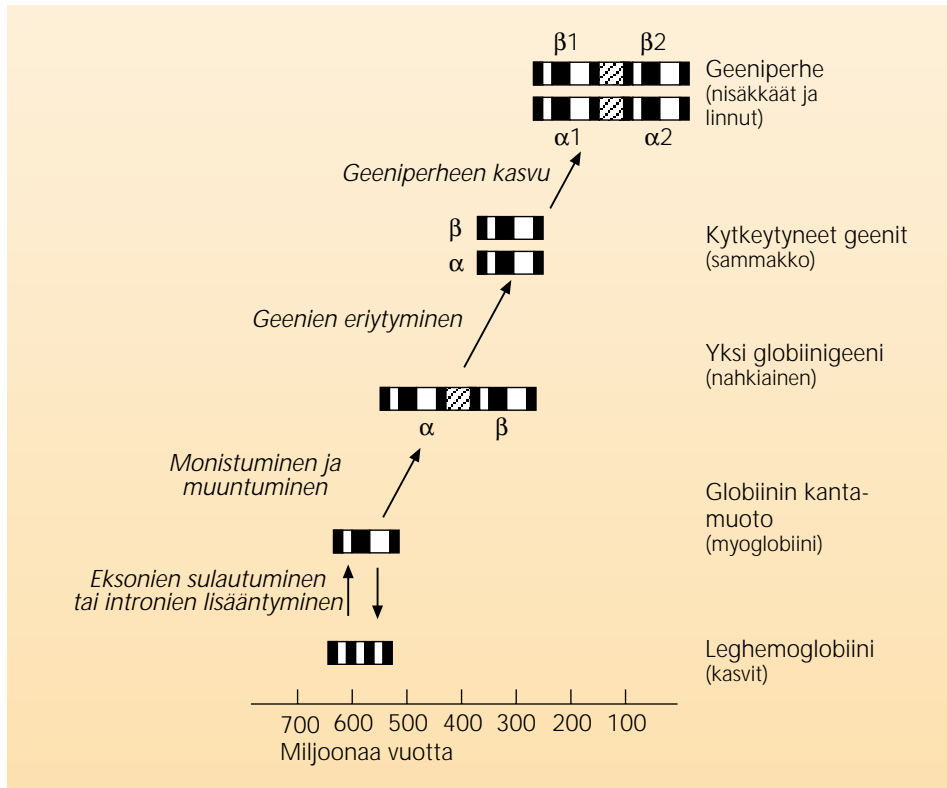
Geeniperheet ja -monistumat. Vertailtaessa malliorganismien geneejiä on huomattu, että suuri osa geneeistä muodostaa perheitä (kuva 2) ja monet niistä esiintyvät kaikissa tutkituissa lajeissa. Geeniperheiden jäsenten määrä vaihtelee eri lajeissa, ja nisäkkäillä niitä on enemmän kuin esimerkiksi banaanikarpäsellä tai sukkulamadolla.

Geenien määrän arvellaan lisääntyvän evoluutiossa monistumien ja uudelleenjärjestymien kautta. Vaikka lopullista varmuutta monistuman keskeisestä asemasta evoluutiossa ei vielä olekaan, siihen viittaa moni havainto. Esimerkiksi voidaan ottaa globiinigeenien evoluutio (kuva 3). Kuitenkaan genomien koon perusteella ei voida ennustaa sen sisältämien geenien määrää, eikä korrelaatio lajin kehitystason ja geenien määrän välillä pidä paikkaansa enää kehittyneissä lajeissa.

Redundanssi. Samaa biologista tehtävää varten solussa on usein rinnakkaisia järjestelmiä. Tätä kutsutaan redundanssiksi. Sen avulla monisolainen organismi pyrkii ilmeisesti turvaamaan keskeiset toiminnot ja suojautumaan mutaatioita vastaan. Yksi geneettinen selitys redundanssin mekanismille ovat geeniperheet, jotka usein vaikuttavat samassa biologisessa tapahtumassa. Samaa tehtävää solussa voivat hoitaa kuitenkin täysin erilaiset molekyylit, joiden vaikutustiet yhtyvät vasta tumatasolla. Redundanssin merkitys elämän ylläpitäjänä on ymmärretty



Kuva 2. Suuri osa geneeistä edustaa geeniperheitä.



Kuva 3. Kaikki globiinigeenit ovat syntyneet yhdestä kantageenistä monistamalla, muuntamalla ja liikkumalla.

konkreettisesti vasta poistogeenisten hiirten ansiosta (Sariola 1996). Vaikka keskeisenä pidetty geeni inaktivoidaan, hiirissä ei aina saadakaan aikaan epämuodostumia tai toimintahäiriöitä. Ne ilmenevät vasta mutatoitaessa useita samaan tapahtumaan vaikuttavia geenejä.

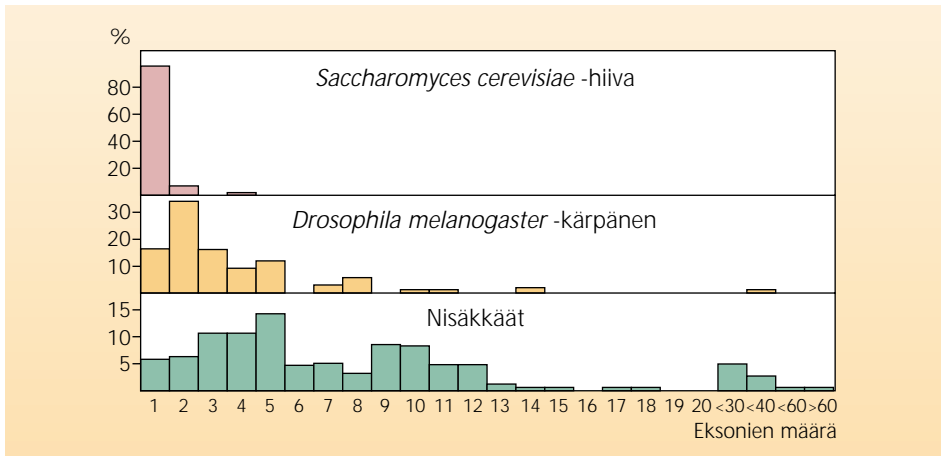
Genomin koodaavat ja ei-koodaavat alueet.

Ainoastaan pieni osa genomista edustaa sitä osaa, jossa tapahtuu lähetti-RNA:n luenta. Tätä osaa kutsutaan koodaavaksi alueeksi. Loppu on ei-koodaavaa aluetta, jolle ovat tyypillisiä DNA:n toistojaksot. Ei-koodaavan alueen osuus genomista on suuri ainoastaan kehittyneissä monisoluisissa lajeissa, eikä sitä juuri esiinny esimerkiksi hiivan genomissa. Ihmisen genomista arviolta yli 90 % on ei-koodaavaa aluetta. Tämän alueen merkitys on vielä suurelta osin hämärän peitossa.

Jatkuvat ja epäjatkuvat geenit. Alkeellisissa lajeissa geenit ovat jatkuvia eli valtaosassa gee-

nejä on ainoastaan yksi eksoni (yhtenäinen proteiinia koodaava alue). Kehittyneiden lajien geenit ovat epäjatkuvia eli ne sisältävät useita eksonia, joiden välissä olevia alueita kutsutaan introneiksi (kuva 4). Niistä tapahtuu luenta esi-RNA:ksi, mutta intronialueiden jaksot poistuvat lähetti-RNA:sta. Intronialueet saattavat olla hyvinkin pitkiä, usein pidempiä kuin eksonit. Vaikka intronien koodi ei siirrykään proteiiniin, niillä on usein keskeinen merkitys geenin ilmentymisen säätelyssä.

Orpogeenit. Tähän mennessä sekvensoitujen eliöiden geneeistä 35–50 % on orpogeeniä. Ne ovat geenejä, joille ei ole löytynyt vastingeeniä mistään organismista ja joiden koodittamien proteiinien tehtävistä ei siis tiedetä mitään. Viime huhtikuussa oli kaikkien sekvensoitujen organismien orpogeenien yhteenlaskettu määrä noin 200 000 (Craig Venter, henkilökohtainen tiedonanto). Niissä riittää tutkijoille työtä. Esimerkiksi



Kuva 4. Useimmat geenit ovat jatkuvia hiivassa (niissä on ainoastaan yksi eksoni) mutta epäjatkuvia kärpäsessä ja nisäkkäissä (niissä on useita eksoneita) (Lewin 2000).

leivinihiivan geneeistä yli kolmasosa on orpogeeniä. Eurooppalaiset ja yhdysvaltalaiset konsortiot kartoittavat parhaillaan niiden tehtäviä tuhoamalla ne yksitellen ja tutkimalla kunkin geenin puutteen vaikutuksia solun toimintaan. Toistaiseksi on selvinnyt, että vain 17 % orgopeeneistä on välttämättömiä elämälle (Winzeler ym. 1999).

Tulevaisuudennäkymiä

Koska hiiren ja ihmisen geenit ovat säilyneet lähes sataprosenttisesti samanlaisina, voidaan käytännössä kaikkien ihmisen geenien vaikutuksia tutkia geeninsiirtotekniikan avulla hiiressä. Hiirimalli on mahdollista haluttaessa tehdä kaikille ihmisen taudeille, joiden geenitausta on löytynyt. Geenien tarkka lukumäärä ei ole toistaiseksi tiedossa yhdenkään nisäkkään osalta, mutta kaikilla nisäkkäillä on arvioitu olevan 70 000–140 000 geeniä. Koska jokainen niistä voi olla tautigeeni ja suuri osa taudeista on multigeenisia, töitä riittää. Viimeisin arvio ihmisen geenien määrästä on onneksi huomattavasti edellä mainittua pienempi. Kun ihmisen kromosomi 21 sekvensoitiin toukokuussa 2000, siitä löytyi vain 225 geeniä, ja tämän perusteella uusi arvio ihmisen geenien määrästä on »enä» 40 000, kolme kertaa suurempi kuin banaani-kärpäsessä (Reeves 2000).

Tautigeenien tutkimiseen tullaan kuitenkin käyttämään yhä enemmän yksinkertaisia malliorganismeja, koska ne tuottavat hiirtä nopeammin uutta tietoa. Kun eri malliorganismien ja ihmisen geenien samankaltaisuudet on selvitetty, ihmisen geenien merkitystä ja toimintaa voidaan tutkia kysymykseen parhaiten soveltuvassa malliorganismissa. Banaanikärpäsen etuna on laaja tieto geeniverkostoista. Sen avulla voidaan selvittää, minkä geenien kanssa tutkittavana oleva ihmisen tautigeeni toimii. Hiivasoluilla kannattaa tutkia esimerkiksi solusykliä ja sen häiriöitä.

Malliorganismien antama geeni-informaatio lisää tietojamme likimain kaikilla biologian ja lääketieteen alueilla. Alkionkehityksen tutkijoille geenikartat tulevat olemaan kultakaivos. Onkin odotettavissa, että kehityksen mekanismit alkavat selvitä vauhdikkaasti, kun eri malliorganismien edut yhdistetään kattavaan geenitietoon. Jännittävää tulee olemaan tieto niistä geneeistä, jotka eivät muistuta mitään nykyisin tunnettua geeniä – esimerkiksi joka kolmas banaanikärpäsen geeni on aiemmin tuntematon. Todennäköisesti avautuu uusia näkymiä solujen toimintaan, ehkä löydetään kokonaan uusia periaatteita tai metabolisia reittejä. Jossain vaiheessa saamme selville myös sen, miten lajille ominaiset käyttäytymispiirteet ja henkiset ominaisuudet periytyvät.

Elämän sukupuusta tehtyä mallia joudutaan varmasti muuttamaan geenitiedon karttuessa. Opimme ymmärtämään paremmin evoluution kulkua, mutta vie kauan, ennen kuin on selvitetty, miten ja minkälaisien geenien muuttuminen on aiheuttanut eri lajien syntyminen. Muu-

tokset genomissa ovat usein tapahtuneet geenien koodaavien alueiden ulkopuolella – niiden säätelyalueilla, jotka määräävät ajan ja paikan, missä geeni toimii. Tämän alueen tutkimus on vasta alkuvaiheessa.

Kirjallisuutta

- Adams MD, Celniker SE, Hol RA, ym. The genome sequence of *Drosophila melanogaster*. *Science* 2000;287:2185–95.
- Botstein D, Chervitz SA, Cherry JM. Yeast as a model organism. *Science* 1997;277:1259–60.
- Brändli AW. Towards a molecular anatomy of the *Xenopus* pronephric kidney. *Int J Devel Biol* 1999;43:381–96.
- Chervitz SA, Aravind L, Sherlock G, ym. Comparison of the complete protein sets of worm and yeast: orthology and divergence. *Science* 1998;282:2022–4.
- Goffeau A, Barrell BG, Bussey H, ym. Life with 6000 genes. *Science* 1996;274:546.
- Hengartner MO. Programmed cell death in the nematode *C. elegans*. *Rec Prog Hormon Res* 1999;54:213–22.
- Kornberg TB, Krasnow MA. The *Drosophila* genome sequence: implications for biology and medicine. *Science* 2000;287:2218–20.
- Lewin B. *Gene* 2000. www.genes.net
- Malakoff D. The rise of the mouse, biomedicine's model mammal. *Science* 2000;288:248–53.
- Makarow M. Hiiwan perimän salat julki. *Duodecim* 1996;112:1737–8.
- Morrow DM, Tagle DA, Shiloh Y, Collins FS, Hieter P. TEL1, an *S. cerevisiae* homolog of the human gene mutated in ataxia telangiectasia, is functionally related to the yeast checkpoint gene MEC1. *Cell* 1995;82:831–40.
- Myers EW, Sutton GG, Delcher AL, ym. A whole-genome assembly of *Drosophila*. *Science* 2000;287:2195–204.
- O'Brian JS, Menotti-Raymond M, Murphy WJ, ym. The promise of comparative genomics in mammals. *Science* 1999;286:458–81.
- Patino MM, Liu J-J, Glover JR, Lindquist S. Support for the prion hypothesis for inheritance of a phenotypic trait in yeast. *Science* 1996;273:622–5.
- Peltonen-Palotie L. Kilpa ihmisen perimästä kuumentaa tiedemaailmaa. *Helsingin Sanomat*. 6.5.2000.
- Plowman GD, Sudarsanam S, Bingham J, Whyte D, Hunter T. The protein kinases of *Caenorhabditis elegans*: a model for signal transduction in multicellular organisms *Proc Natl Acad Sci USA* 1999;96:13603–10.
- Reeves RH. Recounting a genetic story. *Nature* 2000;405:283–4.
- Rubin GM, Yandell MD, Wortman JR, ym. Comparative genomics of the eukaryotes. *Science* 2000;297:2204–15.
- Rubin GM, Lewis EB. A brief history of *Drosophila*'s contributions to genome research. *Science* 2000;287:2216–8.
- Sariola H. Does the kidney express redundant or important molecules during nephrogenesis? *Exp Nephrol* 1996;4:70–6.
- Sariola H, Thesleff I. »Kyllä molekyyliä maailmaan mahtuu«. *Duodecim* 1997;113:1393–400.
- Service RF. Can Celera do it again? *Science* 2000;287:2136–8.
- The *C. elegans* Sequencing Consortium. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science* 1998;282:2012–8.
- Weinberg AM. The birth of big biology. *Nature* 1999;401:738.
- Winzeler EA, Shoemaker DD, Astromoff A, ym. Functional analysis of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* 1999;285:901–6.

HANNU SARIOLA, professori

hannu.sariola@helsinki.fi

IRMA THESLEFF, professori

Biotekniikan instituutti, kehitysbiologian tutkimusohjelma

PL 56, 00014 Helsingin yliopisto

MARJA MAKAROW, professori

Biotekniikan instituutti

PL 56, 00014 Helsingin yliopisto

ja

Kuopion yliopiston biokemian laitos

PL 1627, 70211 Kuopio