

Janna Saarela ja Kaisa Kettunen

## Kenelle eksomisekvensointi?

Uuden sukupolven sekvensointimenetelmät ovat mullistaneet mahdollisuutemme analysoida ihmisen perimää. Genomin proteiinia koodaavan alueen eli eksomin sekvensointia hyödynnetään tutkimuksen lisäksi myös kliinisessä lääketieteessä, erityisesti harvinaisten periytyvien tautien, synnynnäisten kehityshäiriöiden ja syövän diagnostiikassa. Diagnostisena tutkimuksena eksomisekvensointi on perusteltua silloin, kun mahdollinen tulos voi johtaa täsmälliseen diagnoosiin, vaikuttaa merkittävästi sairauden hoitoon ja ohjata potilaan perinnöllisyysneuvonnan piiriin. Menetelmä on luotettava yhden emäksen muutosten ja pienten deleetioiden ja insertioiden havaitsemiseen. Sitä vastoin suurempia kopiolumuutoksia, muutoksia toistojaksojen pituudessa ja alueita, joista on useita kopioita genomissa, on nyky menetelmillä vaativalta analysoida luotettavasti. Koska kaikkien ihmisten eksomissa on runsaasti muutoksia, oleellinen haaste eksomisekvensointitulosten tulkinassa on erottaa merkityksettömät muutokset sairauden aiheuttavista.

**V**iime vuosien nopean sekvensointitekniologian kehittymisen ansiosta nykyisin pystytään tutkimaan yksittäisen henkilön koko genomi muutamassa päivässä vain murto-osan hinnalla siitä, mitä vuonna 2001 valmistunut ensimmäisen ihmisen genomin sekvensointi maksoi. Tautigeenien tunnistaminen tehostui entisestään vuonna 2009, kun markkinoille tuotiin genomin proteiineja koodaavien alueiden eli eksomien rikastamisen mahdollistavia reagenssipaketteja. Näiden avulla sekvensointi voitiin kohdistaa siihen genomin osaan, jossa tauteja aiheuttavat geenivirheet todennäköisimmin sijaitsevat. Eksomisekvensointia on hyödynnetty tehokkaasti yhden geenin aiheuttamien tautien geenivirheiden etsinnässä (1). Viime vuosina tätä teknologiaa on alettu hyödyntää tutkimuksen lisäksi myös kliinisessä lääketieteessä, erityisesti harvinaisten, periytyvien tautien, synnynnäisten kehityshäiriöiden ja syövän diagnostiikassa.

Eksomisekvensointi on tehokas ja luotettava keino tunnistaa, mitkä emäkset henkilön genomin proteiineja koodaavissa alueissa eli eksomeissa poikkeavat maailmanlaajuisesti käytetystä verrokkigenomista. Uusimmat koodaavien alueiden rikastamismenetelmät kohdis-

tuvat noin 180 000 yksittäiseen proteiinia koodaavaan eksoniin keskimäärin 23 000 geenissä ja mahdollistavat lähes kaikkien (95 %) yhden nukleotidin muutosten ja pienten insertioiden ja deleetioiden (alle 10 nukleotidia) tunnistamisen. Haasteita eksomisekvensoinnille luovat alueet, joissa esiintyy toistojaksoja tai useita geenikopioita tai jotka on huonosti kartoitettu verrokkigenomissa. Positiivinen eksomisekvensointitulokset on hyvin tarkka, mutta menetelmän herkkyys vaihtelee eri genomialueilla, minkä vuoksi negatiivinen tulos ei sulje pois geneettistä syytä sairauden taustalla.

Epidemiologisten tutkimusten mukaan jopa 8 %:lla väestöstä ilmenee jokin geneettinen sairaus ennen aikuisikää, kun kaikki synnynnäiset poikkeavuudet otetaan huomioon (2). Kun tunnetut harvinaiset geneettiset sairaudet lasketaan yhteen, ne vaikuttavat merkittävään osaan ihmisistä, ja osa yleisistäkin sairauksista saattaa syntyä yksittäisen geenin virheestä.

### Kenelle eksomitutkimus?

Kustannusvaikuttavuutensa vuoksi kliinisen eksomisekvensoinnin tulisi olla ensivaiheen diagnostinen työkalu tutkittaessa potilaita,

joilla epäillään yhden geenin aiheuttamaa geneettistä sairautta, jonka taustalla tunnetaan useita eri geenejä. Mikäli potilaan taudinkuva on tyypillinen jollekin tietylle taudille tai tautiryhmälle, jonka geneettinen tausta on hyvin kartoitettu, ovat kohdennettu yksittäisen geenin analyysi tai kattava geenipaneeli nykyään vielä edullisin diagnostinen tutkimus. Useita tai pitkiä toistojaksoja sisältävien tautigeenien, kuten lapsuusiän myopatiaa aiheuttavan *NEB*-geenin, osalta kohdennettu geenitutkimus on vielä ainoa vaihtoehto. Niin sanotut kolmannen sukupolven sekvensointilaitteet, jotka pystyvät lukemaan hyvin pitkiä, yksittäisiä DNA-molekyylejä, voivat jatkossa ratkaista toistojaksojen aiheuttamat ongelmat.

Kaikenikäiset potilaat voivat hyötyä eksomisekvensointitutkimuksesta, mutta erityisen hyödyllinen se on pienten lasten epäselvien sairauksien diagnostiikassa, koska valtaosa geneettisistä sairauksista ilmenee lapsuudessa. Tavanomaisten diagnostisten menetelmien hyödyntämistä vaikeuttaa taudin oireiden vain osittainen ilmentyminen. Varhainen molekyyli-diagnoosi mahdollistaa kalliiden jatkotutkimusten välttämisen, taudinkulun paremman ennustamisen, perinnöllisyysneuvonnan, tautiriskin arvioinnin perheessä, sikiö- ja perinataali-diagnostiikan ja perhesuunnittelun sekä parhaissa tapauksissa sairauden täsmällisemmän ja tehokkaamman hoidon.

Aikuispotilaiden eksomisekvensointi on perusteltua, jos tulos voi johtaa täsmälliseen diagnoosiin ja vaikuttaa merkittävästi sairauden hoitoon sekä auttaa ohjaamaan potilaan perinnöllisyysneuvontaan. Yllättävän moni aikuisiässä esiintyvistä harvinaisista sairauksista on taustaltaan geneettinen. Tämä on hyvä pitää mielessä erityisesti, jos potilaan vaste sairauden tavanomaiseen hoitoon on huono. Diagnoosin varmistumisen suurimpia etuja ovat taudinkulun parempi ennustettavuus, oireettomien lähisukulaisten geneettisen riskin kartoittamisen mahdollisuus ja ennakoiva tai varhainen taudin hoito taikka seurannan tehostaminen. Molekyyli-diagnoosi voi mahdollistaa myös aikuispotilaan täsmähoidon ja kaikenikäisille pääsyn mukaan tautiin kohdennetun lääkkeen lääketutkimukseen.

## **Eksomitutkimuksia tulevaisuudessa lähes joka erikoisalalla**

Kliinisiä eksomitutkimuksia tilannevat yhä enemmän perinnöllisyys- ja lastenlääkäreiden lisäksi muutkin lääkärit. Kullekin erikoisalalle relevantin genetiikan tulisi kuulua opetussuunnitelmaan lääketieteen perusopetuksessa ja erikoistumisopinnoissa. **TAULUKOSSA 1** kuvataan asiat, jotka eksomitutkimusta tilaavan lääkärin tulisi huomioida. Potilaan ja perheen kanssa keskusteltaessa on tärkeää sopia, mitä potilaalle kerrotaan tutkimustuloksista. American College of Medical Genetics and Genomics (ACMG) suosittelee, että kaikki lääketieteellisesti hoidettavissa olevia sairauksia aiheuttavat muutokset tulisi raportoida, vaikka sairaus ei olisi vielä ilmennytäkään (3). Sattumalöydösten tulkinta ei kuitenkaan ole yksiselitteistä (tarkemmin jäljempänä) vaan vaatii mittavan lisätyön, perehtyneisyyttä useiden eri tautiryhmien diagnostiikkaan ja väestökohtaista tutkimustietoa, jota ei vielä ole kattavasti saatavilla. Tämän vuoksi usein käytetty vaihtoehto on vain tutkittavaan sairauteen todennäköisesti vaikuttavien muutosten raportointi, mikäli potilas ei erityisesti halua kuulla muistakin mahdollisista löydöksistä, esimerkiksi jos hänen suvussaan esiintyy syöpää.

## **Eksomidatan tulkinta**

Aluksi potilaan emäsjärjestystä verrataan verrokkigenomiin hyödyntämällä sekvensointilaitetuottajien ohjelmistoja ja kansainvälisen tiedeyhteisön kehittämiä algoritmeja. Jokaisen henkilön genomissa on vajaat neljä miljoonaa muunnosta eli varianttia nykyisin käytettävään verrokkigenomiin verrattuna, joten oleellinen haaste genomitiedon tulkinnessa on harmittomien varianttien erottaminen sairauden aiheuttavista muutoksista. Tunnistettujen varianttien tarkka genominen sijainti ja niiden mahdollisesti aikaansaama vaikutus määritetään hyödyntämällä julkisesti saatavissa olevia, kaupallisia tai laboratorioissa kehitettyjä analyysialgoritmeja.

Varianteista noin 10 000–15 000 on proteiinien koodia muuttavia yhden nukleotidin muu-

toksia, joilla voi olla biologinen vaikutus kyseisen proteiinin toimintaan. Muutosten biologisten vaikutusten ennustamiseen on olemassa useita julkisia bioinformatiikan työkaluja, jotka ennenaikaisia lopetuskodoneita, lukukehysmuutoksia ja silmukointivirheitä lukuun ottamatta antavat usein keskenään ristiriitaisia ennusteita. Mikään yksittäinen työkalu ei ole merkittävästi muita täsmällisempi (4). Keskimäärin 2–3 %:n tunnistetuista sekvenssimuutoksista ennustetaan olevan haitallisia proteiinin toiminnalle, joten kunkin henkilön eksomissa on satoja tai jopa tuhansia mahdollisesti haitallisia variantteja. Jopa 100–300 näistä varianteista on proteiinin normaalin toiminnan estäviä loss-of-function-muutoksia. Näistä 10–20 esiintyy kunkin ihmisen genomissa homotsygoottisena (molemmat geenikopiot ovat viallisia), jolloin kyseinen proteiini ei toimi lainkaan (5).

Tietyn proteiinin toiminnan puuttuminen ei kuitenkaan välttämättä ole vahingollista terveydelle. Suomalaisväestössä on enemmän loss-of-function-variantteja kuin heterogeenisemmissä väestöissä. Nykyarvioiden mukaan kunkin henkilön perimässä on 50–100 varianttia, joiden on raportoitu liittyvän johonkin sairauteen, sekä noin 60 uutta (de novo) varianttia, joita ei esiinny potilaan vanhemmilla (6). Uudet variantit ovat väestötasolla yleensä hyvin harvinaisia, usein vain yhdellä ihmisellä esiintyviä, ja niiden oletetaan olevan muita vahingollisempia, koska niihin ei ole kohdistunut luonnonvalintaa.

Kliinisessä eksomitiedon tulkinnassa pyritään tunnistamaan yksilön genomista todennäköisimmin haitalliset muutokset suodattamalla löydetty variantit niiden mahdollisen vaikutuksen, yleisyyden ja taudin periytymisen mallin mukaan hyödyntämällä kaupallisia ja laboratorioissa kehitettyjä bioinformatiikan työkaluja ja mutaatiotietokantoja, kuten ClinVar, Human Genome Mutation Database (HGMD) ja Online Mendelian Inheritance in Man (OMIM). Lopullisen arvion tunnistettujen mutaatioiden vaikutuksesta tekee aina genetiikan asiantuntija tietokantoja ja kirjallisuutta hyödyntäen. Useimmissa laboratorioissa tulkinnan käy vielä läpi laboratorioasiantuntijoista, geneetikoista, bioinformatikosta ja lääkäreistä koostuva työ-

**TAULUKKO 1.** Asiat, jotka eksomitutkimusta tilaavan lääkärin tulisi selvittää.

1. Tutkimuksen kuvaaminen potilaalle tai perheelle, ja jos potilas haluaa tutkimusta, tämän tiedon kirjaaminen sairaskertomukseen.
2. Potilaan taudinkuvan laaja-alainen kartoittaminen, myös oman erikoisalun ulkopuolisten oireiden ja löydöksiin.
3. Tiedon kerääminen potilaan perheestä ja suvusta, erityisesti potilaan taudinkuvan kannalta oleellisista oireista.
4. Tiedonhaku kirjallisuudesta ja tietokannoista tai esimerkiksi perinnöllisyyslääkärin, harvinaistautien yksikön tai eksomianalyysipalveluita tuottavan laboratorion asiantuntijan konsultointi.

ryhmä, ja tautia aiheuttavaksi arvioitu mutaatio varmennetaan potilaan näytteestä myös toisella menetelmällä.

Vaikka eksomidatan tuottaminen on osavassa laboratorioissa nopea ja toistettavissa oleva menetelmä, vajavainen nykytietämys ihmisen genomissa olevien varianttien merkityksestä tekee eksomidatan tulkinnasta vaativaa. Kuinka tunnistaa potilaan taudin aiheuttava mutaatio mahdollisesti vahingollisten mutaatioiden suuresta joukosta? Kliinisten eksomien mutaatioanalyysi kohdistetaan nykyään pääasiassa proteiinia koodaavien ja RNA:n silmukointiin vaikuttavien alueiden toiminnallisiksi ennustettuihin variantteihin, ja ei-koodaavat, geenien toimintaa säätelevät alueet jätetään pääosin analysoimatta, vaikka sekvenssianalyysi osan niistä kattaisikin. Tunnetuista tautia aiheuttavista geenivirheistä 59 % aiheuttaa ennenaikaisen lopetuskodonin, saa aikaan virheellisen lähetti-RNA:n silmukoinnin tai muuttaa lajien välillä konservoituneen aminohapon toiseksi proteiinia koodaavassa geenissä (7).

Yksi selkeä ja analyysityötä nopeuttava tapa on kohdistaa mutaatioanalyysi ensisijaisesti aiemmin kuvattuihin, potilaan taudinkuvaan sopiviin geeneihin (niin sanottu virtuaaligeeni-paneeli). Helpoimmin tulkittavia ovat aiemmin tautia aiheuttaviksi kuvatut mutaatiot tunnetuissa tautigeeneissä, joskin niidenkin tulkinta vaatii erityisosaamista aiemmassa kirjallisuudessa virheellisesti tautia aiheuttaviksi raportoitujen mutaatioiden vuoksi. Kirjallisuudessa

### Ydinasiat

- ▶ Eksomisekvensoinnilla pystytään luotettavasti ja kustannusvaikuttavasti tunnistamaan pääosa genomimme proteiineja koodaavan alueen vaihtelusta.
- ▶ Jokaisen ihmisen genomissa esiintyy vajoja neljä miljoonaa eroavaisuutta, joista 100–300 mahdollisesti estää proteiinin normaalin toiminnan.
- ▶ Keskeisin haaste on tulkintavaihe eli sairautta aiheuttavien muutosten tunnistaminen lukuisten mahdollisesti biologisesti merkittävien muutosten joukosta.
- ▶ Negatiivinen tulos ei sulje pois geneettistä syytä sairauden taustalla, koska kaikkia genomien alueita ja varianttityyppejä ei saada analysoiduksi luotettavasti.
- ▶ Genomitiedon ja väestökohtaisen verrokkidatan lisääntyminen helpottaa tulevaisuudessa eksomidatan tulkintaa.

on esimerkkejä varianttien virheellisestä luokitelusta tautia aiheuttaviksi, kun alkuperäistutkimuksessa on käytetty puutteellista väestökohdataista verrokkidataa (8). Kaikkien alun perin virheellisesti luokiteltujen varianttien statusta ei ole korjattu vielä uusimmissakaan huolellisesti ylläpidetyissä julkisissa mutaatiotietokannoissa (8).

Vain pieni osa suomalaisten potilaiden kliinissä eksomisekvensoinneissa löytyneistä mutaatioista on kuvattu aiemmin. Tämän vuoksi tulosten tulkinnassa on tärkeää ottaa huomioon myös aiemmin kuvaamattomat, proteiinin toiminnan kokonaan estävät ja vahingollisiksi ennustetut aminohappoa muuttavat tunnettujen tautigeenien mutaatiot. Luotettavan ja jatkuvasti lisääntyvän verrokkidatan saatavuus eri väestöistä auttaa aiemmin raportoimattomien mutaatioiden tulkintaa nykyään merkittävästi. Harvinaisten mutaatioiden erilainen esiintyvyys eri väestöissä puoltaakin tarvetta väestökohtaiseen verrokkidataan ja potilaiden etnisen taustan tuntemiseen. Esimerkiksi useat muissa eurooppalaisissa väestöissä hyvin harvinaiset

variantit ovat rikastuneet Suomeen, ja suomalaisen verrokkidatan puuttuminen dataa analysoitaessa voi johtaa virheelliseen tulkintaan variantin yleisyydestä väestössä ja mutaation ennustamiseen tautia aiheuttavaksi. Tiedossamme ei ole, kuinka hyvin ulkomaiset eksomisekvensointipalveluja tuottavat laboratoriot huomioivat suomalaisen väestön geneettisiä erityispiirteitä.

### Taustatieto takaa luotettavimman tulkinnan

Parhaaseen tulokseen eksomidatan tulkinnassa päästään, kun käytettävissä on laaja kuvaus potilaan taudinkuvasta, kartoitus taudin mahdollisesta esiintymisestä suvussa ja tieto etnisestä taustasta. Vanhempienkin näytteet kannattaa liittää analyysiin, mikäli se on mahdollista. Tulkinnan luotettavuutta lisää sujuva vuoropuhelu lähettävän ja dataa analysoivan tahon välillä. Vanhempien näytteiden eksomisekvensointi mahdollistaa de novo -mutaatioiden kartoittamisen ja mahdollisesti löytyvän mutaation periytyvyyden selvittämisen. Tieto taudin esiintymisestä potilaan lähisukulaisilla ja suvussa laajemmin kertoo todennäköisimmän periytyvän mallin ja sen, tulisiko eksomidatasta ensisijaisesti etsiä vallitsevasti vai peittyvästi periytyvää mutaatiota vaiko X-kromosomiin kytkeytyvää.

Tapausesimerkkejä odottamattomastakin periytyvän mallista on: eksomisekvensoinnilla löydettiin de novo -mutaatio vastasyntyneeltä lapselta, jonka vanhemmat olivat serkuksia keskenään, jolloin oletuksena oli peittyvästi periytyvä mutaatio. Tämän vuoksi eksomidata on aina analysoitava niin, että mahdolliset vaihtoehdot periytyvästä mutaatiosta huomioidaan. Myös taudinkuvan vaihtelu ja mahdollinen monimuotoisuus tulee ottaa huomioon tulosten tulkinnassa. Valtavasti lisääntynyt tutkimuksellinen ja kliininen genomidatan hyödyntäminen tautigeenien etsinnässä on tuonut esiin useita esimerkkejä siitä, kuinka aiemmin kuvattu fenotyyppi kattaa vain osan tietyn geenin mutaatioiden aiheuttamasta tautikirjosta (9).

Erityisesti yhden geenin geenivirheiden aiheuttamat harvinaiset sairaudet, joita sairastavat

tavia potilaita on kuvattu vain muutamia yksilöitä, voivat olla tautikirjoltaan merkittävästi alkuperäistä kuvausta monimuotoisempia. Geneettisissä tutkimuksissa on tavanomaisesti hyödynnetty samankaltaisten potilaiden ryhmiä, mikä on saattanut kaventaa nykykäsitystämme tietyn geenin virheiden aikaansaamasta taudinkuvan monimuotoisuudesta.

## Sattumalöydösten tulkinta

Vielä haastavampaa kuin tautia aiheuttavan mutaation tunnistaminen on sattumalöydösten merkityksen tulkinta. Esimerkiksi tästä sopii tutkimus, jossa kolme kokenutta klii-

nistä genomianalyysia tarjoavaa laboratorioita sai analysoitavakseen kahden perinnöllisiä rytmihäiriöitä aiheuttavan geenin (*SCN5A*, pitkä QT -oireyhtymä ja *KCNH2*, Brugada-oireyhtymä) sekvenssidatan 2022 henkilöstä ilman tietoa mahdollisesta taudinkuvasta ja etnisestä taustasta (10). Laboratoriot tai julkinen ClinVar-mutaatietietokanta tunnistivat patogeenisiksi 32 aminohappoa muuttavaa mutaatiota *SCN5A*-geenissä ja kymmenen mutaatiota *KCNH2*-geenissä. Yllättäen kaikki laboratoriot olivat tulkinnoissaan yksimielisiä vain neljän *SCN5A*-mutaation patogeenisuudesta, ja näidenkin kantajista vain kahdella oli rytmihäiriöoireyhtymä elektronisten potilastietojen

**TIETOLAATIKKO.** Aiheeseen liittyviä termejä.

**De novo -variantti** potilaalta tunnistettu, perheessä uusi variantti, jota ei löydy potilaan vanhempien ituradan genomista. Mutaatio on tapahtunut jommankumman vanhemman sukusolussa, hedelmöityneessä munasolussa tai alkionkehityksen varhaisvaiheessa.

**Deleetio** perimäaineksen häviämää

**Eksomisekvenssointi** kaikkien ihmisperimän proteiinia koodaavien alueiden emäsjärjestyksen selvittäminen hyödyntämällä uuden sukupolven sekvenssointimenetelmiä

**Geenin proteiinia koodaava alue**, joka koostuu koodaavista geenalueista eli eksoneista. Näiden lisäksi geenissä on ei-koodaavia intronijaksoja, geenin molemmissa päissä sijaitsevat 5'-UTR- ja 3'-UTR-alueet (untranslated region), joita ei käännetä proteiiniksi sekä geenin säätelystä vastaava promotorialue.

**Geenipaneeli** uuden sukupolven sekvenssointia hyödyntävä tietyn geeniryhmän analyysimenetelmä. Se voi olla suunniteltu kattamaan esimerkiksi kaikki tiettyä tautia aiheuttavaksi kuvatut mutaatiot tai kaikki tiettyä tautia aiheuttavaksi kuvattujen geenien eksonit. Niin sanottu virtuaaligeenipaneeli on kyseessä silloin, kun sekvensoidaan koko genomi tai koko eksomi, mutta data analysoidaan vain tietyn, ennalta valitun geeniryhmän osalta.

**Genomi** ihmisen koko perimä, sisältää kaikki geenit ja myös ne perimän alueet, jotka eivät koodaa proteiineja

**Insertio** perimän lisäys, mutaatio jossa DNA-ketjuun liittyy uutta perimäainesta

**Konservoitunut aminohappo** tietyn proteiinin tai proteiiniperheen aminohappo, joka on sama useilla tai jopa kaikilla eliölajeilla. Tyypillisesti proteiinin toiminnan kannalta tärkeitä aminohappoja ei voida korvata toisilla aminohapoilla, minkä vuoksi sama aminohappo on evoluutiossa säilynyt proteiinissa. Konservoituneiden aminohappojen kodoneissa tapahtuvat mutaatiot johtavat usein proteiinin toiminnalle vakaviin seurauksiin, koska tällaisissa aminohapoissa ei esiinny vaihtelua ilman, että sillä olisi seurauksia proteiinin rakenteeseen tai toimintaan.

**Loss-of-function-mutaatio** estää perintökemian koodaaman proteiinin toiminnan tai vähentää sen vaikutusta. Tällaisiksi mutaatioiksi luetaan ennenaikaisen lopetuskodonin aiheuttavat mutaatiot, RNA:n silmukointiin vaikuttavat mutaatiot ja proteiinin lukukehystä muuttavat mutaatiot.

**Peittyvässä periytymisessä** geenin virheellisen alleelin pitää periytyä molemmilta vanhemmilta, jotta

sairaus ilmenee yksilössä. Yksi virheellinen alleeli ei riitä aiheuttamaan sairautta.

**Sattumalöydös** eksomisekvenssoinnissa tunnistettu mutaatio, jonka tiedetään aikaisempien tulosten perusteella olevan patogeeninen jossain toisessa kuin tutkimuksen kohteena olevassa sairaudessa.

**Toistojako** genomien alue, jossa tietty emäsjakso toistuu useita kertoja ja joka voi sisältää muutamasta emäsparista satoihin emäksiin ja toistua muutamia tai jopa tuhansia kertoja.

**Valegeeni** geenin lähes identtinen kopio, ns. pseudogeeni, joka esiintyy muualla genomissa. Usein täysin toimimaton geeni.

**Vallitsevassa periytymisessä** riittää, että geenistä esiintyy yksi sairaus alleeli, jotta tauti ilmenee yksilöllä. Tauti esiintyy usein monessa peräkkäisessä sukupolvessa.

**Verrokkigenomi** kansainvälisen genomisekvenssointiprojektin tuottama sekvenssidata, jota on sovittu käytettäväksi vertailusekvenssinä genomianalyseissa.

**VUS**, variant of unknown significance: variantti, jonka merkitystä potilaan taudinkuvulle ei pystytä nykytiedon valossa arvioimaan.

tai EKG:n perusteella. Kaiken kaikkiaan 18:lla näistä 32:sta mahdollisen toiminnallisen mutaation kantajasta oli rytmihäiriöihin liittyviä oireita. Tautia aiheuttavaksi heidän mutaation sa raportoi 0–3 laboratoriota ja yhden mutaatioista osalta vain ClinVar-tietokanta.

KCNH2-geenin osalta laboratoriot eivät olleet yksimielisiä yhdestäkään mutaatiosta, ja patogeenisiksi raportoitujen (1–2 laboratoriota) mutaatioiden kantajista neljällä yhdestätoista esiintyi rytmihäiriöoireita. Raportoitu tutkimus kuvaa erinomaisesti genomitiedon tulkinnan vaikeutta etenkin tilanteessa, jossa potilaan taudinkuvaa ja tietoa tämän etnisestä taustasta ei ole käytettävissä. Tutkimus antaakin esimerkiksi oireettomien henkilöiden sattumalöydösten raportoinnin vaativuudesta.

### Diagnoosi tai ei – tavallinen raportti

Eksomisekvensoinnin hyödyntäminen kliinisessä potilastyössä on johtanut selkeään diagnoosiin vaihtelevalla osalla potilaista, sen mukaan minkä tyyppistä potilasmateriaalia on tutkittu (11). Diagnoosin löytyminen on todennäköisempää lapsi- kuin aikuispotilailta, ja merkittäviä eroja sairausryhmien välillä on kuvattu. Lähi-idän väestöissä, joissa serkusavoliitot ovat yleisiä, on raportoitu uuden sukupolven sekvensointia hyödyntävien geenipaneelien diagnostiseksi tehokkuudeksi jopa yli 60 % joissakin tautiryhmissä (ihotaudit, endokriiniset taudit, synnynnäiset aineenvaihduntahäiriöt). Toisia tauteja (hematologiset sairaudet, primääriset immuunipuutokset, sydämen kehityshäiriöt) sekvensoitaessa diagnoosi löytyi toisaalta vain 10–25 %:lta (12). Huomattavaa on, että tyypillisin syy sille, ettei paneelilla tunnistettu potilaan tautimutaatiota, oli geenin puuttuminen tautiryhmälle suunnitellusta paneelistä. Tämä johtui joko taudinkuvan heterogeenisuudesta tai siitä, että tautimutaatiot oli kuvattu vasta paneelin suunnittelun jälkeen.

Uusien tautigeenien tunnistamisen ja käytettävissä olevan verrokkidatan jatkuvan lisääntymisen ennustetaan parantavan diagnoosien saatavuutta koko ajan. Tuoreessa julkaisussa 40 % primaarista immuunipuutosta sairastavista potilaista sai molekyyliidiagnoosin ekso-

misekvensoinnin perusteella (13). Huomattavaa on, että jopa 55 %:n diagnoosi muuttui tai täsmentyi kliinisen kuvan perusteella tehdystä ja että jopa neljännekselle molekyyliidiagnoosi tarjosi mahdollisuuden täsmähoitoon.

Eksomisekvensointitutkimuksen raportin sisältö ei ole vielä laajasti standardoitu, ja siihen saattavat kuulua ainakin **TAULUKOSSA 2** esitellyt varianttiluokat (11).

Mahdollisia syitä sille, että tautia aiheuttavaa geenivirhettä ei löydetä, on monia. Useimmat nykyisin käytössä olevat eksomirikastusmenetelmät kattavat suurimman osan geenien proteiinia koodaavista alueista (yli 95 %), mutta esimerkiksi geenien ei-koodaavat alueet, kuten 5'- ja 3'-UTR, on jätetty rikastuksen ulkopuolelle sekvensointikustannuksen pienentämiseksi. Keskimääräisen lukusyvyyden lisääminen lisää analyysin hintaa mutta ei tuo kaikkia geenialueita analyysin piiriin. Datan analyysiä vaikeuttavat geenit, joista on genomissa useita kopioita tai vagegenejä, koska luettujen sekvenssien rinnastusta ei voida tehdä luotettavasti. Geenivirhe jää mahdollisesti havaitsematta myös, jos se sijaitsee geenien koodaavan alueen

**TAULUKKO 2.** Eksomisekvensointiraportti (11).

Varianttiluokka	Tyypillisesti raportoitu varianttien lukumäärä
Vahingolliset mutaatiot, jotka todennäköisesti aiheuttavat tutkitavan potilaan taudin	0–2
Potilaan sairauteen mahdollisesti liittyvät variantit, joiden merkitystä ei nykytiedon valossa pystytä arvioimaan <sup>1</sup>	2–10
Mutaatiot, jotka aiheuttavat lääketieteellisesti hoidettavissa olevia sairauksia, jotka eivät ole potilaalla vielä ilmenneet, ns. sattumalöydökset <sup>2</sup>	0–1
Peittyvästi periytyvien tautien tunnettujen tautimutaatioiden kantajuus	0–1
Lääkeainemetaboliaan vaikuttavat variantit	0–4

<sup>1</sup>Niin sanottu VUS = variant of unknown significance

<sup>2</sup>Useimmin havaitut sattumalöydökset: rinta- ja munasarjasyöpäalttius (BRCA2, BRCA1), sydämen johtumishäiriöt ja rytmihäiriöt (DSC2, KCNQ1, SCN5A), Marfanin oireyhtymä (FBN1), hypertrofinen kardiomyopatia (MYL2), perinnöllinen ei-polypoottinen paksusuolisyöpä (PMS2)



ulkopuolella tai jos taudin aiheuttaa deletio, monistuma tai muutos toistojaksossa, joita nykyiset sekvensointimenetelmät eivät aina löydä. Negatiivisen tulokseen voi johtaa sekin, että taudin aiheuttavaa geeniä ei ole vielä tunnistettu tautigeeniksi ja että kykymme tunnistaa havaittu variantti tautia aiheuttavaksi on rajallinen. Tämän vuoksi negatiivinen eksomisekvensointitulostus ei sulje pois geneettistä syytä potilaan taudin taustalla.

Useat diagnostista eksomisekvensointia tarjoavat laboratoriot analysoivat potilasdatan säännöllisin tai epäsäännöllisin väliajoin uudestaan, jolloin edellisen analyysin jälkeen raportoidut uudet tautigeenit voidaan ottaa mukaan analyysiin. Tällöin diagnoosiin voidaan päästä vielä kuukausien tai vuosienkin kuluttua tutkimuksen tekemisestä.

## Lopuksi

Eksomisekvensointi on nopea, luotettava ja kustannusvaikuttava menetelmä useimpien genomissamme esiintyvien muutosten havaitsemiseen. Tämän vuoksi sitä tulisi hyödyntää diagnostisena tutkimuksena silloin, kun mahdollinen sekvensointitulostus voi johtaa täsmälliseen diagnoosiin, vaikuttaa merkittävästi sairauden hoitoon ja ohjata potilaan perinnöllisyysneuvonnan piiriin. Merkittävin haaste

eksomisekvensointitulosten tulkinnassa on erottaa merkityksettömät muutokset sairauden aiheuttavista muutoksista. Genomitiedon lisääntyminen ja jatkuvasti lisääntyvä luotettava verrokkipotilasdata eri väestöistä helpottavat jatkossa eksomidatan tulkintaa. Harvinaisten mutaatioiden erilainen esiintyvyys eri väestöissä puoltaa kuitenkin tarvetta väestökohtaiseen verrokkipotilastietoon ja potilaan etnisen taustan tuntemiseen.

Erittäin tärkeää olisi myös pystyä hyödyntämään kliinisenä tutkimuksena tehdyn eksomisekvensoinnin dataa jatkossa tieteellisissä tutkimuksissa, kun potilas antaa siihen suostumuksensa. Tämä voi lyhyellä aikavälillä tuottaa uutta tietoa sairauksien syistä ja nopeuttaa potilaan omankin diagnoosin selviämistä sekä pitkällä tähtäimellä kasvattaa kansallista genomitietovarantoa. Suomessa valmistellaan kansallisen genomistrategian konkreettista toteuttamista, muun muassa kansallisen genomikeskuksen perustamista. Se mahdollistaa olemassa olevan genomidatan tallentamisen ja paremman hyödyntämisen sekä kliinisesti että tutkimuksellisesti. Tämä valmistelutoiminta tarjoaa mahdollisuuksia myös tuotetun genomidatan pitkäaikaiskäytölle ja käytölle. Ennen kuin kansallinen genomikeskus on toiminnassa, tulisi kaikkien genomidataa hyödyntävien tahojen kantaa vastuuta datan säilyttämisestä tulevaisuutta varten lain sallimissa puitteissa. ■

JANNA SAARELA, LT, dosentti, tutkimusjohtaja

KAISA KETTUNEN, FT, vanhempi tutkija

Suomen molekyyliäätieteen instituutti FIMM, Helsingin yliopisto

### SIDONNAISUUDET

Janna Saarela: Apuraha (Sanofi-Genzyme), luentopalkkio (Biotexel, Roche), korvaukset koulutus- ja kongressikuluista (CSL Behring ja Steripolar, Octapharma, Sanofi-Genzyme)

Kaisa Kettunen: Luentopalkkio (PSSH), korvaus koulutus- ja kongressikuluista (CSL Behring ja Steripolar)

### SUMMARY

#### Who would benefit from exome sequencing?

Next-generation sequencing methods have revolutionized the possibilities for analyzing the human genome. Sequencing the exome, the protein-encoding portion of the genome, is used, in addition to research, in clinical medicine, especially in the diagnosis of rare hereditary diseases, congenital developmental disorders and cancer. Using exome sequencing as a diagnostic test is justified when the results could lead to an accurate diagnosis, significantly influence the treatment and genetic counseling. It is a reliable method for detecting single base mutations as well as minor deletions and insertions. However, with current methods the reliable analysis of larger changes in the number of copies, the length of repeats and areas present in multiple copies in the genome is challenging. Every human has many mutations in their exome, and distinguishing between insignificant and pathogenic mutations is thus a key challenge when interpreting the results of exome sequencing.

**KIRJALLISUUTTA**

1. Ku CS, Naidoo N, Pawitan Y. Revisiting Mendelian disorders through exome sequencing. *Hum Genet* 2011;129:351–70.
2. Yang Y, Muzny DM, Reid JG, ym. Clinical whole-exome sequencing for the diagnosis of mendelian disorders. *N Engl J Med* 2013;369:1502–11.
3. Green RC, Berg JS, Grody WW, ym. ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing. *Genet Med* 2013; 15:565–74.
4. Tennessen JA, Bigham AW, O'Connor TD, ym. Evolution and functional impact of rare coding variation from deep sequencing of human exomes. *Science* 2012; 337:64–9.
5. MacArthur DG, Balasubramanian S, Fran- kish A, ym. A systematic survey of loss- of-function variants in human protein- coding genes. *Science* 2012;335:823–8.
6. Kong A, Frigge ML, Masson G, ym. Rate of de novo mutations and the importance of father's age to disease risk. *Nature* 2012; 488:471–5.
7. Botstein D, Risch N. Discovering genotypes underlying human phenotypes: past successes for mendelian disease, future approaches for complex disease. *Nat Ge- net* 2003;33(Suppl):228–37.
8. Manrai AK, Funke BH, Rehm HL, ym. Genetic misdiagnoses and the poten- tial for health disparities. *N Engl J Med* 2016;375:655–65.
9. Lu JT, Campeau PM, Lee BH. Genotype- phenotype correlation – promiscuity in the era of next-generation sequencing. *N Engl J Med* 2014;371:593–6.
10. Van Driest SL, Wells QS, Stallings S, ym. Association of arrhythmia-related genetic variants with phenotypes documented in electronic medical records. *JAMA* 2016; 315:47–57.
11. Yang Y, Muzny DM, Xia F, ym. Molecular findings among patients referred for clinical whole-exome sequencing. *JAMA* 2014;312:1870–9.
12. Saudi Mendeliome Group. Comprehen- sive gene panels provide advantages over clinical exome sequencing for Mendelian diseases. *Genome Biol* 2015;16:134.
13. Stray-Pedersen A, Sorte HS, Samarakoon P, ym. Primary immunodeficiency di- seases: genomic approaches delineate heterogeneous Mendelian disorders. *J Allergy Clin Immunol*. Julkaistu verkossa 16.7.2016. DOI 10.1016/j.jaci.2016.05.042.