

Iida Salonen

ISOSITRAATTIDEHYDROGENAASIMUTAATIO

Vaikutus glioomakasvaimen immuunijärjestelmään

Lääketieteen ja terveysteknologian tiedekunta
Kandidaatin tutkielma
Huhtikuu 2020

TIIVISTELMÄ

Iida Salonen: Isositraattidehydrogenaasimutaatio – Vaikutus glioomakasvaimen immuunijärjestelmään
Kandidaatintutkielma
Tampereen yliopisto
Bioteknologian tutkinto-ohjelma
Huhtikuu 2020

Glioomat ovat eräitä aggressiivisimmista ja tappavimmista aivosyövistä. Niiden aggressiivisuudesta ja tappavuudesta kielii se, että tyypillisesti potilaat selviävät vain 15 kuukautta diagnoosin jälkeen. Tämän vuoksi on erityisen tärkeää löytää niihin mahdollisimman tehokkaita hoitoja. Isositraattidehydrogenaasi (IDH) toimii normaalisti sitruunahappokierrossa entsyyminä α -ketoglutarattia tuottavassa reaktiossa. IDH:n mutaatioilla on kuitenkin todettu yhteyksiä kasvaimen immuunijärjestelmän toimintaan glioomissa. On myös todettu, että mutaatio glioomissa tapahtuu jo syövän alkuvaiheessa niin kutsutuissa syövän kantasoluissa. On todettu myös, että glioomista IDH-villityypin gliooma, eli ei mutatoitunut muoto on kaikista tappavin.

Tämä työn tarkoituksena on selvittää IDH-mutaation vaikutuksia gliooman ja sen ympäristön immuunijärjestelmään ihmisessä. Kyseessä on R-ohjelmakielen avulla tehty ekspressioanalyysi jossa vertaillaan kolmea eri ryhmää keskenään. Ensimmäinen ryhmä on IDH-villityyppi, toinen IDH-mutaatio yhdistettynä 1p/19q deleetioon, ja kolmas IDH-mutaatio ilman 1p/19q deleetiota. Analyysissä käytetty data on peräisin The Cancer Genome Atlas –projektista. Dataa on niin alemmilta glioomaluokituksilta kuin myös korkeamman tason glioblastoomista. Ensin ryhmistä etsittiin ne geenit, joilla ekspressioero on tarpeeksi suuri. Tämän jälkeen tarpeeksi suuren ryhmien välisen ekspressioeron saaneiden geenien tilastollinen merkittävyys varmistettiin t-testin avulla. Lopuksi näitä geeniryhmiä analysoitiin KEGG-reittianalyysin ja Gene Ontology –annotaatioiden avulla. Merkittävimpänä havaintona voidaan pitää interferoni (IFN)- γ :n signaloimana tuotettavien kemokiinien vähenemistä niissä ryhmissä, joissa oli IDH-mutaatio. Tämä kemokiinien väheneminen johtaa siihen, että CD8+ tappaja-T-solut eivät pääse tunkeutumaan syöpäkudokseen yhtä hyvin. Analyysissä tuli myös esille luokan II MHC-kompleksin ekspression muutos. Se ilmeni ekspression pienemiseksi IDH-mutatoituneissa kasvaimissa. Kun luokan II MHC-kompleksien määrä laskee, laskee myös CD4+ auttaja-T-solujen määrä. IDH-mutatoituneissa syövissä siis sekä CD8+ tappaja-T-solujen, että CD4+ auttaja-T-solujen määrä laskee, heikentäen kudoksen immuunijärjestelmän toimintaa.

Avainsanat: Isositraattidehydrogenaasi, mutaatio, gliooma, syöpä, ekspressioanalyysi, immuunijärjestelmä

Tämän julkaisun alkuperäisyys on tarkastettu Turnitin OriginalityCheck –ohjelmalla.

ALKUSANAT

Tämä kandidaatintutkielma on kirjoitettu Tampereen yliopiston Lääketieteen ja terveysteknologian tiedekuntaan keväällä 2020. Niin tutkimustyön tekeminen kuin kirjoittaminen on opettanut paljon sekä itse aiheesta, että laskennallisen biologian tutkimustyön tekemisestä.

Haluan esittää kiitokseni ohjaajalleni professori Matti Nykterille ohjauksesta työn ohella, sekä tästä mahdollisuudesta tehdä tämä mielenkiintoinen tutkimus. Haluan kiittää myös professori Anne Kalli-oniemeä ohjauksesta ja kaikesta yleisestä kandidaatintyön tekemiseen liittyvästä avusta. Kiitokset myös kaikille niille ystäväilleni ja perheenjäsenille, jotka tätä työtä tehdessäni ovat tarjonneet minulle apuaan ja tukeaan.

Tampereella, 28.4.2020

Iida Salonen

SISÄLLYSLUETTELO

1. JOHDANTO.....	4
2. MATERIAALIT JA MENETELMÄT	5
2.1 TCGA datamatriisi.....	5
2.2 Alkutestaus	5
2.3 Tilastollinen testaus	6
2.4 KEGG ja GO testit.....	6
3. TULOKSET	7
4. PÄÄTELMÄT	13
4.1 Tulosten oikeellisuus.....	13
4.2 Biologiset merkitykset	15
LÄHTEET	17
LIITTEET	18
Koodit	18

1. JOHDANTO

Glioomat ovat eräitä yleisimmistä ja tappavimmista aivosyövistä (Hartmann ym. 2010, Bleeker ym. 2012). Glioomat, ja erityisesti glioblastoomat, ovat aivosyövistä kaikkein aggressiivisimpia. Niiden aggressiivisuudesta kertoo paljon se, että mediaani selviytymisaika potilailla diagnoosin jälkeen on 15 kuukautta (Bleeker ym. 2012). Ne jaotellaan WHO:n mukaan tasoihin I-IV niiden histologisten ja molekulaaristen piirteiden perusteella. Tason I kasvaimiin lukeutuu mm. pilosyyttiset astrozytoomat, tason II kasvaimiin mm. oligodendroglioomat, tason III kasvaimiin mm. anaplastiset oligodendroglioomat ja tason IV kasvaimiin glioblastoomat. Eri glioomia voidaan tunnistaa erilaisten molekulaaristen markkereiden perusteella. Eräs niistä on 1p/19q deleetio yhdistettynä mutaatioon, joka toimii markkerina oligodendroglioomalle. Tässä tutkimuksessa eräänä ryhmänä ovat ne näytteet, joista tuo deleetio yhdistettynä mutaatioon löytyy. Kuitenkin eri glioomatyypit, kuten oligodendroglioomat voivat kuulua markkereidensa perusteella eri tasolle (Louis ym. 2016). Esimerkiksi siinä, missä anaplastiset oligodendroglioomat kuuluvat tason III kasvaimiin, kuuluvat tavalliset oligodendroglioomat tasolle II.

Isositraattidehydrogenaasi (IDH) puolestaan toimii sitruunahappokierrossa entsyyminä isositraatin α -ketoglutaraattia tuottavassa reaktiossa. IDH:lla on useampi isoformi, mutta niistä vain IDH1:n ja IDH2:n mutaatiot on todettu yleisiksi glioomissa. Useimmissa tapauksissa mutatoituneena on IDH1:n 132. aminohappo tai analoginen IDH2:n R172 aminohappo (Yan ym. 2009). Nämä mutaatiot tapahtuvat usein jo syövän alkuvaiheessa, sillä ne ovat havaittavissa sekä tason II että III glioomissa (Yan ym. 2009) jotka muuten eroavat toisistaan.

Näitä kahta yhdistävänä tekijänä toimii se, että IDH:lla ja sen mutaatiolla glioblastoomissa on todettu yhteyksiä immuunijärjestelmään (Kohanbash ym. 2017). Immuunijärjestelmä vaikuttaa syövän leviämiseen, sillä usein immuunijärjestelmä tuhoaa tunnistamansa syöpäsolut. Siispä immuunijärjestelmän puutokset voidaan yhdistää suurempaan syöpäriskiin (Schreiber ym. 2011).

Tässä tutkimuksessa lähtöaineistona oli The Cancer Genome Atlas (TCGA) –projektista peräisin oleva ekspressiodata, jota tutkimalla selvitettiin, onko jonkin proteiinin tuotanto kasvanut vai vähentynyt IDH-mutaation seurauksena. Datasta löytyy tason II ja III glioomien, sekä tason IV glioblastoomien ekspressioita. Data on tuotettu koko genomilla laajuisella transkriptomin, eli kaikkien solun sen hetkisten mRNA molekyylien sekvensoinnilla. Ennen ekspressioanalyysiä datalle tehdään seuraavia työvaiheita. Ensimmäisenä sekvensointidatasta määritetään eri tavoin silmukoituneet RNA-sekvenssit. Tämä tehdään vertaamalla sekvensointidataa valittuun referenssigenomiin. Näin saaduista geenien lukemista määritetään seuraavaksi RNA-ekspressiotasot. Ekspressiotasojen määrittämisen jälkeen data vielä normalisoidaan haluttuun muotoon. Jokaisessa työvaiheessa on monia

eri tapoja ja vaihtoehtoisia menetelmiä, jotka tuottavat keskenään erilaista dataa. Tässä tutkimuksessa käynteille datalle tehdyt työvaiheet on esitelty luvussa 2.1.

2. MATERIAALIT JA MENETELMÄT

Tässä luvussa on esitetty materiaalille tehdyt työvaiheet alalukuihin jaettuna. Materiaalien alkuperä ja ennen tätä tutkimusta tehdyt työvaiheet on esitetty yleisesti johdannossa, mutta tässä käytetyt menetelmät esitellään tarkemmin.

2.1 TCGA datamatriisi

Ennen tätä tutkimusta aineistolle on tehty seuraavat työvaiheet (https://docs.gdc.cancer.gov/Data/Bioinformatics_Pipelines/Expression_mRNA_Pipeline/, 4.4.2020). Ensimmäisenä koko genomista laajuisesta transkriptomin sekvensoinnista saatu data on käännetty FASTQ-muotoon, jonka jälkeen siitä on eroteltu eri tavoin silmukoituneet RNA-sekvenssit vertaamalla niitä GRCh38 referenssigenomiin. Tämän jälkeen geenien lukemista on määritetty RNA:n ekspressiotasot HT-Seq-Count:ia apuna käyttäen. Lopuksi, jotta data on käytettävää, on se normalisoitu. Tässä analyysissä data on TPM (transcripts per million) -normalisoitua. Näitä menetelmiä käyttäen data on siis saatettu tässä analyysissä käytettävään muotoon.

Tässä tutkimuksessa suurin osa työstä on tehty R-ohjelmalla käyttäen sen peruspaketteja ja Matrix Tests -pakettia. Datamatriisissa on 57 133 geenin ekspressioarvot yhteensä 507 eri näytteestä. Näytteistä 204 kpl oli IDH-villityyppi (IDH-WT) -näytteitä, 111 kpl IDH mutaatio yhdistettynä deleetioon (IDH-CD) -näytteitä ja 192 kpl IDH-mutaatio ilman deleetiota (IDH-NCD) -näytteitä. Datalle oli myös taulukko, josta selvisi näytteen koodilla se, mihin näistä ryhmistä data kuuluu. Näytteistä 150 kpl oli GBM näytteitä ja 357kpl LGG näytteitä. LGG data on tason II ja III gliomien dataa, ja GBM puolestaan tason IV glioblastoomien dataa. Data on ladattu joulukuussa 2019 (<https://portal.gdc.cancer.gov/repository>, 18.4.2020).

2.2 Alkutestaus

Jotta saataisiin selville, onko TPM-normalisoitu data tähän analyysiin sopivaa, suoritettiin sille seuraavat alkutestaukset. Aluksi kaikki näytteiden arvot muutettiin ensin niiden kymmenkantaiseen logaritmuotoon. Sitten data jaettiin kahteen ryhmään, näytteisiin 1–255 ja 256–507, sillä kaikkia 507 näytettä ei pystytty laittamaan samaan kuvaajaan muodostuneen vektorin koon vuoksi. Näytteitä ja

niiden arvojen jakautumista tarkasteltiin myös näytetasolla. Näistä ryhmistä tehtiin Boxplot-kuvaajat, joista määritettiin poikkeamat ja todettiin, onko datan TPM-normalisaatio riittävä, vai pitääkö sille tehdä muutoksia. Mikäli merkittäviä poikkeamia olisi ollut, olisi dataa pitänyt käsitellä uudestaan, jotta se olisi oikein normalisoitunutta.

2.3 Tilastollinen testaus

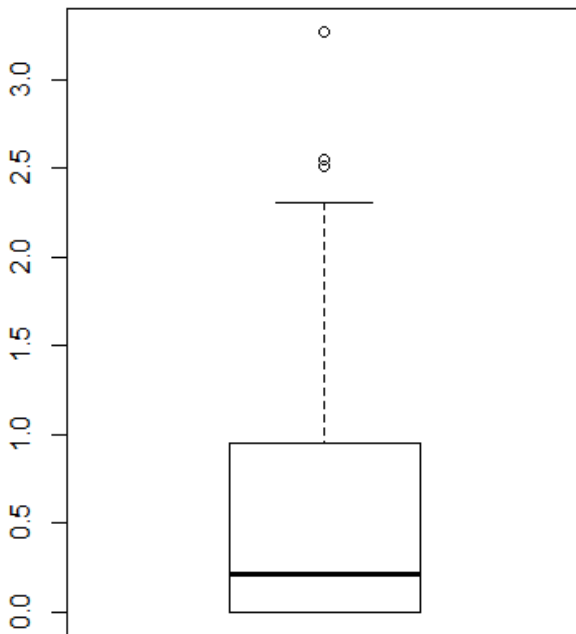
Tämän jälkeen alkuperäiset arvot muutettiin niiden kaksikantaisiksi logaritmeiksi. Sitten näytteet jaettiin niiden mutaatiotyypin perusteella uusiin matriiseihin, ensimmäisenä ryhmänä IDH-WT, toisena IDH-CD ja kolmantena IDH-NCD. Jokaisesta ryhmästä laskettiin jokaiselle geenille sen ekspression mediaani. Näitä geenien mediaaneja vertailtiin ryhmissä toisiinsa laskemalla niiden välisen eron itseisarvo. Ensin vertailu tehtiin ryhmille, joissa mediaanien välinen ero oli yli 1,5. Kuitenkin jotta vertailuun saataisiin enemmän geenejä mukaan ja sitä kautta reittejä paremmin näkyviin, vertailtiin myös ryhmiä, joissa mediaanien välinen ero oli yli 1. Nämä ryhmien väliset erot vastaavat 2- ja 2,83-kertaisia ekspressioeroja lineaarisella asteikolla. Saatuja geeniryhmiä vertailtiin pareittain t-testien avulla. Saaduille tuloksille tehtiin Bonferronin korjaukset (p -arvo $< 0,0005$). Näiden koodit löytyvät liitteistä. T-testin ja sen korjauksen avulla saatiin selville ne geenit, joilla on tilastollisesti merkittävä ero kahden vertailun ryhmän välillä. Käytetty MatrixTests-paketti mahdollisti sarakkeiden väliset t-testit.

2.4 KEGG ja GO reittianalyysit

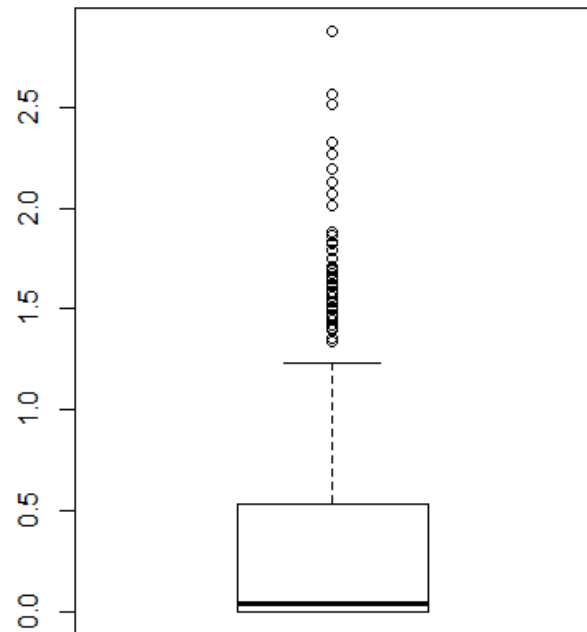
Jokaisen ryhmän lopulliset geenit eroteltiin muista, ja ryhmille tehtiin rikastusanalyysi Enrichr-verketyökalulla (Chen ym. 2013). Enrichr vertailee annettua geenilistaa sen tietokannan geeni- ja proteiinilistoihin. Tällä vertailulla saadaan rikastettua näkyviin ne reitit, joihin annetun geenilistan geenit vaikuttavat. Rikastusanalyysissä geenejä verrattiin normaaliin ihmisen geenitaustaan. Reittianalyysistä tutkittiin geenien vaikutuksia Gene Ontology (GO) 2018-annotaatioiden ja KEGG 2019 –reittianalyysin avulla. GO-annotaatiot kertovat eri tasoilla sen, mihin ekspressioerot solussa vaikuttavat. Tasoja on kolme ja ne ovat biologiset prosessit, molekulaariset funktiot ja solun osat. Tässä analyysissä tutkittiin kaikkia kolmea, jotta saataisiin mahdollisimman hyvä kuva siitä miksi ja miten ekspressioerot vaikuttavat immunitettiin. Kuitenkin rikastusanalyysissä osa geeneistä jäi pois, sillä ne tunnistettiin pseudogeeneiksi, joille ei tiedetä mitään funktiota. Siispä ne eivät edes olisi vaikuttaneet reittianalyysin tuloksiin. Saaduista taulukoista tutkittiin korjattujen p -arvojen ja rikastuneiden geenien määrien avulla löytyykö joukosta immuunijärjestelmään vaikuttavia tekijöitä. Myös rikastuneiden geenien määrä suhteessa kaikkiin reittiin vaikuttaviin geeneihin huomioitiin.

3. TULOKSET

Materiaalit ja menetelmät –kohdassa esitellyn alkutarkistuksen jälkeen saatujen kuvaajien yhteenveto on esitelty kuvissa 1 ja 2. Koska merkittäviä poikkeamia ei ollut, voitiin tutkimusta jatkaa dataa muokkaamatta.



Kuva 1: Kuvassa on boxplot, joka sisältää näytteiden 256–507 kymmenkantaisiksi logaritmeiksi muutetut arvot.



Kuva 2: Kuvassa on boxplot, joka sisältää näytteiden 1–255 kymmenkantaisiksi logaritmeiksi muutetut arvot.

Koska data on kymmenkantaisessa logaritmuodossa, ovat sen arvot siksi erilaisia työn muihin kaksikantaisiin arvoihin verrattuna. Tämä ei kuitenkaan vaikuta lopputulokseen.

Seuraavassa taulukossa (taulukko 1) on esitetty jokaisen vertailuryhmän geenien määrä. Taulukossa olevat geenit on jo jaettu niihin kolmeen ryhmään ja pareihin, joissa niitä vertaillaan. Tilastol-

Taulukko 1: Taulukossa on esitetty vertailuryhmien geenien määrät

X kertainen muutos	vertailuryhmä	geenien määrä
1	CD-NCD	403
1,5	CD-NCD	85
1	WT-CD	3615
1,5	WT-CD	1645
1	WT-NCD	2464
1,5	WT-NCD	985

listen testausten yhteydessä ei yhtäkään geeniä jouduttu poistamaan listasta, sillä kaikki alittivat asetetun p-arvon.

Taulukko 2: Taulukossa on esitetty KEGG-reittien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
CD-NCD	Complement and coagulation cascades	5,13E-07	7,90E-05	11/79
CD-NCD	Platelet activation	9,26E-04	3,56E-02	9/124
CD-NCD	Phagosome	1,05E-03	3,58E-02	10 /152
WT-CD	Phagosome	1,12E-08	1,15E-06	57 /152
WT-CD	Complement and coagulation cascades	6,36E-08	3,27E-06	35/79
WT-CD	Leukocyte transendothelial migration	6,67E-06	1,37E-04	40/112
WT-NCD	Complement and coagulation cascades	1,57E-04	6,92E-03	22/79
WT-NCD	Phagosome	7,74E-04	2,17E-02	33/152

Kaikissa seuraavissa taulukoissa on esitetty vain suoraan immunologiaan liittyvät reitit. Myös muita reittejä nousi esille, mutta ne eivät olleet suoraan liitettävissä immunologiaan. Näiden reittien lukumäärä on ilmoitettu taulukon yhteydessä. Taulukoissa näkyy ne reitit, joihin eri tavoin ekspressoituvat geenit vaikuttavat. Taulukoissa näkyy myös reitin p-arvot, sekä niiden geenien määrä joiden ekspressio on muuttunut suhteessa kaikkiin reitin geeneihin. Taulukossa 2 on esitetty KEGG-reittianalyysin tulokset ryhmille, joissa vertailtavien ryhmien välinen muutos on 1. KEGG-reittianalyysi antaa tuloksena ne metabolian ja soluprosessien reitit, joihin valikoidut geenit vaikuttavat. Reittejä, jotka eivät immunologiaan liittyneet ja joiden p-arvo oli alle 0,05 oli WT-CD vertailussa 299 kappaletta, WT-NCD vertailussa 293 kpl ja CD-NCD vertailussa 200 kpl.

Taulukko 3: Taulukossa on esitetty KEGG-reittien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1,5.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
WT-CD	Complement and coagulation cascades	2,01E-09	2,07E-07	25/79
WT-CD	Phagosome	1,77E-08	9,09E-07	35/152
WT-NCD	-	-	-	-
CD-NCD	Cytokine-cytokine receptor interaction	3,67E-02	9,41E-01	4/294

Taulukossa 3 on esitetty KEGG-reittianalyysin tulokset ryhmille, joissa vertailtavien ryhmien välinen muutos on 1,5. WT-NCD vertailussa ei saatu ollenkaan immunologiaan liittyviä tuloksia, sillä reitillä ei ollut tarpeeksi geenejä. CD-NCD vertailussa reitin p-arvo oli suurempi kuin 0,05, mutta tulos otet-

tiin kuitenkin huomioon ja se on esitetty taulukossa. Reittejä, jotka eivät liity immunologiaan ja joiden p-arvo jäi liian alhaiseksi, jotta ne olisi otettu varsinaisiin tuloksiin mukaan oli WT-CD vertailussa 279 kappaletta, WT-NCD vertailussa 258 kpl ja CD-NCD vertailussa 90 kpl.

Taulukko 4: Taulukossa on esitetty GO-annotaation biologisten prosessien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
CD-NCD	positive regulation of phagocytosis (GO:0050766)	9,03E-07	4,61E-03	8/39
CD-NCD	positive regulation of myeloid leukocyte mediated immunity (GO:0002888)	1,89E-05	2,41E-02	4/9
CD-NCD	positive regulation of leukocyte degranulation (GO:0043302)	4,79E-05	4,89E-02	4/11
CD-NCD	positive regulation of apoptotic cell clearance (GO:2000427)	5,42E-06	1,38E-02	4/7
WT-CD	cytokine-mediated signaling pathway (GO:0019221)	4,28E-12	1,09E-08	184/633
WT-CD	interferon-gamma-mediated signaling pathway (GO:0060333)	9,49E-12	1,61E-08	38/70
WT-CD	neutrophil mediated immunity (GO:0002446)	2,59E-10	3,31E-07	144/487
WT-CD	neutrophil activation involved in immune response (GO:0002283)	2,71E-10	2,77E-07	143/483
WT-CD	neutrophil degranulation (GO:0043312)	2,84E-10	2,41E-07	142/479
WT-CD	cellular response to cytokine stimulus (GO:0071345)	1,18E-08	8,60E-06	131/456
WT-CD	inflammatory response (GO:0006954)	2,06E-07	9,56E-05	79/252
WT-CD	response to cytokine (GO:0034097)	2,01E-06	6,82E-04	48/138
WT-NCD	cytokine-mediated signaling pathway (GO:0019221)	3,76E-06	2,40E-03	117/633
WT-NCD	cellular response to type I interferon (GO:0071357)	1,98E-05	1,01E-02	21/65
WT-NCD	type I interferon signaling pathway (GO:0060337)	1,98E-05	9,19E-03	21/65

Taulukossa 4 on esitetty GO-annotaation biologisten prosessien tulokset ryhmille, joissa niiden välinen muutos on 1. GO-annotaatio painottuu geeneihin ja niiden tuotteisiin, jonka takia sitä on käytetty tässä analyysissä KEGG-reittianalyysin lisäksi. Näin on saatu kattavat tulokset, joiden pohjalta asiaa tulkita. Taulukosta näkyvät ne biologiset prosessit, joihin geenien on todettu liittyvän. Tulosten

avulla pystytään tutkimaan mihin reitteihin IDH-mutaatio vaikuttaa. Reittejä, joiden p-arvo jäi liian alhaiseksi, tai ei liittynyt immunologiaan oli WT-CD vertailussa 4506 kappaletta, WT-NCD vertailussa 4064 kpl ja CD-NCD vertailussa 540 kpl.

Taulukko 5: Taulukossa on esitetty GO-annotaation biologisten prosessien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1,5.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
WT-CD	platelet degranulation (GO:0002576)	5,60E-08	7,14E-05	30/124
WT-CD	cytokine-mediated signaling pathway (GO:0019221)	1,02E-07	1,04E-04	91/633
WT-CD	neutrophil degranulation (GO:0043312)	1,96E-07	1,67E-04	73/479
WT-CD	neutrophil activation involved in immune response (GO:0002283)	2,73E-07	1,99E-04	73/483
WT-CD	interferon-gamma-mediated signaling pathway (GO:0060333)	5,28E-07	2,69E-04	20/70
WT-CD	cellular response to cytokine stimulus (GO:0071345)	1,14E-06	5,28E-04	68/456
WT-NCD	cytokine-mediated signaling pathway (GO:0019221)	8,34E-06	6,08E-03	57/633
CD-NCD	positive regulation of toll-like receptor 3 signaling pathway (GO:0034141)	2,52E-02	1,00E+00	1/6

Taulukossa 5 on esitetty GO-annotaation biologisten prosessien tulokset ryhmille, joissa niiden välinen muutos on 1,5. Koska valikoituneet geenit oli tiukemmin rajattu, oli reittejä kaiken kaikkiaan myös huomattavasti vähemmän. Kuitenkin kaikista pareista löytyi tuloksia, joten reitti ei ollut liian tiukasti rajattu. Rajaus auttoi löytämään joukosta helpommin ne reitit, joilla ero ekspressiossa on suurempi, joskaan tämä ei aina korreloi biologisen merkittävyyden kanssa. Reittejä, jotka eivät joko liittyneet immuunijärjestelmään tai joilla ei ollut tarpeeksi suuri p-arvo oli WT-CD vertailussa 3575 kappaletta, WT-NCD vertailussa 2899 kpl ja CD-NCD vertailussa 543 kpl.

Taulukossa 6 on annettu GO-annotaation molekulaaristen toimintojen tulokset ryhmille, joiden välinen ero on 1. Tulokset kertovat mihin geenit vaikuttavat molekulaarisella tasolla. Siispä IDH-mutaatio sekä 1p/19q deleetion vaikutuksia mm. erilaisiin proteiineihin on hyvä tutkia myös tämän GO-

Taulukko 6: Taulukossa on esitetty GO-annotaation molekulaaristen toimintojen tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
CD-NCD	G-protein coupled chemoattractant receptor activity (GO:0001637)	3,09E-03	3,95E-01	3/15
WT-CD	chemokine binding (GO:0019956)	2,06E-04	2,16E-02	10/17
WT-NCD	-	-	-	-

annotaation avulla. WT-NCD vertailussa suoraan immunologiaan liittyviä reittejä ei tullut esiin tiukahkon rajauksen takia. Kaiken kaikkiaan pois jääneitä reittejä oli WT-CD vertailussa 971 kappaletta, WT-NCD vertailussa 884 kpl ja CD-NCD vertailussa 367 kpl.

Taulukko 7: Taulukossa on esitetty GO-annotaation molekulaaristen toimintojen tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1,5.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
WT-CD	Toll-like receptor binding (GO:0035325)	3,56E-04	2,73E-02	5/9
WT-NCD	-	-	-	-
CD-NCD	cytokine receptor binding (GO:0005126)	2,07E-02	1,00E+00	3/137
CD-NCD	cytokine activity (GO:0005125)	2,84E-02	1,00E+00	3/155

Taulukossa 7 on annettu GO-annotaation molekulaaristen toimintojen tulokset ryhmille, joiden välinen ero on 1,5. Kuten edellisessä, myöskään tässä WT-NCD vertailussa suoraan immunologiaan liittyviä reittejä ei tullut esiin, sillä tässä käytetty rajausta ekspression suhteen oli vielä tiukempi. Tämä ei kuitenkaan aina korreloi sen kanssa, ettei sellaisia biologisesti merkittäviä reittejä olisi joihin geenin toiminta molekulaarisella tasolla vaikuttaa. Taulukosta pois jääneitä geenejä oli WT-CD vertailussa 716 kappaletta, WT-NCD vertailussa 573 kpl ja CD-NCD vertailussa 141 kpl.

Taulukko 8: Taulukossa on esitetty GO-annotaation solunsisäisten komponenttien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
CD-NCD	tertiary granule (GO:0070820)	5,08E-04	3,24E-02	11/164
CD-NCD	tertiary granule membrane (GO:0070821)	6,65E-04	3,29E-02	7/73
CD-NCD	specific granule membrane (GO:0035579)	2,43E-03	1,08E-01	7/91
WT-CD	tertiary granule (GO:0070820)	1,63E-05	1,04E-03	52/164
WT-CD	MHC class II protein complex (GO:0042613)	1,69E-04	5,39E-03	9/14
WT-CD	phagocytic vesicle (GO:0045335)	1,83E-04	5,45E-03	29/83
WT-CD	tertiary granule membrane (GO:0070821)	2,74E-04	7,62E-03	26/73
WT-NCD	MHC class II protein complex (GO:0042613)	6,65E-04	2,97E-02	7/14
WT-NCD	MHC protein complex (GO:0042611)	3,94E-03	9,77E-02	7/18

Taulukossa 8 on esitetty tulokset GO-annotaation solunsisäisille komponenteille. Tulokset kertovat missä osassa solua geenien tuottamat proteiinit toimivat. Taulukon ryhmien välinen ero on 1. Taulukossa esitetyn luokan II MHC-kompleksin ekspression todettiin olevan korkeampaa IDH-WT tyyppin kasvaimissa IDH-CD ja IDH-NCD tyyppin kasvaimiin verrattuna. Taulukosta pienen p-arvon tai immunologiaan liittymättömyyden takia pois jääneitä reittejä oli WT-CD vertailussa 371 kappaletta, WT-NCD vertailussa 330 kpl ja CD-NCD vertailussa 134 kpl.

Taulukko 9: Taulukossa on esitetty GO-annotaation solunsisäisten komponenttien tulokset IDH-WT, IDH-CD ja IDH-NCD ryhmille, joissa vertailtavien ryhmien välinen muutos on 1,5.

Vertailuryhmä	Reitin nimi	p-arvo	Korjattu p-arvo	Geenien määrä per reitin koko
WT-CD	MHC class II protein complex (GO:0042613)	3,94E-06	3,51E-04	8/14
WT-CD	MHC protein complex (GO:0042611)	4,25E-05	1,72E-03	8/18
WT-NCD	MHC class II protein complex (GO:0042613)	3,00E-05	2,68E-03	6/14
WT-NCD	MHC protein complex (GO:0042611)	1,56E-04	7,75E-03	6/18
CD-NCD	-	-	-	-

Taulukossa 9 on esitetty tulokset GO-annotaation solunsisäisille komponenteille, ja taulukon vertailuryhmien välinen ero on 1,5. CD-NCD vertailussa ei suoraan immunologiaan yhdistettäviä reittejä löytynyt ollenkaan tiukan rajauksen takia. Tiukka rajausta vaikutti myös muiden vertailuryhmien vähäisempiin tuloksiin. Kuitenkin tulosten perusteella voitiin suoraan todeta luokan II MHC-kompleksin merkittävyys. Taulukosta pois jääneitä reittejä oli WT-CD vertailussa 278 kappaletta, WT-NCD vertailussa 221 kpl ja CD-NCD vertailussa 68 kpl.

4. PÄÄTELMÄT

4.1 Tulosten oikeellisuus

Näistä tuloksista voidaan todeta, että tilastollinen yhteys syöpäkudoksen immuunijärjestelmän toiminnan ja IDH-mutaation välillä löytyy. Vaikka GO-annotaation kaikissa ryhmissä suoria yhteyksiä ei löytynyt, immuunijärjestelmään vaikuttaa myös siihen epäsuorasti liittyvät asiat, kuten solu-soluliitokset ja solun pintaproteiinit.

Alkutarkastuksesta huomataan, että merkittäviä poikkeamia datassa ei näytetasolla ole. Siispä dataa on käytetty sellaisenaan. Data on ryhmissä hieman eri tavoin jakautunutta, sillä kuvassa 2 $\frac{3}{4}$ datasta on jakautunut lähemmäs nollaa verrattuna kuvaan 1. Kuitenkaan tämä ei sinällään ole merkittävää, sillä arvot pysyvät samoissa mittaluokissa. Tuloksiin saattaa kuitenkin vaikuttaa se, että käytetty kriteeri ekspressiomuutoksille oli hyvin tiukka. Tämä nähdään siitä, ettei t-testillä ja sen korjauksella yksikään geeni pudonnut aineistosta.

KEGG-reiteiltä saatiin melko erilaiset tulokset niiden ryhmien välillä, joissa mediaanien välinen ero oli 1 verrattain ryhmään, joissa ero oli 1,5. Taulukossa 2 WT-CD ja WT-NCD vertailuryhmissä tulokset vaikuttavat p-arvojensa ja geenimääriensä puolesta oikeellisilta. CD-NCD vertailuryhmän biologista merkittävyyttä voidaan kyseenalaistaa, sillä siinä eri tavoin ekspressoituneiden geenien määrät ovat vähäisiä reittien kokoihin verrattuna. Fagosomireitillä on ekspressioero vain 10 geenissä, vaikka reitillä on 152 geeniä. Toki, mikäli nämä geenit ovat kriittisiä reitin toiminnan kannalta, voi toiminta muuttua. Kuitenkin koska alle 10 % reitin geneistä on muuttunut, ei mutaatio tässä suhteessa välttämättä ole niin merkittävä p-arvosta huolimatta. Taulukossa 3 esitetyissä tuloksissa on samaa ongelmaa. CD-NCD vertailussa vain korjaamaton p-arvo osuu tilastollisen merkittävyyden piiriin, sillä korjattu p-arvo sekä geenien määrä molemmat tekevät tuloksen merkittävyydestä oikeastaan melko epäluotettavan. Koska kyseisessä sytokiini-sytokiini reseptori-interaktioreitissä on 294 geeniä, joista tässä tapauksessa olisi mutatoitunut vain neljä, vaikuttaa tulos melko merkityksettömältä biologisella

tasolla. WT-NCD vertailusta ei saatu ollenkaan immunologiaan liittyviä tuloksia, joten selkeästi taulukossa 2 esitetyt Komplementti- ja koagulaatiokaskadireitin ja Fagosomireitin geeniekspressioiden muutos ei ole niin iso, että se näkyisi myös toisessa taulukossa. WT-CD vertailun tuloksia taulukossa 3 voidaan kuitenkin myös pitää oikeellisina p-arvojen ja geenien määrien suhteen.

Mielenkiintoista on, että taulukossa 3 näkyvää sytokiini-sytokiini reseptori interaktioreittiä ei näy taulukossa 2. Todennäköistä on, että koska reitillä on rikastunut vain neljä geeniä, ei reitin p-arvo ole noussut tarpeeksi suureksi, jotta se olisi otettu mukaan taulukkoon 2. GO-annotaation biologisten reittien tarkastelussa taulukossa 4 esitetyjä vertailuja ja niihin liittyviä reittejä voidaan pitää tilastollisesti merkittävänä korjattujen p-arvojensa ja geenimääriensä puolesta. Kaikkien taulukon tulosten voidaankin siis olettaa olevan oikeellisia.

Taulukossa 5 miltei kaikki tulokset ovat p-arvojensa puolesta tilastollisesti merkittäviä. Kuitenkin CD-NCD vertailussa esille nousseen Toll reseptori 3:n positiiviseen säätelyyn vaikuttavan reitin korjattu p-arvo on 1, eli siis ei tilastollisesti merkittävä. Kuitenkin tavallinen p-arvo osoittaisi sen olevan tilastollisesti merkittävä ja lisäksi reitillä on vain kuusi geeniä. Siispä voisi ajatella, että kun vain kuudesta geenistä yhden toiminta muuttuu, on sillä merkitystä reitin kannalta. Puolestaan WT-CD reitillä on useampi reitti, johon liittyy satoja geenejä (sytokiinisäädely signaalintireitti 633 geeniä, neutrofiilien degranulaatio 479 geeniä ja neutrofiilien aktivaatio 483 geeniä), jolloin eri tavoin ekspressoituneiden geenien määrän on oltava myös suurempi. Tämän vuoksi koska reiteissä on paljon enemmän geenejä, joiden ekspressio on muuttunut, joten reittien toiminnan voitaisiin olettaa muuttuneen.

GO-annotaation molekulaaristen reittien suhteen taulukossa 6 CD-NCD vertailussa korjattu p-arvo ei ole tilastollisesti merkittävä. Kuitenkin, koska tavallinen p-arvo olisi tilastollisesti merkittävä ja reitillä on vain 15 geeniä, joista kolme on mutatoitunut, voisi sillä kuitenkin ajatella olevan vaikutusta kemokiinien toimintaan ja siten immuunijärjestelmään. Taulukon 7 tulosten biologista merkitystä voidaan epäillä. Koska vain pienessä osassa geenejä on ekspresioeroja, vain WT-CD vertailun voidaan olettaa olevan oikeellinen. CD-NCD vertailuissa molemmat korjatut p-arvot ovat 1 ja yli sadan geenin reitissä vain kolmen geenin ekspressio on muuttunut. Siispä voidaan ajatella, että nämä geenit tuskin vaikuttavat reitin toimintaan merkittävästi.

Mitä tulee solunsisäisiin komponentteihin, taulukossa 8 olevien CD-NCD vertailun merkitys voidaan kyseenalaistaa. Vaikka vain spesifisen jyväsän membraaniin vaikuttavan reitin korjattu p-arvo ei ole tilastollisesti merkittävä, ei missään CD-NCD vertailussa esille tulleessa reitissä ollut paljoa geenejä suhteessa reitin kokoon. Voidaankin siis olettaa, etteivät CD-NCD vertailun tulokset välttämättä ole biologisesti merkittäviä. Taulukossa 8 olevien WT-CD ja WT-NCD tulosten voidaan kuitenkin olettaa olevan oikein niin p-arvojen kuin geenien määrien suhteen. Taulukossa 9 on samaa vikaa kuin taulukossa 8. CD-NCD vertailun tulokset tuskin siinäkään ovat biologisesti merkittäviä korjatun p-arvon

ja geenien määrän perusteella. WT-CD vertailussa esille tulleen Toll-reseptorin sitoutumisreitin voidaan kuitenkin olettaa olevan oikeellinen p-arvojen ja geenien määrän perusteella.

Tuloksista voidaan myös nähdä, että KEGG-reitit antoivat paljon yleisemmät ja laajemmat reitit, kun taas GO-annotaatio spesifimmät. Voidaankin todeta, että tämän tutkimuksen kannalta GO-annotaatiotulokset ovat hyödyllisempiä.

4.2 Biologiset merkitykset

IDH-mutaatiolla on todettu yhteys T-soluihin, tarkemmin CD8+ tappaja T-soluihin (Kohanbash ym. 2017). Kuitenkaan tämä yhteys ei nouse tuloksissa suoraan reittinä esille. Tämä voi johtua siitä, ettei datassa ole mukana siihen vaikuttavia geenejä tai siitä, ettei niiden ekspressioero ollut tarpeeksi suuri, jotta ne olisivat erottuneet mukaan niihin geeneihin, jotka ajettiin rikastusanalyysistä läpi. Kuitenkin Kohanbash ym. ovat tutkimuksessaan todenneet, että LGG-tyypin glioomissa IDH-mutaatiot johtavat matalampaan CD8+ sytotoksisten T-solujen kudoksen infiltraatioon. IDH-mutaatio johtaa soluissa pienempään interferoni- γ (IFN- γ) indusoituvien kemokiinien ekspressioon verrattuna IDH-WT tapauksiin. IFN- γ on tyypin II interferoni, joka tulehduksen aikana lisää makrofagiin aktivaatiota sekä HLA I ja II luokkien ilmentymistä. Tämä parantaa antigeeniä ilmentävien solujen kykyä esitellä vieraita antigeenejä T-soluille. IFN- γ voi myös estää Th2-tyyppisten lymfosyyttien proliferaatiota. IFN- γ :n ekspression muutos IDH-WT ja IDH-CD ryhmien välillä voidaankin todeta tämän tutkimuksen tulosten pohjalta. Reitti näkyy taulukoissa 4 ja 5.

Koska ero on noussut esille WT-CD ja WT-NCD vertailuissa, ei 1p/19q deleetion siis pitäisi suuresti muuttaa tilannetta. Ekspressioeron onkin todettu aiheutuvan IDH-mutaatiosta (Kohanbash ym. 2017). Siksi ekspressioero ei tule CD-NCD ryhmien vertailussa esille. Kyseisten IFN- γ :n indusoivien kemokiinien ekspressio siis laskee sekä CD- että NCD-ryhmissä. GO-annotaatiokoodin (<https://www.ebi.ac.uk/QuickGO/term/GO:0060333>, viitattu 28.3.2019) perusteella voidaan tutkia ja todeta, että tämä IFN- γ :n ekspressiomuutos vaikuttaa myös muihin sen reitin annotaatioihin, jotka myös tulevat esille taulukoissa 4 ja 5. Näitä reittejä ovat mm. sytokiinisäädely signaalintireitti ja vaste sytokiineille. INF- γ :n lisäksi WT-NCD vertailussa myös tyypin 1 interferonien signaalintireitillä on tapahtunut ekspressiomuutos. Tämä voisi viitata siihen, etteivät myöskään IFN- α ja IFN- β signaloitu normaalisti. Syöpäkasvaimissa kilpailu ravinteista on kovaa, ja tämä vaikuttaa erityisesti siihen, kuinka paljon solut saavat sokeria käytettäväkseen. On todettu, että kun T-solut eivät saa tarpeeksi sokeria, eivät ne pysty kunnolla tuottamaan IFN- γ :aa (Chang ym. 2015). CD ja NCD ryhmissä todettu IFN- γ –ekspression lasku voidaan siis osittain olettaa johtuvan myös T-solujen sokerin puutoksesta syöpäkasvaimessa IDH-mutaation lisäksi.

Koska IFN- γ vaikuttaa osaltaan CD8+ T-solujen toimintaan, vaikuttavat ne myös ihmisen immuunivasteeseen. Tämä on todettavissa taulukosta 4, sillä siellä havaittuna on ekspressioero tulehdus-

vasteessa IDH-WT ja IDH-CD ryhmien välillä. Esille noussut neutrofiilien aktivaatio viittaa myös immunijärjestelmän muutokseen. Kuitenkaan IFN- γ :n aktivaatio ei selitä tätä, sillä se aktivoi makrofageja, ei neutrofiileja.

Muutoksia ekspressiotasoissa on myös MHC-kompleksin suhteen, kuten taulukoista 7 ja 8 huomataan. IFN- γ :n ekspression lasku voi osaltaan johtua siitä, että luokan I MHC:n ekspressio laskee. Tästä kielii se, että kun näitä ekspressiotasoja keinotekoisesti nostetaan (Szatmari ym. 2006), saadaan T-soluja stimuloitua. Tällöin loogisesti luokan I MHC-kompleksien määrän lasku vaikuttaisi T-solujen vähenemiseen. Kuitenkaan tämä ei ole taulukoissa nähtävissä, sillä ekspressioerot ovat MHC luokan II geneeissä, vaikka muutos ekspressiossa on ilmoitettu koko MHC-kompleksille. Tämä nähdään siitä, että kyseessä on samat geenit molemmissa reiteissä ja siitä, että MHC-kompleksi on yläluokkana luokan II MHC-kompleksille. Luokan II MHC proteiini toimii yhdessä CD4+ auttaja-T-solujen kanssa, ei CD8+ tappaja-T-solujen. Auttaja-T-solut mm. erittävät sytokiineja. Siispä taulukossa 4 ja 5 esiintyvät erot sytokiinireiteissä voisivat olla yhteydessä juuri tähän tason II MHC ekspressioeroon. Hussain ym. (Hussain ym. 2006) tutkimuksessa on osoitettu, että kun luokan II MHC ekspressio on kasvanut, on CD4+ auttaja T-solujen todettu tunkeutuvan paremmin syöpäkasvaimen samalla kun CD8+ tappaja-T-solun tunkeutuminen vähenee. Koska tuloksissa todettiin luokan II MHC-kompleksin olevan korkeammin ekspressoitunut IDH-WT tyyppin kasvaimessa, voitaisiin olettaa, että IDH-mutaatio vähentää luokan II MHC-kompleksien ekspresoitumista, ja siten CD4+ auttaja T-solujen pääsyä syöpäkudokseen. Tällöin myöskään muita valkosoluja paikalle houkuttelevia kemokiineja ei erity. Tämä olisi loogista ottaen huomioon, että IDH-mutaatio vähentää myös CD8+ tappaja T-solujen määrää IFN- γ :n vähenemisen kautta.

Kun kaikkia ryhmiä vertaillaan keskenään, voidaan huomata, että rikastuneet reitit ovat hyvin erilaisia vertailtavien ryhmien kesken. Tämä osaltaan vahvistaa mm. Berghoffin (Berghoff ym. 2017) huomaamaa hypoteesia siitä, että tulehduksen mikroympäristö vaihtelee sen mukaan, millainen glioman alatyypin on kyseessä. Mielenkiintoista kuitenkin on, että taulukossa 2 kaikissa kolmessa ryhmässä KEGG-reittien ryhmävertailussa nousi esille komplementti- ja koagulaatiokaskadit. Tästä voidaan päätellä, että komplemettijärjestelmän ekspressio on kaikkien kolmen ryhmän välillä eri tasolla – joissakin se on matalampi kuin toisissa, toisissa korkeampi. Selkein ero näkyy taulukossa 3: komplementti ja koagulaatiokaskadit ovat antaneet tuloksen vain WT-CD ryhmille. Korkein ekspressioero löytyy siis sieltä.

CD-NCD ryhmien välillä tilastollisesti merkittäviä ekspressioeroja löytyy myeloidisten leukosyyttien säätelemästä immunitetistä ja leukosyyttien degranulaatiosta. Onkin siis mahdollista sanoa näiden tulosten perusteella, että 1p/19q mutaatio voisi vaikuttaa leukosyyttien toimintaan tulehduksessa. Myös ei-tilastollisesti merkittävät tulokset, kuten Toll-reseptori 3 signalointiin vaikuttava reitti, voisi tässä tapauksessa vaikuttaa juuri tuohon leukosyyttien säätelemään immunitettiin.

Yhteenvetona voidaan todeta IDH-mutaatiolla olevan vaikutusta syöpäkasvaimen ja sitä ympäröivän kudoksen immuunijärjestelmään. Mutaation on todettu vähentävän IFN- γ :n ekspressoitumista, mikä johtaa CD8+ tappaja T-solujen määrän vähenemiseen kudoksessa. Myös luokan II MHC-kompleksien määrä vähenee IDH-WT tyyppin kasvaimen verrattuna, mikä johtaa myös vähäisempään CD4+ auttaja T-solujen määrään kudoksessa. Mutaation lisäksi myös syövän mikroympäristö voi vaikuttaa näiden solujen määrän vähenemiseen, sillä syöpäkudoksessa kilpailu ravinteista on kovaa. Se miten 1p/19q deleetio vaikuttaa immuunijärjestelmään ei ole yhtä selvää, mutta tuloksista voidaan todeta sen vaikuttavan kudoksen tulehdusvasteeseen leukosyyttien toiminnan kautta.

LÄHTEET

- Berghoff A.S., Kiesel B., Widhalm G., Wilhelm D., Rajky O., Kurscheid S., Kresl P., Wöhrer A., Marosi C., Hegi M.E. & Preusser M. (2017) Correlation of immune phenotype with IDH mutation in diffuse glioma. *Neuro-Oncology* **19**(11), 1460-1468.
- Bleeker F., Molenaar R. & Leenstra S. (2012) Recent advances in the molecular understanding of glioblastoma. *Journal of Neuro-Oncology* **108**(1), 11-27.
- Chang C., Qiu J., O'Sullivan D., Buck M., Noguchi T., Curtis J., Chen Q., Gindin M., Gubin M., van der Windt G.W., Tonc E., Schreiber R., Pearce E. & Pearce E. (2015) Metabolic Competition in the Tumor Microenvironment Is a Driver of Cancer Progression. *Cell* **162**(6), 1229-1241.
- Chen E.Y., Tan C.M., Kou Y., Duan Q., Wang Z., Meirelles G.V., Clark N.R. & Ma'ayan A. (2013) Enrichr: interactive and collaborative HTML5 gene list enrichment analysis tool. *BMC Bioinformatics* **14**(1), 128.
- Hartmann C., Hentschel B., Wick W., Capper D., Felsberg J., Simon M., Westphal M., Schackert G., Meyermann R., Pietsch T., Reifenberger G., Weller M., Loeffler M. & von Deimling A. (2010) Patients with IDH1 wild type anaplastic astrocytomas exhibit worse prognosis than IDH1-mutated glioblastomas, and IDH1 mutation status accounts for the unfavorable prognostic effect of higher age: implications for classification of gliomas. *Acta Neuropathologica* **120**(6), 707-718.
- Hussain S.F., Yang D., Suki D., Aldape K., Grimm E. & Heimberger A.B. (2006) The role of human glioma-infiltrating microglia/macrophages in mediating antitumor immune responses¹. *Neuro-Oncology* **8**(3), 261-279.
- Kohanbash G., Carrera D.A., Shrivastav S., Ahn B.J., Jahan N., Mazor T., Chheda Z.S., Downey K.M., Watchmaker P.B., Beppler C., Warta R., Amankulor N.A., Herold-Mende C., Costello J.F. & Okada H. (2017) Isocitrate dehydrogenase mutations suppress STAT1 and CD8+ T cell accumulation in gliomas. *The Journal of Clinical Investigation* **127**(4), 1425-1437.
- Louis D., Perry A., Reifenberger G., von Deimling A., Figarella-Branger D., Cavenee W., Ohgaki H., Wiestler O., Kleihues P. & Ellison D. (2016) The 2016 World Health Organization Classification of Tumors of the Central Nervous System: a summary. *Acta Neuropathologica* **131**(6), 803-820.

Schreiber R.D., Old L.J. & Smyth M.J. (2011) Cancer Immunoediting: Integrating Immunity's Roles in Cancer Suppression and Promotion. *Science (Washington)* **331**(6024), 1565-1570.

Szatmari T., Lumniczky K., Desaknai S., Trajcevski S., Hidvegi E.J., Hamada H. & Safrany G. (2006) Detailed characterization of the mouse glioma 261 tumor model for experimental glioblastoma therapy. *Cancer Science* **97**(6), 546-553.

Yan H., Parsons D.W., Jin G., McLendon R., Rasheed B.A., Yuan W., Kos I., Batinic-Haberle I., Jones S., Riggins G.J., Friedman H., Friedman A., Reardon D., Herndon J., Kinzler K.W., Velculescu V.E., Vogelstein B. & Bigner D.D. (2009) IDH1 and IDH2 Mutations in Gliomas. *The New England Journal of Medicine* **360**(8), 765-773.

LIITTEET

Koodit

T-testi ja Bonfferonin korjaus. Esimerkkinä CD-NCD ryhmä, muut vertailut tehty samalla koodilla.

```
> set.seed(1)
> alpha <- 0.05
> m <- 10000
> N <- 12
> pvals <- replicate(m, {
+   IDHmCD = sample(logIDHmutCoDel, N)
+   IDHmNCD = sample(logIDHmutNonCoDel, N)
+   t.test(IDHmCD, IDHmNCD)$p.value})

> sum(pvals < 0.05)
```

Fold change eli ryhmien välisten ekspressioerojen selvitys. Esimerkkinä CD-NCD vertailu, kaikki muutkin tehtiin samalla tavalla

```
> genesCDNCD1 <- which(CD_NCD_FC > 1)
> rnamesCDNCD <- rownames(logWT)[genesCDNCD1]
> write.table(rnamesCDNCD, file = "C:/Users/Iida/Desktop/Opiskelu/Kandi/genesCDNCD1.txt", sep = "\t")
> genesCDNCD15 <- which(CD_NCD_FC > 1.5)
```

```
> rnamesCDNCD <- rownames(logWT)[genesCDNCD15]
> write.table(rnamesCDNCD, file = "C:/Users/Iida/Desktop/Opiskelu/Kandi/genesCDNCD15.txt", sep = "\t")
```

MHC-kompleksin ekspressioeron selvitys

```
> FC2_WTCD <- medianLogWT - medianLogCD
> which(rownames(logWT) == "CD74")
[1] 16721
> which(rownames(logWT) == "HLA-DRB5")
[1] 18153
> which(rownames(logWT) == "HLA-DMB")
[1] 18175
> which(rownames(logWT) == "HLA-DPB1")
[1] 18183
> which(rownames(logWT) == "HLA-DQA1")
[1] 18157
> which(rownames(logWT) == "HLA-DPA1")
[1] 18182
> which(rownames(logWT) == "HLA-DQB1")
[1] 18158
> MHC_rownums <- c(16721, 18153, 18175, 18183, 18157, 18182, 18158)
> MHC_values <- FC2_WTCD[MHC_rownums]
> MHC_values
[1] 2.504225 2.621340 1.940016 2.412654 1.921736 2.745716
[7] 1.850328
```