

1 2015 Luonnon Tutkija



- 3 **Pääkirjoitus** Lintu-Palménin jäljillä
Anto Leikola

- 4 **Esbo-Lövön kesälaboratorio 1889—99**
Aikamatka Tvärminnen eläintieteellisen aseman juurille
professori J. A. Palménin muistoksi
Marjatta Rautiala

- 12 **DNA ja taksonomian vallankumous**
Marko Mutanen

Kirjoja

- 22 Anderson, J. T. & Davis, C. A. (toim.) 2013: Wetland techniques. 1. Foundations. 2. Organisms. 3. Applications and management
Pertti Koskimies
- 22 Birmelin, I. 2012: Von wegen Spatzenhirn! Die erstaunlichen Fähigkeiten der Vögel
Pertti Koskimies
- 23 Stringer, C. 2014: Vain yksi jäi. Miten meistä tuli ainoa ihmislaji
Anto Leikola
- 24 Shrubbs, M. 2013: Feasting, fowling and feathers. A history of the exploitation of wild bird
Pertti Koskimies

Vanamo

- 25 Kertomus Suomen Biologian Seuran Vanamon ry toiminnasta vuonna 2014
- 27 Suomen Biologian Seuran Vanamon ry toimintasuunnitelma vuodeksi 2015
- 27 Vanamon apurahat 2015

Väitöksiä

- 28 Siirtogeeninen kolibakteeri auttaa havaitsemaan rakennusten homevaurioita
(Janne Atosuo)
Lisäksi 18 väitöskirjaviitettä
Koonnut Ahti Launis

Nuori Anssi

- 31 Kukka

Luonnon Tutkija

Perustettu vuonna 1897
Vuoteen 1946 **Luonnon Ystävä**

Julkaisija
Suomen Biologian Seura Vanamo ry

Päätoimittaja
Juhani Mänttari

Toimituksen osoite
PL 44 (Jyrängöntie 2)
00014 HELSINGIN YLIOPISTO
puhelin 02941 50059, 0400 585 720
sähköposti juhani.manttari@helsinki.fi

Taitto Juhani Mänttari

Avustava toimittaja Ahti Launis

Toimitusneuvosto
Jari Valkonen (puh.joht.)
Helena Korpelainen Jussi Kukkonen
Satu Lampila Riitta Savolainen
Ilari Sääksjärvi Teemu Teeri

Ilmestyminen
Vuonna 2015 ilmestyy neljä numeroa.

Tilaaminen
Vanamon jäsenet saavat lehden jäsenetuna. Muille vuosikerran 2015 tilaushinta on 42 euroa. Tilataan toimituksesta tai maksamalla tilaushinta Vanamon tilille Danske Bank FI51 8000 1100 0702 50. Maksun tiedotteeksi on merkittävä tilaajan osoite sekä tilauskausi. Irtonumeron hinta 11 euroa.

Kirjoitusohjeet nrossa 1/2015 s. 42

Ilmoitushinnat	kokosivu	400 euroa
	puolisivu	250 euroa
	takakansi	600 euroa

Numeroon 3/2015 tarkoitetun aineiston tulee olla toimituksessa viimeistään 14. 8. 2015.

Painopaikka Kirjapaino Uusimaa, Porvoo
ISSN 0024-7383

Vanamon Internet-sivut www.vanamo.fi

Kansikuva Esbo-Lövön kesälaboratoriona palvelleet kalastajien mökit Pienessä Lehtisaareessa Espoossa 1890-luvulla. — Sivulla 4—11 kerrotaan Suomen ensimmäisestä biologisesta kenttäasemasta.
Valok. J. A. Palmén (Tvärminnen arkisto)



Luonnon Tutkija ottaa mielellään julkaistavaksi kirjoituksia biologian alalta.

Tilaukset ja osoitteenmuutokset
Luonnon Tutkijan toimitukseen

Irtonumeromyynti
Tiedekirja, Snellmaninkatu 13, 00170 HELSINKI,
puhelin (09) 635 177, sähköp. tiedekirja@tsv.fi ja
Akateeminen Kirjakauppa, Keskuskatu 1, Helsinki,
puhelin (09) 121 4322, www.akateeminen.com

DNA ja taksonomian vallankumous

Marko Mutanen

DNA:n emäsjärjestyksen lukemistekniikan nopea kehitys on mullistamassa lajintunnistusta ja taksonomista tutkimusta. Monissa maissa, myös Suomessa, kootaan kattavaa DNA-viivakoodikirjastoa, joka mahdollistaa lajien automaattitunnistuksen standardi-DNA-jakson perusteella. Eliöiden sukulaisuussuhteita ja lajirajoja voidaan selvittää vertaamalla kokonaisia genomeja toisiinsa, ja laajoja aineistoja voidaan määrittää nopeasti ja automaattisesti.

DNA:n emäsjärjestys määrittelee eliölajien kaikki perinnölliset ominaisuudet. Biologit, taksonomit aivan erityisesti, ovat tavallisesti kiinnostuneita juuri lajien tai yksilöiden perinnöllisistä ominaisuuksista. Lajien tunnistaminen ja rajaaminen perustuvat periytyvien ominaisuuksien vertailuun. Niinpä genomi sisältää kaiken oleellisen tiedon myös eliöiden morfologisista ja muista periytyvistä ominaisuuksista. DNA:n neljän emäksen tulkitseminen on yksiselitteistä, ja tieto on siinä säilyttynä alkuperäismuodossaan. Taksonomian tieteenala käy parhaita todellista DNA-vallankumousta. Vallankumous johtuu ennen kaikkea kahdesta valtavasta, jopa paradigmaattiseksi kutsutusta kehitysaskeleesta (Hajibabaei 2012): DNA-viivakoodit ovat mahdollistamassa standardoidun lajintunnistuksen, ja uuden sukupolven sekvensointitekniikat mahdollistavat mikroskooppisen eliöstön ja laajojen aineistojen automaattisen mää-

rittämisen ja yksilön genomien tutkimisen nopeasti ja halvalla.

Biodiversiteettikriisi ja taksonomian kriisi

Monenlaisten ympäristöuhkien, ennen kaikkea laajamittaisen elinympäristöjen tuhoutumisen, vuoksi suuri osa maapallon eliölajeista uhkaa kuolla sukupuuttoon. Synkältä näyttävää tilannetta on kutsuttu biodiversiteettikriisiksi. Karkeasti arvioiden vain noin 10 % maailman eliölajeista on tieteellisesti kuvattu, joten suuri osa sukupuuttoon kuolevista lajeista häviää ilman, että niiden olemassaolosta koskaan tullaan edes tietämään. Luonnon monimuotoisuudella on myös ihmisen kannalta arvaamattoman suuria vaikutuksia. Esimerkiksi allergioiden runsastuminen on voitu yhdistää monimuotoisuuden vähenemiseen (Hanski ym. 2012).

Sen jälkeen, kun Carl von Linné kehitti nykyaikaisen luokittelujärjestelmän noin 250 vuotta sitten,

lajeja on kuvattu arviolta hiukan alle kaksi miljoonaa. Valtaosa lajinkuvaustyöstä on siis tekemättä, ja tätä vauhtia kestäisi vielä vuosisatoja, ennen kuin kaikki lajit on kuvattu. Sitä paitsi liikkeelle on lähdetty suurikokoisista ja hyvin tunnistettavista lajeista, siis helpoimmasta päästä. On arvioitu, että mikäli lajeja on maailmassa 10–15 miljoonaa (mikä lienee aika realistinen arvio), niiden tunteminen perinteisin menetelmin edellyttäisi 15 000 lajiasiantuntijan osaavaa joukkoa (Hebert ym. 2003). Lisäksi osaamisen pitäisi olla tasaisesti jakaantunut eri eliöryhmiin, ja tietämyksen pitäisi siirtyä sukupolvelta toiselle. Maailmassa on tällä hetkellä arviolta 4 000–6 000 ammattimaista taksonomia eli lajien luokitteluun ja lajinkuvauksiin perehtynyttä asiantuntijaa. Taksonomien osaaminen painottuu voimakkaasti tiettyihin eliöryhmiin, kun taas toiset jäävät kokonaan ilman erityisosajia. Lukuisat runsaslajiset eliöryhmät, esimerkiksi pienet maaselkärangattomat ja mikroskooppiset sienet, tunnetaan erittäin huonosti, koska perinteisin morfologisin menetelmin lajien erottamiseen ei löydy riittävästi tuntomerkkejä. Näin ollen maailman koko lajiston kuvaaminen perinteisin keinoin on epärealistista. Tilannetta on kutsuttu taksonomian kriisiksi, ja sitä vaikeuttaa entisestään lajien tunnistamisen ja lajirajojen määrittämisen vaikeus. Koska biodiversi-

teettitutkimuksen ja luonnonsuojelun keskeisenä edellytyksenä on riittävä tietämys ja ymmärtämys lajeista, biodiversiteettikriisin ratkaiseminen edellyttää myös taksonomian kriisin ratkaisua.

DNA:n merkitys taksonomisessa tutkimuksessa on kasvanut nopeasti

Nature-lehdessä julkaisemassaan kuuluisassa tutkimuksessa Francis Crick and James Watson (1953) kuvasivat DNA-molekyylin kaksoiskierrekenteen ja loivat tutkimuksellaan pohjan genetiikan tutkimuksen nopealle kasvulle. Kesti kuitenkin vielä yli 20 vuotta, ennen kuin Frederick Sanger kollegoineen (1977) kehitti tavallisissa laboratorioissakin käyttökelpoisen menetelmän DNA:n emäsjärjestyksen lukemiseen. Pitkään senkin jälkeen molekyylogeneettinen tutkimus perustui muihin DNA:n ominaisuuksiin kuin sen emäsjärjestykseen, esimerkiksi DNA-fragmenttien kokojakaumien vertailuun. Keskeisenä syynä oli se, että emäsjärjestyksen lukeminen oli kallista ja nykyiseen verrattuna hidasta ja työlästä. Sekvensointi on sittemmin halventunut niin nopeasti, että kehitystä on verrattu ns. Mooren lakiin, ts. sekvensoinnin hinta on puolittunut suunnilleen kahden vuoden välein. Huiman teknisen kehityksen johdosta on viime aikoina edetty vieläkin nopeammin. Myös tietokoneiden laskentakapasiteetin kasvamisen ja uusien analysointimenetelmien ja tietokoneohjelmien merkitys on ollut suuri DNA-sekvenssitiedon käytön nopeassa yleistymisessä taksonomiassa.

Taksonomit ovat käyttäneet DNA:han perustuvaa tietoa tutkimuksessaan käytännössä siitä lähtien, kun se tuli mahdolliseksi. DNA:n emäsjärjestystä on käytetty erityisesti fylogeneetisiin tutkimuksiin eli taksonien evoluutiohistorian selvittämiseen, mutta myös lajikysymysten ratkaisuun. DNA:ta on joko käytetty yksinään, tai sitten sen tarjoamaa tietoa on yhdistetty muuhun, ennen kaikkea morfologiseen tietoon. Mitä nopeammaksi ja halvemmaksi DNA:n emäsjärjes-

tyksen lukeminen on tullut, sitä merkittävämmäksi sen osuus kaikesta taksonomisesta tutkimuksesta on muodostunut. Fylogeneettisissä tutkimuksissa suuntaus on ollut tasaisen kasvava siten, että valtaosa nykyään julkaistavista fylogeneettisistä tutkimuksista pohjautuu kokonaan tai osittain DNA:han. DNA:n avulla on ratkaistu monia pitkäkestoisia mysteereitä, kuten valaiden sukulaisuussuhteet muihin nisäkkäisiin (valaat polveutuvat sorkkaeläimistä, ja virtahepo on lähisempää sukua valaille kuin muille sorkkaeläinryhmille). Taksonomisessa tutkimuksessa DNA:n merkitys on viimeisen kymmenen vuoden aikana kasvanut vielä nopeammin. Käänteentekevä vuosi oli 2003, jolloin julkaistiin lähes samanaikaisesti kaksi sittemmin runsaasti keskustelua herättänyttä artikkelia. Toisessa artikkelissa esitettiin toive siitä, että taksonomia, ennen kaikkea lajien rajaaminen, perustuisi tulevaisuudessa ainoastaan DNA:han (Tautz ym. 2003). Toisessa esitettiin ajatus DNA-viivakoodista, vakioiduista DNA-jaksoista, jotka mahdollistaisivat automatisoidun lajintunnistuksen (Hebert ym. 2003). Samantapaisia ajatuksia oli esitetty aiemminkin, mutta idea viivakoodistandardista ja kaikki lajit kattavasta viivakoodien referenssikirjastosta sai muotonsa juuri tässä artikkelissa, joka julkaisunsa jälkeen on ollut siteerautimpia biologian alan julkaisuja. Hebertin ja hänen kollegoidensa artikkelilla oli valtava merkitys taksonomisessa yhteisössä. Se paitsi herätti voimakkaita vastareaktioita myös innosti lukuisia taksonomeja ja toimi pontmina kautta aikain suurimmalle genomien biodiversiteettihankkeelle, iBOL:lle (International Barcode of Life; <www.ibol.org>), jonka osana kansallinen FinBOL-hanke (Finnish Barcode of life; <www.finbol.org>) on tuottanut DNA-viivakoodin jo reilusti yli 10 000 Suomessa tavattavasta lajista. Hyvin pian huomattiin, että paitsi lajien määrittämisessä viivakoodit tarjoavat arvokasta informaatiota myös itse taksonomiseen työhön, esimerkiksi lajirajojen määrittelyyn. DNA-viivakoodien ansiosta on havaittu lukuisia hankalasti toi-

sista lajeista erotettavia piilolajeja. Esimerkkejä löytyy Suomenkin lajistosta. Paljolti juuri DNA-viivakoodien ansiosta taksonominen tutkimus elää monissa maissa todellista uutta kukoistusta. Laajojen kansallisten ja kansainvälisten hankkeiden ja DNA-viivakooditietokannan (BOLD; <www.boldsystems.org>) helppokäyttöisyyden ansiosta monet amatööritaksonomit ovat omaksuneet DNA-viivakoodien käytön uudeksi informaation lähteeksi tutkimuksissaan.

DNA-viivakoodit ja niiden hyöty lajintunnistuksessa

DNA-viivakoodista eli DNA-lajintunnisteista on kirjoitettu Luonnon Tutkijassa useaan otteeseen aiemmin (Korpelainen 2008, 2010, 2013 a). Näin ollen en käy yksityiskohtaisesti läpi idean perusteita ja pyrin samalla minimoimaan päällekkäisyyttä edellisten kirjoitusten kanssa. Koska olen hyönteistutkija, otan esimerkkitapaukset hyönteismaailmasta, kun taas Korpelaisen käyttämät esimerkit koskevat kasveja. Käsittelen tässä kirjoituksessa kuitenkin laajemmin DNA-viivakoodien etuja ja haittoja, koska niiden käsittely on aiemmin jäänyt melko suppeaksi.

DNA-viivakoodi on standardi, lajien erotteluun soveltuva jakso DNA:ta. &Eläimissä DNA-viivakoodiksi on vakiintunut osa mitokondrion c-sytokromioksidaasiproteiinin 1. alayksikköä koodittavaa geeniä (COI tai *cox1*). Kyseinen geenijakso muuntuu ajan kuluessa aika nopeasti, mutta vaihtelee lajin sisällä yleensä vähän ja toimii siten hyvin lajintunnisteena. COI-geenijaksolla on lisäksi kaksi muuta hyvin merkittävää etua. Ensinnäkin siitä on useimmissa soluissa lukuisia kopioita, koska soluissa on tavallisesti useita mitokondrioita. Tämä helpottaa sen sekvensoimista esimerkiksi huonokuntoisista näytteistä tai vähäisestä lähtömateriaalista. Toiseksi c-sytokromioksidaasi on soluhengityksessä välttämätön entsyymi, joten sitä koodittava geeni löytyy taatusti jokaisesta eläinlajista ja -yksilöstä. Ilman toimivaa COI-geeniä mikään soluissaan hengittävä eliö ei voi elää.

Määrittelyn suhteen tarkastelen DNA-viivakoodien etuja ja haittoja siitä hypoteettisesta lähtökohdasta, että on olemassa yksi ”oikea” lajiluokittelu eli että lajit ovat yksiselitteisesti rajattavia luonnonyksiköitä. Tosiasiassa tilanne ei ole koskaan näin yksinkertainen ja selvä, kuten myöhemmin käy ilmi.

DNA-viivakoodipohjaisessa määrittelyssä on monia merkittäviä etuja verrattuna perinteiseen morfologiseen lajintunnistukseen. DNA-viivakoodit mahdollistavat jopa laajojen aineistojen tunnistuksen automatisoinnin ja vähentävät riippuvaisuutta lajiasiantuntijoista, joita ei aina ole edes käytettävissä.

Kuva 1 Kaksi esimerkkiä DNA-viivakodein määritetystä tuntemattomasta hyönteistoukasta. — Ylemmässä kuvassa oleva toukka löytyi siiankärsämön juuren ja varren taitteesta syömässä kasvia sisäpuolelta. DNA-viivakodein laji osoittautui *Epiblema graphana* -kääriäisperhoslajiksi. — Alemman kuvan toukka löytyi yhteen kudottujen pihlajan lehtien välistä. Tämäkin oli kääriäisperhosiin kuuluva *Ancylis selenana*. Kumpaakaan lajia ei ollut koskaan kasvatettu Suomessa aiemmin ja jälkimmäisen osalta ravintokasvitieto oli kokonaan uusi.



Viivakoodimäärittysten oikeellisuus ei ole asiantuntijan pätevyyden varassa. Perinteisessä määrittämisessä virhemarginaali voi olla huomattavan iso ja määrittästyön laatua on vaikea arvioida. Perinteinen määrittäminen edellyttää yleensä sitä, että määritettävä yksilö on riittävän hyväkuntoinen. Määrittäminen vaikeutuu tuntomerkkien kuluessa tai puuttuessa muusta syystä. Määrittäminen DNA-viivakodein onnistuu yhtä lailla niin kokonaisesta ja hyväkuntoisesta yksilöstä kuin pienestä kappaleesta tai kuluneesta yksilöstä. Määrittämisen epävarmuus on laskennallisesti arvioitavissa paljon paremmin kuin perinteisessä määrittämisessä. Valtava etu DNA-viivakodeissa on se, että lajinmäärittämisessä käytetään aina yhtä ja samaa tuntomerkkiä, oli kysymyksessä mikä tahansa laji tai mikä tahansa yksilön kehitysvaihe. Ei tarvita erillisiä määrittäyskaavoja vaikkapa hyönteisten aikuisista, toukista ja koteloista (munista sellaisia ei juuri edes ole tehty). Nuorten kasvien määrittäminen on usein hankalaa, ja sienirihmastoja tai sukkulamatojen määrittäminen säännönmukaisesti lähes mahdotonta. Tutkijoilla on kuitenkin erilaisia tarpeita määrittää nuoria kehitysvaiheita. Esimerkiksi määrittämällä hyönteistoukat DNA-viivakodein elintapojen ja ravintoverkkojen selvittäminen helpottuu (kuva 1). Aiemmin loisitusta hyönteistoukasta voitiin määrittää vain joko isäntä- tai loislaji. DNA-viivakoodien avulla voidaan määrittää paitsi loinen myös isäntä sen jäännösten perusteella. Nykitekniikalla on mahdollista selvittää loishyönteisen isäntälaji jopa analysoimalla sen mahansisältö (Rougerie ym. 2011). Suomessakin akatemiaprofessori Ilkka Hanskin johtamassa tutkimuksessa on vastikään viivakodein saatu tietoa *Ophion*-suvun loispistiäisten isäntälajeista tutkimalla loisten mahansisältöä (julkaisematon tieto). Aiemmin tätä oli yritetty vuosikausia selvittää kasvattamalla mahdollisten isäntälajien toukkia — onnistumatta.

DNA-viivakoodien kiistattomaksi hyödyksi on osoittautunut niiden potentiaali paljastaa aiemmin huomaamatta jääneitä ulkonäöltään

samanlaisia piilolajeja. Usein tällaisissa tapauksissa DNA-viivakoodista on löytynyt yhden lajin sisältä kahta tai useampaa jyrkästi toisistaan poikkeavaa haplotyyppiä. DNA-viivakoodi yksistään ei ole tae piilolajista, mutta monissa tapauksissa myöhemmät tutkimukset ovat varmistaneet asian. Suomestakin tunnetaan useita tällaisia tapauksia. Esimerkiksi ulkonäöltään hyvin muuntelevan *Epinotia nisel-la*-kääriäisperhosen havaittiin sisältävän kahta aivan erilaista DNA-viivakoodia, ja myöhemmässä tutkimuksessa tämä ero kyettiin yhdistämään genitaalituntomerkkieroihin, ulkonäköeroihin ja ekologisiin eroihin (Mutanen ym. 2012 a). Maailmalta tunnetaan raflavampiakin esimerkkejä. Esimerkiksi Pohjois- ja Etelä-Amerikassa laajalle levinneen, suurikokoisen ja näyttävän *Astraptes fulgurator* -paksupääperhosen todettiin sisältävän kymmenen lajia (myöhemmin löytyi vielä yhdestoista), jotka eroavat toisistaan paitsi DNA-viivakoodin myös toukkien ulkonäön ja ekologian puolesta (Hebert ym. 2004).

Paitsi tutkimuksessa DNA-viivakoodia voidaan soveltaa myös moni aloilla yhteiskunnassa, esimerkiksi vieraslajien seurannassa, elintarvikevalvonnassa (vrt. <<http://phe.rockefeller.edu/barcode/sushigate.html>>), uhanalaisten lajien kaupan valvonnassa ja vaikkapa tunnistettaessa sisäilman homeitiöitä (Korpelainen 2013 b).

DNA-viivakoodien ongelmat lajintunnistuksessa

DNA-viivakoodien käyttö lajinmäärityksessä ei aina ole ongelmattonta. Ongelmiin voivat johtaa sekä tekniset seikat että biologiset prosessit. Edellä mainittuihin kuuluvat ennen kaikkea erilaiset virheet, joita voi tapahtua luettaessa DNA-viivakoodia. Tyypillinen virhe on kontaminaatio eli näytteen saastuminen vieraasta DNA:sta. Kontaminaatiosta on vaikea päästä kokonaan eroon, mutta sen riskiä voidaan monin keinoin pienentää esimerkiksi parantamalla huolellisuutta eri työvaiheiden aikana. Kontaminaation riski kasvaa, kun analysoitavan näytteen DNA:n laa-

tu heikkenee. Tällöin pienikin määrä vierasta DNA:ta voi tuottaa väärän viivakoodin. Onneksi kontaminaatio on yleensä helppo todeta. Mikäli esimerkiksi hyönteisnäyte tuottaakin hyönteisen sisällä olleen loisen viivakoodin, kontaminaation tunnistaminen on yleensä vaivatonta. Tunnistaminen tulee sitä helpommaksi, mitä laajemmaksi DNA-viivakoodien vertailukirjasto kasvaa. DNA-viivakoodijaksosta voi olla myös toimimattomia kopioita tuman DNA:ssa (NUMT, < engl. *nuclear mitochondrial DNA*). Vaikka tällaiset ”valegeenit” ovat hyvin yleisiä, ne aiheuttavat harvoin ongelmia mm. siitä syystä, että niistä on vähän identtisiä kopioita kussakin solussa ja niissä voi olla mutaatioita myös alukealueella, jolloin ne eivät monistu. Joskus valegeeni voi kuitenkin monistua oikean viivakoodin sijasta tai ohessa, mutta silloinkin sen on yleensä helppo tunnistaa. Koska valegeenit eivät koodita proteiinia, niihin ei yleensä kohdistu valintaa, jolloin niiden evoluutionopeus on suuri. Tämä auttaa niiden tunnistamisessa. Valegeenit sisältävät yleisesti nukleotidien deleetioita, insertioita sekä päatekodoneja (aminohappoa koodittamattomia kodoneja, joihin lähetti-RNA:n translaatio aminohappoketjuksi päättyy). Teknisiin ongelmiin voidaan laskea myös se, että DNA-viivakoodin lukemista vaikeuttavat mutaatiot alukejaksoissa. Muutokset alukejaksoissa ovat kaksiteräinen miekka: ne mm. vähentävät merkittävästi hyönteisnäytteiden kontaminoitumista vaikkapa ihmisen tai loiseliöiden DNA:sta, mutta ne hankaloittavat viivakoodin lukemista samanaikaisesti laajasta yksilöjoukosta. Esimerkiksi perhosen ja pistiäisen DNA-viivakoodin lukemiseen tarvitaan erilaiset alukkeet. Joihinkin eliöryhmiin ei vielä ole kehitetty hyvin toimivia alukkeita. Joissakin ryhmissä DNA-viivakoodin lukeminen on muista syistä hyvin vaikeaa. Esimerkiksi kulta- ja mutipistiäisten DNA-viivakoodijakso muodostaa monistusvaiheessa eräänlaisen lenkin, mikä tekee sekvenssin lukemisesta normaalein menetelmin tavallista hankalampaa.

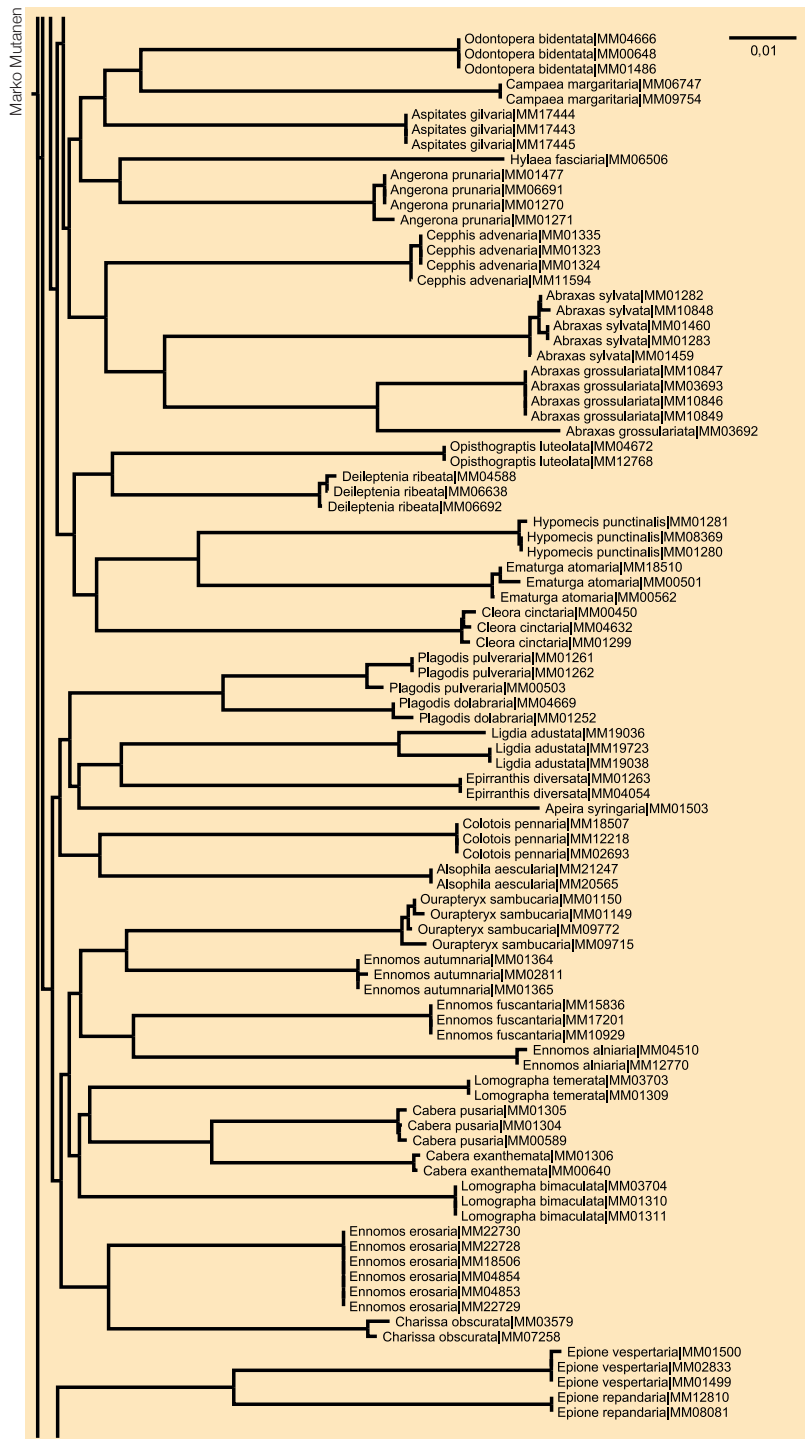
DNA-viivakoodien päätehtävä on mahdollistaa lajien erottaminen toisistaan. Useimmissa tutkimuksissa DNA-viivakoodien on todettu erottelvan yli 90 % lajeista, monissa huomattavasti suuremmankin osuuden. Joskus kaksi tai useampi laji voi kuitenkin omata samanlaisen tai hyvin samankaltaisen DNA-viivakoodin. Tällöin kyse voi olla siitä, että lajit ovat niin nuoria, etteivät kehityslinjat ole lajiutumisen jälkeen vielä geneettisesti erilaistuneet riittävästi. Ilmiön taustalla voi olla myös introgressio eli DNA:n ”vuotaminen” lajista toiseen risteytymisen ja takaisinristeytymisen seurauksena. Koska mitokondrio-DNA periytyy ainoastaan äidiltä eikä siinä yleensä tapahdu rekombinaatiota (se periytyy kokonaisuena ”pakettina”), lajien välisen risteytymisen ja takaisinristeytymisen seurauksena voi olla koko mitokondrion DNA:n siirtyminen lajista toiseen. Toisinaan lajissa on tällöin kahta haplotyyppiä, ”omaa” ja ”vierasta”, yhtä aikaa. Toisinaan tulokashaplotyyppi voi ”vallata” toisen lajin nopeastikin, erityisesti jos asiaan pääsee vaikuttamaan endosymbioottinen *Wolbachia*-bakteeri. *Wolbachia* on hyvin tavallinen niveljalkaisissa ja sukkulamadoissa, ja se voi aiheuttaa monenlaisia erikoisia ilmiöitä. Niistä DNA-viivakoodien kannalta ehkä merkittävin on se, että tartunta voi johtaa vieraan mitokondriohaplotyyppin nopeaan fiksoitumiseen (leviämiseen kauttaaltaan) uudessa lajissa (Hurst & Jiggins 2005). Lopputuloksena kahden lajin DNA-viivakoodit eivät eroa, kunnes aikaa kuluu taas niin paljon että uusia mutaatioita pääsee kertymään. Tämä aika vaihtelee mm. sukupolven pituuden ja efektiivisen populaatiokoon mukaan ja on käytännössä aina vähintään tuhansia vuosia.

Vaikka edellä mainitun kaltaiset ongelmat ovat tavallisimpia, tavaataan myös muunlaisia ongelmia. Yksi tällainen ongelma on heteroplasmia. Sillä tarkoitetaan usean erilaisen mitokondriohaplotyyppin samanaikaista esiintymistä yhdessä yksilössä ja solussa. Luonnollisesti DNA-viivakoodin määrittäminen tällöin hankaloituu. Heteroplasmia-

assa syntyy harha heterotsygotias-
ta, vaikka mitokondrio-DNA on
haploidi eli todellista heterotsygo-
tiaa ei siinä ole koskaan. Hetero-
plasmiaa on erityisesti muutamissa
nilviäisissä, joiden mitokondriot
periytyvät säännönmukaisesti mo-
lemmilta vanhemmilta.

Kuinka hyvin viivakoodit toimivat? Esimerkkejä perhosista ja kovakuoriaisista

On yllättävän vaikeaa arvioida
DNA-viivakoodien toimivuutta
määrittäessä, koska toimivuutta
joudutaan vertaamaan johonkin
olemassa olevaan luokitteluun. Yh-
denkään lajiluokittelu ei kuiten-
kaan voida sanoa olevan yksiselit-
teisesti totta. Luokittelu joko hei-
jastaa sen tekijän subjektiivista nä-
kemystä asiasta tai sitten perustuu
johonkin laskennalliseen standar-
diin, mikä on vähemmän subjektiiv-
vinen, mutta kuitenkin pohjimil-
taan keinotekoinen tapa rajata laje-
ja. Arvioon DNA-viivakoodien
toimivuudesta vaikuttavat heiken-
tävästi paitsi epätarkkuus ja suora-
naiset virheet vertailuluokittelussa
myös virhemääritykset. Esimerkik-
si geenipankissa on paljon virheel-
lisesti määritettyjä sekvenssejä, jo-
ten ei ole yllättävää, että geenipan-
kista imuroituun tietoon perustu-
vissa tutkimuksissa on saatu varsin
huonoja arvioita viivakoodien toi-
mivuudesta (Meier ym. 2006).
Yleensä taksonomeja kiinnostavat
enemmän hankalat lajiryhmät kuin
helpot, minkä vuoksi suhteettoman
suuri osuus tehdyistä tutkimuksista
on keskittynyt jo ennalta vaikeiksi
tiedettyihin lajiryhmiin, esimerkki-
nä *Grammia*-suvun siilikkäistä teh-
ty tutkimus (Schmidt & Sperling
2008). Vastaavasti ylioptimistiseen
estimaattiin on johtamassa se, että
otanta on harvoin edes paikallisesti
saati maantieteellisesti kattava, jol-
loin erot lajien välillä voivat vai-
kuttaa suuremmilta kuin ne todelli-
suudessa ovat. Kuten aiemmin kävi
ilmi, näistä seikoista riippumatta
DNA-viivakoodit eivät aina erotte-
le edes selvästi erillisiä, kiistatto-
masti ”hyviä” lajeja. Mutta kuinka
paljon tällaisia lajeja lopulta on?
Arvioin seuraavassa tilannetta Suo-
men perhos- ja kovakuoriaislajis-



Kuva 2 Esimerkki DNA-viivakoodien erottelukyvystä mittariperhosissa. — Kuva on ote laajemmasta neighbor-joining-menetelmällä tehdystä sukupuusta. Vaakasuorien osien yhteispituus kuvastaa geneettisen eron suuruutta lajien ja yksilöiden välillä (mittakaava kuvan yläreunassa: jana kuvastaa 1 %:n geneettistä eroa). Kuvasta ilmenee hyvin ns. barcoding gap -ilmiö: lajin sisäinen muuntelu on tyypillisesti suuruusluokkaa pienempää kuin lajien välinen muuntelu. Tästä on poikkeuksia.

ton näkökulmasta ja laajennan
maantieteellistä aluetta perhosten
osalta. Korostan samalla, että arvi-

oon vaikuttavat edellä mainitut
subjektiiviset tekijät. Kovakuo-
riaisten osalta arvio perustuu jo jul-



Kuva 3 Kuvan 2 neighbor-joining-puussa esitettyjä perhoslajeja. — Lajit vasemmalta oikealle ja ylhäältä alaspäin: *Campaea margaritaria*, *Aspitates gilvaria*, *Hylaea fasciaria*, *Angerona prunaria*, *Cepphis advenaria*, *Abraxas sylvata*, *Abraxas grossulariata*, *Opisthagraptis luteolata*, *Deileptenia ribeata*, *Hypomecis punctinalis*, *Ematurga atomaria*, *Cleora cinctaria*, *Plagodis pulveraria*, *Plagodis dolabraria*, *Epirrhanthis diversata*, *Apeira syringaria*, *Colotois penaria*, *Alsophila aescularia*, *Ourapteryx sambucaria*, *Ennomos autumnaria*, *Ennomos fuscantaria*, *Ennomos alniaria*, *Lomographa temerata*, *Cabera pusa*, *Cabera exanthemata*, *Lomographa bimaculata*, *Ennomos erosaria*, *Charissa obscurata*, *Epione vespertaria*, *Epione repandaria*.

kaistuun tutkimukseen (Pentinsaari ym. 2014) ja sisältää 1 872 Suomessa tai lähialueella elävää lajia. Perhosten osalta käytän omaa pääosin julkaisematonta aineistoani, joka tässä analyysissä kattaa 2 573 lajia. Ulkomaisten kollegoideni kanssa olen tutkinut myös 122 allopatriisesti elävän alpiinis-arktisen perhoslajin DNA-viivakoodimuuteltua (Mutanen ym. 2012 b) sekä äskettäin noin 1 000 perhoslajin viivakoodimuuteltua Suomen ja Itävallan välillä (Huemer ym. 2014).

Yleensä suomalaiset perhoslajit voidaan yksiselitteisesti erottaa toi-

sistaan pelkän viivakoodin perusteella. Kuvassa 2 on neighbor-joining-menetelmällä kuvattu satunnaisesti valitun mittariperhosjoukon lajinsisäistä ja lajienvälistä muuntelua DNA-viivakoodissa (kuvia perhosista kuvassa 3). Kuvasta ilmenee hyvin se, että yleensä lajien väliset erot viivakoodissa ovat suuruusluokkaa suuremmat kuin yksilöiden väliset erot lajien sisällä. Vaihdellessa hieman arvioitavan ja käytetyn vertailuluokittelun mukaan tutkittujen 2 573:n Suomessa tavattavan perhoslajin joukosta löytyy 28 lajiparia tai lajiryhmää (lajiryhmiä mukana vain

kaksi), joissa DNA-viivakoodi ei aina erottele lajeja toisistaan. Tällöin määritysvarmuus olisi noin 97,5 %. Näistä 28 tapauksesta kuitenkin vain 15 lajiryhmän kohdalla kyse on oman arvioni mukaan kutakuinkin varmasti erillisistä lajeista. Useita näistä lajipareista tai -ryhmistä on jo muodollisestikin synonyisoitu (katsottu kuuluvaksi yhteen lajiin) eri yhteyksissä. Yhden tässä arvioissa mukana olevan lajiparin toisen osapuolen (*Schrankia intermedialis*; nimikin tarkoittaa välimuotoista) on todettu olevan kahden lajin risteymä. Mutta joukossa on myös kiistattomasti

erillisiä lajeja, mukaan lukien viiden lajin kompleksi *Apotomis*-kääriäissuvusta. Suku on tunnetusti hankala, ja yksi näistä viidestä lajista on ehkä toisen lajin synonyymi, mutta neljä muuta lajia eroavat toisistaan monessa suhteessa sekä morfologisesti että ekologisesti (kuva 4). Säännön mukaan viivakoodien erottamattomuus on yhteydessä ilmeiseen lähisukulaisuuteen ja morfologiseen samankaltaisuuteen. Vaikka onkin vaikeaa tietää, johtuuko erottamattomuus lajien nuoresta iästä vai mitokondrioiden introgressiosta (vai molemmista), uskon jälkimmäisen ilmiön olevan tavallisempi syy. Esimerkiksi kahden mittariperhoslajin (*Scopula ternata* ja *S. frigidaria*) viivakoodi on sama lähellä lajien päällekkäistä levinneisyysaluetta. Keski-Euroopassa, mutta myös Suomessa *S. ternata* -perhosissa tavataan toinen, ehkä alkuperäinen viivakoodityyppi. Tässä tapaukses-

sa siis *S. frigidaria* -mitokondrio olisi jossain vaiheessa ”hypännyt” *S. ternata* -lajiin. Mittarilajistossamme on muutamia muitakin samankaltaisia tapauksia. Edellä mainittujen lajiparien ja -ryhmien lisäksi Suomen lajistossa on noin 15 lajiparia, joiden viivakoodi on erilainen, mutta hyvin samankaltainen. Laajempi otanta voi osoittaa päällekkäisyyttä myös näiden lajien välillä. Perhoslajistossamme on lisäksi muutamia kymmeniä lajeja, joissa tavataan kahta huomattavasti (yli 2 %) toisistaan eroavaa haplotyyppiä. Jotkin näistä voivat ilmentää piilolajin olemassaoloa, toisissa tapauksissa taustalla voivat olla jotkin toistaiseksi tuntemattomat biologiset tekijät. Uuden sukupolven sekvensointitekniikka (ks. jäljempänä) tarjoaa erittäin lupavia uusia työkaluja tutkia tällaisia tapauksia tarkemmin.

Kovakuoriaisissa DNA-viivakoodit toimivat ehkä vielä perho-

siakin paremmin. Pentinsaaren ym. (2014) 1 872 lajia käsittävässä tutkimuksessa viivakoodit erottelivat 98,3 % lajeista mukaan lukien epäselvät lajiparit. Hyvä toimivuus kovakuoriaisissa on luultavasti yhteydessä siihen mielenkiintoiseen seikkaan, että lajien geneettiset erot viivakoodialueella näyttävät olevan kovakuoriaisissa keskimäärin selvästi suuremmat (lähes 12 %) kuin perhosissa (keskimäärin noin 6 %). Kuitenkin myös kovakuoriaisissa on lajipareja ja yksi kolmen lajin ryhmä, joissa viivakoodit eivät erottele lajeja toisistaan. Kuten perhosissa myös kovakuoriaisissa on useita tapauksia, joissa lajin sisältä löytyi yllättäen kaksi tai useampi erilainen haplotyyppi. Esimerkiksi varsin näyttävän aurorakuoriaisen (*Dictyoptera aurora*) voidaan lähes varmasti sanoa koostuvan kahdesta lajista, joissa todettiin huolellisessa tarkastelussa myös morfologisia eroja

Kuva 4 Neljä selvästi biologisen lajikäsityksen näkökulmasta erillistä *Apotomis*-suvun kääriäisperhoslajia, joita ei voi luotettavasti erottaa DNA-viivakoodien. — Taustalla voi olla lajienvälinen mitokondrioiden introgressio. Tämä on Suomen perhoslajistossa ainoa esimerkki useamman kuin kahden lajin DNA-viivakoodien samanlaisuudesta. Lajit vasemmalta oikealle ja ylhäältä alas: *Apotomis semifasciana*, *A. sororculana*, *A. betuletana* ja *A. inundana*.





Marko Mutanen

Kuva 5 DNA-viivakooditus paljasti yllättäen, että aurorakuoriainen (*Dictyoptera aurora*) käsittää kaksi hyvin läheistä laji, joiden myöhemmissä tutkimuksissa on havaittu eroavan myös morfologisesti. Kuvassa molemmat taksonit.

(kuva 5). Nämä tapaukset osoittavat oivallisesti sen, että DNA-viivakoodit todella paljastavat aiemmin kokonaan huomaamatta jääneitä lajeja jopa hyvin tutkituista lajiryhmistä.

Mitä tapahtuu, kun maantieteellistä aluetta laajennetaan? Asiaa on tutkittu vielä melko vähän. Yksi kattavimmista on tekemämme vertailu 1 004 perhoslajista, jotka elävät sekä Suomessa että Itävallassa (Huemer ym. 2014). Tutkimuksessa havaittiin, että vaikka laajempi maantieteellinen otanta oletusten mukaisesti kasvattaa lajinsisäistä muuntelua viivakoodialueella, peräti 98,8 %:ssa tutkituista lajeista DNA-viivakoodi on lajille tyypillinen. 124 lajissa kokonaismuuntelu ylittää 2 %:n arvon, mikä on lajinsisäiseksi muunteluksi varsin paljon. Joukossa on melko varmasti lajeja, jotka itse asiassa sisältävät kaksi laji. Hiukan yllättäen erot eivät ole tilastollisesti suurempia sellaisissa lajeissa, joiden levinneisyys on epäyhtenäinen, kuin sellai-

sis, joiden levinneisyys jatkuu yhtenäisenä Suomesta Itävaltaan. Tunnettujen vaeltajien välillä erot ovat pienimpiä, mikä tukee sitä ajatusta, että vähemmän levittäytyvien eläinlajien alueelliset erot korostuvat. Vaikka levinneisyystyypillä ei havaittukaan olevan yhteyttä maantieteelliseen muunteluun, muuntelu on suurempaa sellaisissa lajiryhmissä, joiden oletimme olevan huonoja leviäjiä. Toisessa tutkimuksessa selvitimme viivakoodieroja 122:sta puhtaasti alpiinis-arktisesta lajista, jotka siis elävät sekä Alppien ympäristössä että arktisella alueella, mutta puuttuvat tältä väliltä (Mutanen ym. 2012 b). Tässä tutkimuksessa korostuivat allopatriaan liittyvät ongelmat. DNA-viivakoodien toimivuutta on näissä olosuhteissa hyvin vaikea tutkia objektiivisesti, koska allopatristen populaatioiden lajinrajaus on erityisen hankalaa ja subjektiivista, jopa vailla pelisääntöjä. Tutkimuksessa käy selvästi ilmi, että joskus hyvinkin samankaltaiset populaa-

tiot on katsottu kuuluviksi eri lajeihin, kun taas toisinaan selvästi erilaistuneet populaatiot katsotaan kuuluviksi samaan lajiin. Seuraukset voivat olla merkittävät. Esimerkiksi omaksi lajikseen katsottu, Pohjois-Skandinaviassa endeeminen ruijannokiperhonen (*Erebia polaris*) on saanut EU-direktiivilajiin aseman. Laji on kuitenkin sekä geneettisesti että morfologisesti käytännöllisesti katsoen identtinen monilla vuoristoalueilla, mm. Alpeilla, elävän *Erebia medusa* -lajin kanssa.

DNA-pohjainen lajinrajaus

Tunnetaan toinen toistaan ongelmallisempia lajiryhmiä, joissa lajien rajaaminen ei ole helppoa. Tähän vaikuttaa suuresti lajikäsityksen vaihtelu eli se, että on olemassa monia tapoja määritellä laji. Kun mukaan liitetään maantiede, asiat mutkistuvat entisestään, koska lajeilla on taipumusta sopeutua eri alueiden erilaisiin olosuhteisiin ja koska geenivirta on populaatioiden välillä eri syistä monesti hidasta. Erityisen hankalaa on sellaisten lajien tai populaatioiden luokittelu, joiden välillä ei ole ollut geenivirtaa ollenkaan pitkään aikaa tai sitä on ollut hyvin vähän maantieteellisen eristäytymisen vuoksi. Tällöin lajit elävät allopatristesti (vastakohta on sympatria, jossa lajit elävät samalla alueella) ja mm. lisääntymisisolaation tutkiminen on vaikeaa tai ainakin epäkäytännöllistä.

DNA-pohjaisia lajinrajausmenetelmiä on kehitelty viime vuosina innokkaasti. Osaksi tämä on johtunut laajojen DNA-viivakoodikirjastojen synnystä, mutta vielä enemmän siitä, että geneettistä aineistoa on nykyään helposti saatavissa tai kerättävissä. Jotkin lajinrajausmenetelmät, etenkin ABGD (< engl. *automated barcode gap discovery*; Puillandre ym. 2012) ja ReSL (< engl. *refined single linkage*; Ratnasingham & Hebert 2013), on erityisesti suunniteltu rajaamaan DNA-viivakoodisekvenssien perusteella ns. toiminnallisia lajiyksiköitä (OTU, < engl. *operational taxonomic unit*). Uusia lajeja todetaan koko ajan paljon nopeammin kuin niitä ehditään kuvata, sil-

lä lajinkuvaus on usein valitettavan hidas ja kankea prosessi, joten toiminnallisten lajiyksiköiden merkitys kasvaa. Niinpä on pohdittu, tarvitsevaiko OTUt standardoidun nimistön (Schindel & Miller 2009; artikkelissa käytetään kuvaamattomasta lajista nimitystä *taxon label*). ReSL-menetelmällä luodut OTUt saavat ns. BIN-tunnukset (< engl. *barcode index number*); niistä on tulossa DNA-viivakoodien havaittujen OTUjen tilapäisnimistön standardi. OTUja ei ole tarkoitettu sellaisinaan kuvattaviksi lajeiksi, vaan lähinnä aputyökaluiksi lopulliseen lajinrajaukseen, jonka tulisi hyödyntää myös muuta tietoa kuvattavien lajien morfologiasta ja mahdollisuuksien mukaan biologiasta, fenologiasta ym. Toiset menetelmät, esimerkiksi GMYC (< engl. *general mixed Yule-coalescence*; Pons ym. 2006) ja PTP (< engl. *Poisson tree processes*; Zhang ym. 2013), eivät kohdistu tiettyyn markkeriin, eli ne voivat käsitellä monen eri geenin tietoja yhtä aikaa ja periaatteessa mahdollistavat täten jopa yksinomaan DNA-aineistoon perustuvan lajinrajauksen. Viime vuosina on tutkittu lajinrajausmenetelmien ominaisuuksia ja eroja (mm. Kekkonen & Hebert 2014). Menetelmät tuottavat yleensä hiukan toisistaan eroavia, mutta pääpiirteissään samankaltaisia luokitteluja. Esimerkiksi BIN-menetelmällä pelkästään DNA-viivakoodeihin perustuva lajinrajaus tuottaa lopputuloksen, joka on yllättävän lähellä perinteisin menetelmin rajattuja lajeja (Ratnasingham & Hebert 2013) osoittaen DNA-viivakoodien informaatiivisuuden myös lajinrajausta tehdessä. Todennäköisesti tulevaisuudessa näemme enenevässä määrin lajinkuvauksia, jotka perustuvat kokonaan DNA-sekvenssitietoon. Jo tällä hetkellä joissakin ryhmissä, erityisesti prokaryoottisissa eliöissä, lajinrajaus perustuu lähes yksinomaan DNA-sekvensseihin. Ei ole mitään teoreettista perustetta sille, miksei myös vaikkapa hyönteislajin kuvaus voisi perustua kokonaan DNA-sekvenssitietoon, erityisesti jos tämä tieto on kattavaa, ei yhteen DNA-alueeseen perustuvaa.

Uuden sekvensointitekniikan mahdollisuudet

Frederick Sangerin ryhmän kehittämä periaate DNA-sekvenssien emäsjärjestyksen lukemiseksi pysyi vallitsevana yli kaksi vuosikymmentä ja on edelleen laajasti käytössä. Menetelmä on toki kehittynyt ajan saatossa koko ajan. Sangerin sekvensointiin perustuvien menetelmien laajan genomiaineiston kokoaminen on hidasta, koska menetelmä vaatii yhä melko paljon käsityötä ja perustuu laajasti yksittäisten näytteiden ja analyysiin valittujen 500–800 emäsparin geenimarkkerien monistamiseen polymeraasiketjureaktion avulla. DNA:n monistaminen ja sekvensointi ovat Sangerin sekvenssoinnissa erillisiä tapahtumia. Yhdellä sekvensointiajolla saatavien sekvenssien määrä on kymmeniä tai enintään satoja, ja vuorokauden aikana menetelmällä voidaan lukea maksimissaan noin 115 000 emäsparia DNA-sekvenssiä (Mardis 2011). Näin ollen menetelmä ei sovellu nopeaan ja laajaan koko genomien analysointiin. Esimerkiksi ihmisen koko tuman genomi käsittelee noin kolme miljardia emäsparia ja noin 20 000–25 000 geeniä.

Vuodesta 2005 lähtien käyttöön on tullut kokonaan uudenlaisia menetelmiä, joita yhdessä nimitetään uuden sukupolven sekvenssoinniksi (engl. *next-generation sequencing*, NGS). Näillä menetelmillä voidaan lukea DNA-sekvenssiä monta suuruusluokkaa nopeammin kuin Sangerin menetelmällä. Menetelmät ovat nopeasti mullistamassa kaikkea genomitutkimusta, ja menetelmien kehittyminen on ollut huikaisevan nopeaa. Samalla sekvenssitiedon tuottamisen hinta emästä kohden on tippunut murto-osaan aiemmasta. Hinnanlaskuun on vaikuttanut suuresti myös se, että toisistaan eroavia uuden sukupolven sekvenssointialustoja (engl. *platform*) on useita erilaisia, mikä on lisännyt kilpailua valmistajien välillä.

En käy tässä kirjoituksessa läpi eri menetelmien yksityiskohtia. Aiheesta on kirjoitettu useita artikkeleita (Metzker 2010; Mardis 2011, 2013). Oleellisin ero Sangerin sek-

vensointiin on se, että näillä menetelmillä DNA:n monistus ja lukeminen tapahtuvat samanaikaisesti ja yhtä aikaa voidaan lukea miljoonia DNA-fragmentteja (jotka kuitenkin ovat keskimäärin hiukan lyhyempiä kuin Sangerin sekvenssit). Tehokkaimmilla tällä hetkellä käytössä olevilla menetelmillä voidaan yhdessä ajossa lukea yli biljoona emäsparia ja tuottaa satoja miljoonia tai jopa yli miljardi erillistä sekvenssiä. Näin ollen vaikkapa ihmisen koko genomien lukeminen lyhyessä ajassa on tullut mahdolliseksi, mikä on synnyttänyt kaupalliset markkinat mm. perinnöllisten sairauksien tunnistamiselle. Tällä hetkellä oman genomien lukemisen voi teettää kaupallisesti noin 1 000 eurolla.

Taksonomiassa ja lajien tunnistamisessa uusilla menetelmillä on luonnollisesti valtavat mahdollisuudet. Menetelmillä voidaan saada laaja genomikirjasto yksittäisestä yksilöstä tai lajista, mikä mahdollistaa eliöryhmien evoluutiohistorian (fylogenetian) tutkimisen kattavan genomiaineiston perusteella. Tällaisia satojen geenien informaatioon perustuvia tutkimuksia on jo julkaistu (esimerkiksi Kawahara & Breinholt 2014). Menetelmillä voidaan myös tuottaa sekvenssitietoa kohdistetusti vaikkapa DNA-viivakoodijaksosta suuresta yksilömäärästä yhtäaikaisesti. Tätä uutta tieteenhaaraa kutsutaan metagenomiikaksi ja sen soveltamista systematiikkaan metasytematiikaksi ja DNA-viivakoodien avulla lajien tunnistamiseen metaviivakooditukseksi (engl. *meta-barcoding*). Metaviivakoodituksella on lukuisia sovelluskohteita, ja menetelmää hyödyntäviä artikkeleita julkaistaan tätä nykyä runsaasti. Sitä on käytetty esimerkiksi ihmisen suoliston mikrobidiversiteetin ja maaperän sieniyhteisöjen tutkimisessa. Metagenomiikan merkitys tulee olemaan erityisen suuri yhteisöekologisissa tutkimuksissa. Siinä, missä aiemmin runsasyksilöisten ja -lajisten yhteisöjen tutkiminen on ollut mahdotonta tai ainakin vaikeaa, epätarkkaa ja työlästä, voidaan lajikoostumusta tutkia metagenomiikan ja uuden sukupolven sekvenssoinnin avulla paljon aiem-

paa laajemmin, tehokkaammin ja halvemmalla. Menetelmä mahdollistaa esimerkiksi laajan hyönteisaineiston yksilöiden samanaikaisen määrittämisen viivakoodeja hyväksi käyttäen ilman asiantuntijoita, mikäli käytettävissä on kattava viivakoodien vertailukirjasto. Mikäli referenssikirjastoa ei ole, tutkimuskysymysten selvittämiseen voidaan käyttää toiminnallisia lajiyksiköitä. Menetelmällä saadaan tällä hetkellä vain karkea arvio lajien runsaudesta aineistossa, mutta luultavasti myös yksilömäärän tarkempi arviointi on tulevaisuudessa mahdollista. Vastikään julkaistu Gibsonin ym. (2014) tutkimus on hyvä esimerkki metaviivakoodituksen mahdollisuuksista. Tutkijat analysoivat siinä yhtaikaisesti yli tuhat monipuolisesti lentäviä hyönteisiä pyydystävällä ns. Malaisen pyydyskällä kerättyä yksilöä ja sekvensoivat ne sekä Sangerin menetelmällä että uuden sukupolven sekvensointimenetelmin. Sangerin sekvensoinnilla todetuista lajeista löytyi metaviivakooditusmenetelmällä noin 96 %. Tulokseen vaikutti huomattavasti ennen kaikkea erilaisten alukkeiden määrä, mikä osoittaa, että eri ryhmille soveltuvien alukkeiden kehittämisellä ja niiden lisäämisellä päästään lopulta erinomaiseen lopputulokseen.

Uuden sukupolven sekvensointimenetelmät kehittyvät edelleen nopeasti, ja jää nähtäväksi, mitä ne mahdollistavat vaikkapa kymmen vuoden kuluttua. Se on jo nyt selvää, että ne tulevat mullistamaan kaiken systematiikan alan tutkimuksen fylogenetikasta lajien määrittämiseen ja niiden rajaukseen.

Kiitokset: Kiitän Mikko Pentinsaarta käsikirjoituksen kommentoinnista sekä luvasta käyttää aurorakuoriaista erimerkkinä kryptisestä lajiparista.

Kirjallisuus

- Gibson, J., Shokralla, S., Porter, T. M., King, I., Konyonenburg, S. van, Janzen, D. H., Hallwachs, W. & Hajibabaei, M. 2014: Simultaneous assessment of the macrobiome and microbiome in a bulk sample of tropical arthropods through DNA metasytematics. — *Proc. National Acad. Sci. USA* 111: 8007–8012.
- Hajibabaei, M. 2012: The golden age of DNA metasytematics. — *Trends Genet.* 28: 535–537.
- Hanski, I., Herten, L. von, Fyhrquist, N., Koskinen, K., Torppa, K., Laatikainen, T., Karisola, P., Auvinen, P., Paulin, L., Mäkelä, M. J., Vartiainen, E., Kosunen, T. U., Alenius, H. & Haahtela, T. 2012: Environmental biodiversity, human microbiota, and allergy are interrelated. — *Proc. National Acad. Sci. USA* 109: 8334–8339.
- Hebert, P. D. N., Cywinska, A., Ball, S. L. & deWaard, J. R. 2003: Biological identifications through DNA barcodes. — *Proc. R. Soc. London B* 270: 313–321.
- Hebert, P. D. N., Penton, E. H., Burns, J. M., Janzen, D. H. & Hallwachs, W. 2004: Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astraptes fulgerator*. — *Proc. National Acad. Sci. USA* 101: 14812–14817.
- Huemer, P., Mutanen, M., Sefcs, K. & Hebert, P. D. N. 2014: Testing DNA barcode performance in 1000 species of European Lepidoptera: large geographic distances have small genetic impacts. — *PLoS ONE* 9: e115774 (doi: 10.1371/journal.pone.0115774).
- Hurst, G. D. D. & Jiggins, F. M. 2005: Problems with mitochondrial DNA as a marker in population, phylogeographic and phylogenetic studies: the effects of inherited symbionts. — *Proc. R. Soc. London B* 272: 1525–1534.
- Kawahara, A. Y. & Breinholt, J. W. 2014: Phylogenomics provides strong evidence for relationships of butterflies and moths. — *Proc. R. Soc. London B* 281: 20140970 (doi: 10.1098/rspb.2014.0970).
- Kekkonen, M. & Hebert, P. D. N. 2014: DNA barcode-based delineation of putative species: efficient start for taxonomic workflows. — *Molec. Ecol. Resources* 14: 716–715 (doi: 10.1111/1755-0998.12233).
- Korpeläinen, H. 2008: Geneettiset viivakoodit tunnistamisen avuksi. — *Luonnon Tutkija* 112: 3.
- Korpeläinen, H. 2010: DNA-lajitunnisteilla tarkkuutta lajinmäärittelyyn. — *Luonnon Tutkija* 114: 174–177.
- Korpeläinen, H. 2013 a: Suomen kasvien DNA-viivakoodikirjasto. — *Luonnon Tutkija* 117: 132–134.
- Korpeläinen, H. 2013 b: What is in the air? Using DNA barcodes to detect indoor fungal composition. — *Barcode Bull.* 4: 19.
- Mardis, E. R. 2011: A decade's perspective on DNA sequencing technology. — *Nature* 470: 198–203.
- Mardis, E. R. 2013: Next-generation sequencing platforms. — *Annual Rev. Anal. Chem.* 6: 287–303.
- Metzker, M. L. 2010: Sequencing technologies — the next generation. — *Nat. Rev. Genet.* 11: 31–46 (doi: 10.1038/nrg2626).
- Meier, R., Shiyang, K., Vaidya, G. & Ng, P. K. 2006: DNA barcoding and taxonomy in Diptera: a tale of high intraspecific variability and low identification success. — *Syst. Biol.* 55: 715–728.
- Mutanen, M., Aarvik, L., Landry, J.-F., Segerer, A. & Karsholt, O. 2012 a: *Epinoia cinerea* (Haworth, 1811) bona sp., a Holarctic tortricid distinct from *E. nisella* (Clerck, 1759) (Lepidoptera: Tortricidae: Eucosmini) as evidenced by DNA barcodes, morphology and life history. — *Zootaxa* 3318: 1–25.
- Mutanen, M., Hausmann, A., Hebert, P. D. N., Landry, J.-F., Waard, J. R. de & Huemer, P. 2012 b: Allopatry as a Gordian knot for taxonomists: patterns of barcode divergences in arctic-alpine Lepidoptera. — *PLoS ONE* 7: e47214 (doi: 10.1371/journal.pone.0047214).
- Pentinsaari, M., Hebert, P. D. N. & Mutanen, M. 2014: Barcoding beetles: a regional survey of 1872 species reveals high identification success and unusually deep interspecific divergences. — *PLoS ONE* 9: e108651 (doi: 10.1371/journal.pone.0108651).
- Pons, J., Barraclough, T. G., Gomez-Zurita, J., Cardoso, A., Duran, D. P., Hazell, S., Kamoun, S., Sumlin, W. D. & Vogler, A. P. 2006: Sequence-based species delimitation for the DNA taxonomy of undescribed insects. — *Syst. Biol.* 55: 595–609.
- Puillandre, N., Lambert, A., Brouillet, S. & Achaz, G. 2012: ABGD, automatic barcode gap discovery for primary species delimitation. — *Molec. Ecol.* 21: 1864–1877.
- Ratnasingham, S. & Hebert, P. D. N. 2013: A DNA-based registry for all animal species: the barcode index number (BIN) system. — *PLoS ONE* 8: e66213 (doi: 10.1371/journal.pone.0066213).
- Rougerie, R., Smith, M. A., Fernandez-Triana, J., Lopez-Vaamonde, C., Ratnasingham, S. & Hebert, P. D. N. 2011: Molecular analysis of parasitoid linkages (MAPL): gut contents of adult parasitoid wasps reveal larval host. — *Molec. Ecol.* 20: 79–186.
- Sanger, F., Nicklen, S. & Coulson, A. R. 1977: DNA sequencing with chain-terminating inhibitors. — *Proc. National Acad. Sci. USA* 74: 5463–5467.
- Schindel, D. E. & Miller, S. E. 2009: Provisional nomenclature. The on-ramp to taxonomic names. — *Teoksessa: Polaszek, A. (toim.), Systema naturae 250. The Linnaean ark: 109–115.* 300 s. CRC Press. Boca Raton (FL).
- Schmidt, B. C. & Sperling, F. A. H. 2008: Widespread decoupling of mtDNA variation and species integrity in *Grammia* tiger moths (Lepidoptera: Noctuidae). — *Syst. Entomol.* 33: 613–634.
- Tautz, D., Arctander, P., Minelli, A., Thomas, R. H. & Vogler, A. P. 2003: A plea for DNA taxonomy. — *Trends Ecol. Evol.* 18: 70–74.
- Watson, J. D. & Crick, F. H. C. 1953: Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid. — *Nature* 171: 737–738.
- Zhang, J., Kapli, P., Pavlidis, P. & Stamatakis, A. 2013: A general species delimitation method with applications to phylogenetic placements. — *Bioinformatics* 29: 2869–2876.

Kirjoittaja

Marko Mutanen, biologian laitos, biodiversiteettiyksikkö, PL 3000, 90014 OULUN YLIOPISTO; marko.mutanen@oulu.fi

FT Marko Mutanen toimii yli-intendentinä Oulun yliopiston biodiversiteettiyksikössä. Hänen tutkimusalaan on molekyyliatonomi ja systematiikka. Hän johtaa Suomen Akatemian rahoittamaa tutkimusryhmää, joka keskittyy hyönteisten DNA-pohjaiseen lajitunnistukseen ja lajinrajaukseen sekä perhosten fylogenetikkaan. Hän myös koordinoi kansallista FinBOL-hanketta, joka pyrkii tuottamaan DNA-viivakoodivertailukirjaston koko Suomen eläölajistosta.