

<https://helda.helsinki.fi>

---

## Koko transkriptomisekvensoinnin (RNAseq) käyttö neuromuskulaaritautilien tutkimuksessa ja diagnostiikassa

Hackman, Peter

2020

---

Hackman , P 2020 , ' Koko transkriptomisekvensoinnin (RNAseq) käyttö neuromuskulaaritautilien tutkimuksessa ja diagnostiikassa ' , Kliin lab : kliinisen laboratoriaoalan julkaisu : SKKY:n jäsenlehti. , Vuosikerta. 37 , Nro 1 , Sivut 6-8 . < [https://www.skky.fi/sites/skky.fi/files/media/Kliinlab\\_1\\_2020\\_screen.pdf](https://www.skky.fi/sites/skky.fi/files/media/Kliinlab_1_2020_screen.pdf) >

---

<http://hdl.handle.net/10138/318116>

---

publishedVersion

---

*Downloaded from Helda, University of Helsinki institutional repository.*

*This is an electronic reprint of the original article.*

*This reprint may differ from the original in pagination and typographic detail.*

*Please cite the original version.*

# Koko transkriptomisekvensoinnin (RNAseq) käyttö neuromuskulaaritautilien tutkimuksessa ja diagnostiikassa

*Peter Hackman*

**PERINNÖLLISIÄ** neuromuskulaaritauteja on aikaisemmin pidetty suhteellisen harvinaisina Suomessa ja täten vähemmän tärkeinä terveyden kannalta yleisesti. Itse asiassa näiden tautien esiintyvyys Suomessa on kuitenkin arvioitu olevan noin 2/1000. Tämä on samalla tasolla joidenkin muiden yleisten hermoston rappeutumishäiriöiden kanssa, kuten esimerkiksi Parkinsonin taudin ja multippeliskleroosin (MS) (1, 2). Tämän perusteella Suomessa löytyisi noin 10 000 henkilöä, jotka kärsivät jostakin perinnöllisestä neuromuskulaaritaudista. Nykyiset arvioinnit viittaavat siihen, että oikea määrä voisi olla jopa korkeampi, varsinkin kun monet potilaat jäävät ilman tarkkaa molekyyllitason diagnoosia. Tämä on merkittävä terveysongelma, koska oikea geneettinen diagnoosi on erittäin tärkeää neuvonnan ja hoitojen kannalta sekä myös virheellisten hoitojen välttämiseksi (3, 4).

Lihassairaudet kuuluvat neuromuskulaaritauteihin. Suurin osa lihassairauksista on perinnöllisiä, yhdestä geenistä johtuvia. Tällä hetkellä on yli 200 tunnettua geeniä, joiden tiedetään liittyvän lihassairauksiin. Lihassairauksia on kuitenkin kuvattu maailmalla paljon enemmän, mikä viittaa siihen, että kaikkia lihassairauksille altistavia geenejä ja niiden variantteja ei vielä ole tunnistettu. Monilla potilailla on oireita, jotka ovat yhteensopivia tunnettujen lihassairauksien kanssa, mutta tautia ei ole vielä pystytty yhdistämään tiedossa oleviin alttiusgeeneihin.

Perinnöllisissä lihassurkastumistaudeissa nähdään usein valikoivaa lihasten vaurioitumista huolimatta tautigeenien ilmentymisestä kaikissa tai monissa lihaksissa, sekä usein

myös muissa kudoksissa. Surkastumista nähdään aluksi usein vain tietyissä lihaksissa. Esimerkiksi tibiaalisessa lihasdystrofiassa tauti alussa ilmenee vain ”jalan-nosto” lihaksessa, tibialis anteriorissa. Yksi tavoitteistamme onkin selvittää miksi taudit ilmenevät vain tietyissä lihaksissa vaikka geenit ilmenevät kaikissa tai monessa lihaksessa.

Niin kutsuttujen seuraavan sukupolven sekvensointimenetelmien (eng. Next Generation Sequencing, NGS) kehittyminen on merkittävästi vauhdittanut tauteja aiheuttavien mutaatioiden tunnistusprosessia. Kohdennetut NGS-paneelit (TNGS), koko eksomisekvensointi (WES) ja koko genomin sekvensointi (WGS) ovat nykyään rutiinimenetelmiä sekä perinnöllisten tautien perustutkimuksessa että yhä enemmässä määrin myös näitten sairauksien diagnostiikassa. Koko genomin transkriptomisekvensointi (RNAseq), joka lasketaan myös NGS-menetelmäksi on erinomainen uudempi menetelmä, jolla voidaan analysoida biologisten näytteiden geenien ilmentymisprofiilia ja sekvenssejä samanaikaisesti. Kaikki nämä menetelmät ovat huomattavasti halventuneet viime vuosien aikana (ks. esimerkki kustannuskehityksestä National Human Genome Research Institute; <https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost>).

Jotta voisi tutkia geenien ilmentymisen poikkeavuuksia sairaista kudoksista, täytyy ensin tuntea kyseisten geenien ilmentyminen normaalikudoksissa. Tätä varten RNAseq:llä on analysoitu näytteitä, jotka ovat peräisin alaraajojen luustolihasnäytteistä yksilöiltä, jotka eivät tietävästi ole sairastaneet

lihastauteja. Analyysit osoittivat merkittäviä eroavuuksia geenien ilmentymisessä anatomisesti eri lihaksissa sekä geenien useiden erilaisen isomuotojen ilmentymistä samassa lihaksessa. Tulokset osoittivat myös mahdollista yhteyttä ilmentymiserojen ja perinnöllisissä lihastaudeissa nähtävän tiettyjen lihasten sairastumisen välillä. Tämä viittaa siihen, että lihasgeenien erilainen ilmeneminen eri lihaksissa on yksi määräävä tekijä tiettyjen lihasten osallisuudesta lihassurkastumataudeissa (5, 6).

Titiinigeenin (TTN) mutaatioiden on tähän mennessä raportoitu olevan useiden erilaisten lihassairauksien syytä. Näihin sairauksiin, joita kutsutaan kollektiivisesti titinopatioiksi, kuuluu luustolihas- ja sydänsairauksia tai niiden yhdistelmiä (7). TTN-geenin koodaama titiiniproteiini on pisin tunnettu proteiini luonnossa. Titiiniproteiini on hyvin keskeinen lihaksen perusrakenteen - sarkomeerin - rakenteen ja toiminnan kannalta. Geenin valtava koko (yli 100 kiloemästä) tekee analysoinnin perinteisellä Sanger-sekvensointimenetelmällä käytännössä mahdottomaksi. Geenin suuresta koosta johtuen samalla yksilöllä voi olla useita harvinaisia TTN-variantteja, ja oikean aiheuttavan mutaation tunnistaminen vaarattomien varianttien joukosta voi näin ollen olla haastavaa. Proteiinista tunnetaan useita vaihtoehtoisia isomuotoja, ja näiden isomuotojen ilmentyminen on osittain myös lihasspesifistä. Eri lihaksissa siis ilmenee vaihtelevia määriä tai erilaisia isomuotoja lihasgeneista. TTN:n silmukointia on tutkittu RNAseq:llä ja on löydetty aikaisemmin tunnistamattomia isomuotoja (ks. kuva 1).



**Peter Hackman**  
dosentti  
Hallinnollinen ryhmänjohtaja  
Folkhälsan  
peter.hackman@helsinki.fi

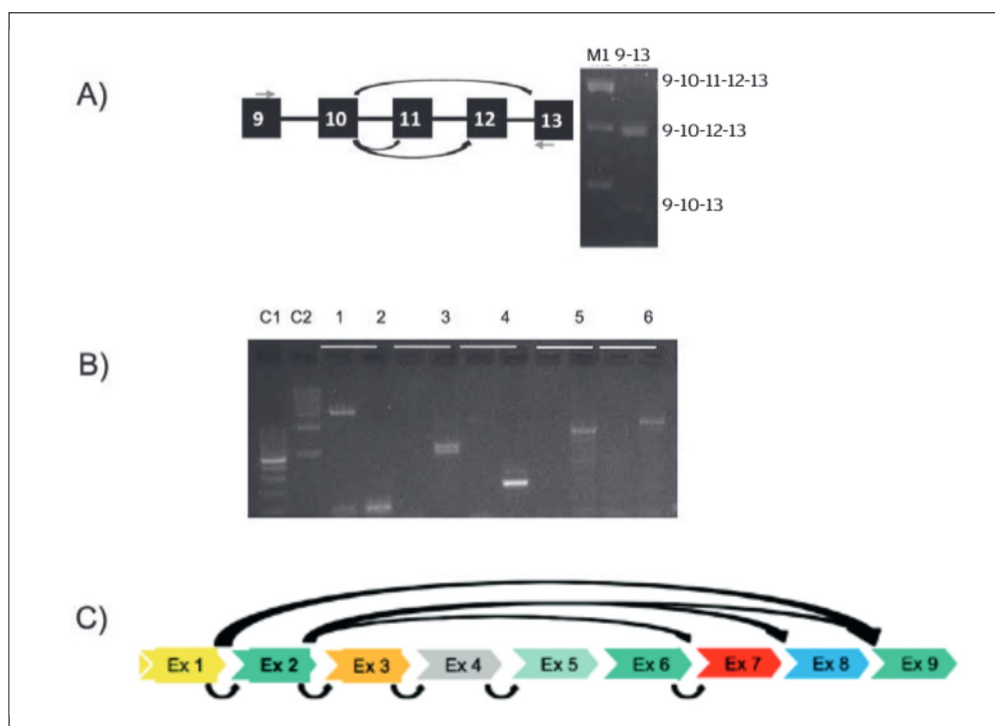
Niiden tarkkaa funktiota ei vielä tunneta, mutta tutkimus osoitti, että tiiniin silmukointi luustolihasissa on monimuotoisempaa kuin aikaisemmin arveltiin. Tämä tieto on tärkeää, kun halutaan tutkia TTN-varianttien kliinistä merkitystä lihaksissa. Vaikka taustalla olevat molekyyli-tason mekanismit, jotka selittävät lihasten valikoivaa osallistumista lihastaudeissa, tunnetaan edelleen huonosti, auttavat nämä tulokset tautimekanismien selvityksessä.

RNAseq on tärkeä ja täydentävä työkalu neuromuskulaaristen sekä myös kaikkien muiden perinnöllisten tautien diagnostiikassa ja tutkimuksessa. Menetelmää ei ole toistaiseksi käytetty järjestelmällisesti perinnöllisiä lihastauteja tutkittaessa. RNAseq:llä on kuitenkin hiljat-

tain raportoitu löytyneen joidenkin perinnöllisten sairauksien aiheuttajamutaatioita. Viimeaikaisia julkaistuja esimerkkejä tästä on tiettyjen mitokondriopatioiden syyn löytyminen potilaiden fibroblasteja tutkimalla. Toinen esimerkki on kollageeni VI -tyyppinen dystrofia, joka todettiin RNAseq:llä käyttäen lihasbiopsioita (8, 9). Menetelmä laajentaa tutkimusmahdollisuuksia ja on tehokas työkalu myös geenien koodaavien alueiden ulkopuolisten varianttien sekä ns. hiljaisten koodaavien alueiden muutosten havaitsemiseksi. "Hiljaiset muutokset" eivät aiheuta muutosta koodaavissa aminohapossa, mutta voivat aiheuttaa muutoksia esimerkiksi geenin ilmentymistasossa tai silmukoitumisessa. RNAseq:llä voidaan myös tutkia geenien ilmen-

tymisen muutoksia yksilössä eri ajan-kohtina tai elämänvaiheissa, ottamalla näytteitä eri ajanjaksoina samasta kudoksesta. Koska menetelmää voidaan käyttää taudin seurannassa, voidaan ymmärtää paremmin taudin mekanisme molekyylitasolla. RNAseq mahdollistaa siis täysin uuden analyysitason, joka lisää onnistuneen diagnoosin mahdollisuutta ja voi myös helpottaa uusien tautien aiheuttavien geenien tunnistamisessa.

RNAseq:llä voidaan siis tutkia sairauksia aiheuttavia genejä ja variantteja tapauksissa, joissa aiemmat DNA-pohjaiset NGS-menetelmät ovat epäonnistuneet. Sitä voidaan myös käyttää ensimmäisenä testinä joissakin tapauksissa. Menetelmällä voidaan lisäksi seurata geenien erilais- ta ilmentymistä eri ajanjaksoina sekä



**Kuva 1. Esimerkkejä geenin transkriptien vaihtoehtoisesta silmukoinnista ja uusien isomuotojen löytymisestä.**

A) Geenin eksoni 10 yhdistyy/silmukoituu uudella tavalla eksonin 12 ja 13 kanssa, sekä agarosigeelikuva kyseisistä RNA isomuotofragmenteista.  
B) Agarosigeeli kuva eri pituisista RNA isomuotofragmenteista.  
C) Esimerkki miten eksonit 1–9 geenissä voivat yhdistyä uudella tavalla RNA isomuodoissa.

→ analysoida paremmin eri lihasten välisiä eroja. Saatua datan määrä on kuitenkin huomattavasti suurempi kuin DNA-pohjaisilla NGS-menetelmillä. Tämä tekee tulosten analyysistä huomattavasti haastavampaa. Bioinformaattisia analyysimenetelmiä kehitetään siksi jatkuvasti sekä diagnostiikka- että tutkimustarkoituksia varten.

Tällä hetkellä suuri osa lihastautipotilaista ei saa lopullista geneettistä diagnoosia, jos käytetään DNA-pohjaisia NGS-menetelmiä. RNAseq:llä voidaan nyt testata suoraan näytteitä sairaasta kudoksesta, yleensä lihasbiopsioista. Tällä tavalla voidaan havaita sairauten mahdollisesti liittyvien geenien poikkeava ilmentyminen tietyissä lihaksissa sekä poikkeavuuksia myös muiden lihasgeenien ilmentymisessä. Voimme tutkia poikkeavaa silmukointia, tiettyjen isomuotojen poikkeavaa esiintymistä sekä esimerkiksi eri geenivarianttien yli- tai aliekspressiota. RNAseq:llä on mahdollista havaita myös geneettinen mosaikismi, eli eri alleelien esiintyminen ja ekspressio eri kudoksissa. Potilailla saattaa myös olla erilaisia häiriöitä lihasgeenien ilmentymisessä eri elämän vaiheessa. Esimerkiksi sikiövaiheessa voi ilmentyä ongelmia, jotka vaikuttavat lihasten kehitykseen haitallisesti, mutta jotka palautuvat myöhemmin elämässä. Lihakset, jotka eivät muodostu sikiö-

kehityksen aikana, voivat kuitenkin puuttua tai jäädä alikehittyneiksi. Yksi RNAseqin rajoituksista on se, että menetelmään tarvitaan kudoksenäytteitä. Niitä voi olla vaikea saada yksilöiltä, joita halutaan tutkia. Aina ei kuitenkaan tarvita tutkimusnäytteitä. Hyödyllisiä vertailutietoja voi hakea myös julkisista arkistoista kuten Gene Expression Omnibus (GEO; <https://www.ncbi.nlm.nih.gov/geo/>) ja (<https://www.encodeproject.org/>).

RNAseq-analyysi perustuu suhteellisen lyhyisiin sekvensseihin. Alustavia RNA-tason tuloksia, jotka löytyvät RNAseq:llä, voidaan varmistaa ja tarkentaa muilla niin sanotun kolmannen sukupolven sekvensointimenetelmillä. Näitä ovat esimerkiksi hyvin pitkien PCR-tuotteiden sekvensointi yli eksonirajojen, PACbio- (enintään 30 KB) tai Oxford Nanopore Technologies (yli 100 KB) -menetelmillä. Pitkien sekvenssien avulla voidaan tunnistaa paitsi yksittäiset silmukoinnin liitoskohdat myös liitostapahtumien ryhmät, sekä miten ne on yhdistetty toisiinsa pitkissä transkripteissa. RNAseq:llä löydetty mahdolliset patogeeniset muunnokset voidaan tarkistaa edelleen esimerkiksi eri toiminnallisten testien avulla. Tämä voidaan tehdä proteiinitalon testeillä potilasnäytteistä tai esimerkiksi solulinjoilla.

Uusi kehittyvä menetelmä on niin

sanottu single-cell RNAseq, joka periaatteessa mahdollistaa geenien ilmentymisen analyysin yhdestä ainoasta solusta. Tämä mahdollistaa käytännössä analyysin hyvin pienestä lähtömateriaalista. Samaa solulinjaa tai kudosta voidaan seurata eri aikapisteissä, esimerkiksi sairauten etenemisen mukaan. Voidaan myös analysoida solujen erilaistumisen aikana tapahtuvaa geenien ilmentymisen muutosta tietyssä kudoksessa. Toinen kehittyvä menetelmä on ribosomin profilointi (RiboSeq), joka on ribosomiin sitoutuneiden mRNA-fragmenttien sekvensointi. Tämä tekniikka voi paljastaa miten mRNA muunnetaan proteiiniksi, sekä geenien todellisemmat ilmentymistasot.

RNAseq:iä käytetään yhä enemmän molekyylibiologisten prosessien tutkimuksessa. Voidaan tutkia koko transkriptomia, kohdennetusti tiettyjen geenien, yksittäisen solun RNA ilmentymistä tai erilaisten geenien säätelyyn liittyvien RNA luokkien, miRNA, siRNA ym. ilmentymistä. Sillä on siis monia sovelluksia genetiikan ja biolääketieteen alalla, ja uusia on tulossa. Yhteenvetona voi todeta, että RNAseq on muuttamassa perinnöllistä tutkimusta ja diagnostiikkaa nykyisestä geeni-keskeisestä enemmän yksilöllisen geneettisen ilmentymisen tutkimisen suuntaan. ♦

## Kirjallisuusluettelo

1. Kaakkola S. Parkinsonin tauti. Lääkärin käsikirja 9.5.2018
2. MS-tauti. Käypä hoito 16.12.2015. Suomalaisen Lääkäriseuran Duodecimin ja Suomen Neurologinen Yhdistys ry:n asettama työryhmä
3. Udd B. Distal myopathies. *Curr Neurol Neurosci Rep.* 2014;14(3):434
4. Udd B. Lihastautien diagnostiikka tarkentuu – potilaita Suomessa yli 10 000 *Suomen Lääkärelehti* 2007;62(6):519-22
5. Huovinen S, Penttilä S, Somervuo P, Keto J, Auvinen P, Vihola A et al. Differential Isoform Expression and Selective Muscle Involvement in Muscular Dystrophies. *Am J Pathol.* 2015;185(10):2833-42.
6. Savarese M, Jonson PH, Huovinen S, Paulin L, Auvinen P, Udd B et al. The complexity of titin splicing pattern in human adult skeletal muscles. *Skelet Muscle.* 2018;8(1):11.
7. Savarese M, Sarparanta J, Vihola A, Udd B, Hackman P. Increasing Role of Titin Mutations in Neuromuscular Disorders. *J Neuromuscul dis.* 2016;3(3):293-308
8. Scotton C, Bovolenta M, Schwartz E, Falzarano MS, Martoni E, Passarelli C et al. Deep RNA profiling identified CLOCK and molecular clock genes as pathophysiological signatures in collagen VI myopathy. *J Cell Sci.* 2016;129(8):1671-84.
9. Haberman Y, Karns R, Dexheimer PJ, Schirmer M, Somekh J, Jurickova I et al. Ulcerative colitis mucosal transcriptomes reveal mitochondriopathy and personalized mechanisms underlying disease severity and treatment response. *Nat Commun.* 2019;10(1):38