TARTU ÜLIKOOL

LOODUS- JA TÄPPISTEADUSTE VALDKOND

MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT

GENOOMIKA INSTITUUT

PSORIAATILISE ARTRIIDI ÜLEGENOOMNE ASSOTSIATSIOONIUURING

Magistritöö

30 EAP

Aleksandra Judin

Juhendajad: PhD Ene Reimann

McS Tanel Traks

McS Galadriel Lucia Velazquez Silva

TARTU 2025

INFOLEHT

Psoriaatilise artriidi ülegenoomne assotsiatsiooniuuring

Kokkuvõte. Psoriaatiline artriit (PsA) on krooniline põletikuline liigesehaigus. Siinse magistritöö eesmärk oli tuvastada PsA geneetilisi riskitegureid ning uurida haigusega seotud bioloogilisi mehhanisme suuremahulise ülegenoomse assotsiooniuuringu (GWAS) metaanalüüsi kaudu. Analüüs viidi läbi kolmes Põhja-Euroopa kohordis: Eesti geenivaramu, Soome FinnGen ja UK Biopanga andmete põhjal, hõlmates kokku ligikaudu 900 000 osalejat. GWAS-i metaanalüüsi tulemused annoteeriti FUMA platvormi abil, kandidaatgeenide prioritiseerimine tehti *Open Targets Genetics* (OTG) platvormil ning seda toetati täiendavalt FUMA ja *Mouse Genome Informatics* (MGI) andmebaasi funktsionaalsete annotatsioonidega. Leitud geneetilised signaalid annavad uusi vihjeid PsA molekulaarsete mehhanismide kohta.

Märksõnad: psoriaatiline artriit, GWAS, metaanalüüs, geneetika, kandidaatgeenid

CERCS kood: B110 Bioinformaatika, meditsiininformaatika, biomatemaatika, biomeetrika;

B790 Kliiniline geneetika

Genome-wide association study of psoriatic arthritis

Abstract. Psoriatic arthritis (PsA) is a chronic inflammatory joint disease. This master's thesis aimed to identify genetic risk factors and explore biological mechanisms underlying PsA through a large-scale genome-wide association study (GWAS) meta-analysis. The analysis included data from three Northern European cohorts: the Estonian Biobank, FinnGen and UK Biobank, totalling nearly half a million participants. The GWAS meta-analysis results were annotated using the FUMA platform, candidate genes prioritisation was carried out using *Open Targets Genetics* (OTG) platform and further supported by functional annotations derived from FUMA and the *Mouse Genome Informatics* (MGI) database. The identified genetic signals and potential causal genes provide new insights into the molecular underpinnings of PsA.

Keywords: psoriatic arthritis, GWAS, meta-analysis, genetics, candidate genes

CERCS code: B110 Bioinformatics, medical informatics, biomathematics, biometrics.

B790 Clinical genetics

SISUKORD

INFOLEHT 2

SISUKORD 3

KASUTATUD LÜHENDID 5

SISSEJUHATUS 7

1 KIRJANDUSE ÜLEVAADE 9

1.1 Psoriaatilise artriidi klassifikatsioon ja definitsioon 9

1.2 Psoriaatilise artriiti avaldumine 10

1.3 Psoriaatilise artriidi patogenees 11

1.4 Nahabarjäär ja lipiidid 11

1.5 Liigesekahjustus psoriaatilise artriidi korral 12

1.6 Psoriaatilise artriidi ravimeetodid 13

1.7 Psoriaatilise artriidi geneetika 14

1.8 *HLA* geenid ja PsA risk 14

1.9 Ülegenoomsed assotsiatsiooniuuringud 16

1.10 Psoriaatilise artriidi senised ülegenoomsed assotsiatsiooniuuringud 18

2 EKSPERIMENTAALOSA 19

2.1 Töö eesmärgid 19

2.2 Materjal ja metoodika 20

2.2.1 Valim 20

2.2.2 Genotüpiseerimine, kvaliteedikontroll, imputatsioon 22

2.2.3 Assotsiatsioonianalüüs Eesti geenivaramu, FinnGen ja UK Biopanga andmetel 23

2.2.4 GWAS metaanalüüs Eesti geenivaramu, FinnGen ja UK Biopanga andmetel 25

2.2.5 GWAS metaanalüüsi tulemuste annotatsioon ja visualiseerimine FUMA platvormi kaudu 25

2.2.6 Psoriaatilise artriidiga seotud võimalike uute riskilookuste määratlemine 28

2.2.7 Kandidaatvariantide *in silico* annotatsioon FUMA platvormi abil. 29

2.2.8 Geenide prioritiseerimine *Open Targets Genetics* ja *Mouse Genome Informatics* platvormide abil 29

2.2.9 MAGMA koespetsiifiline geenieksрessiooni analüüs 32

2.3 Tulemused 33

2.3.1 Genoomsete riskilookuste iseloomustamine 33

2.3.2 Võimalike uute riskilookuste määratlemine 36

2.3.3 Kandidaatgeenide prioritiseerimine FUMA, *Open Target Genetics* ja *Mouse Genome Informatics* abil. 38

2.3.4 *Variant-to-gene* skoori tõendusmaterjal 38

2.3.5 Tõendusmaterjal geeni paiknemise kohta juht-SNP läheduses 38

2.3.6 Tõendusmaterjal eQTL-seoste kohta 39

2.3.7 Tõendusmaterjal kromatiininteraktsioonide kohta 39

2.3.8 *Missense*-variantide tõendusmaterjal 42

2.3.9 Fenotüübiline tõendusmaterjal hiire geeni väljalülitamise (*knockout*) mudelitest 42

2.3.10 Prioritiseerimise tõendusmaterjali kokkuvõte oluliste lookuse kohta 43

2.3.11 MAGMA koespetsiifiline geenieksрessiooni analüüs 45

2.4 Arutelu 48

KOKKUVÕTE 52

SUMMARY 53

TÄNUSÕNAD 56

KASUTATUD KIRJANDUS 57

KASUTATUD VEEBIAADRESSID 62

LISAD 63

LISA 1. 63

LISA 2 69

LISA 3. 70

LIHTLITSENTS 73

KASUTATUD LÜHENDID

AMP – adenosiinmonofosfaad

CI – usaldusintervall (ingl *confidence interval*)

EAF – efektialleeli sagedus (ingl *effect allele frequency*)

eQTL – ekspressiooni taset mõjutav lookus (ingl *expression quantitative trait loci*)

FUMA – funktsionaalne kaardistamine ja annoteerimine (ingl *Functional Mapping and Annotation*)

GWAMA – ülegenoomsete assotsiatsiooniuuringute metaanalüüs (ingl *Genome-Wide Association Meta Analysis*)

GWAS – ülegenoomne assotsiatsiooniuuring (ingl *Genome-Wide Association Study*)

h2 – päritavuskoefitsient (ingl *heritability*)

Hi-C – kõrge läbilaskevõimega kromosoomide konformatsiooni jäädvustamine (ingl *high-throughput chromosome conformation capture*)

HLA – inimese leukotsüütide antigeen (ingl *human leukocyte antigen*)

IFN-γ – interferoon-gamma

IL-17 – interleukiin-17

IL-22 – interleukiin-22

IL-23 – interleukiin-23

kb – tuhat aluspaari (ingl *kilobase pair*)

LD – ahelduse tasakaalutus (ingl *linkage disequilibrium*)

MAF – harvema alleeli sagedus (ingl *minor allele frequency*)

MAGMA – geenipõhise analüüsi meetod (ingl *Multi-marker Analysis of Genomic Annotation*)

Mb – miljon aluspaari (ingl *megabase pair*)

MGI – hiire genoomi informaatika andmebaas (ingl *Mouse Genome Informatics*)

NF-kB – nukleaarfaktor kappa-B

OTG – *Open Targets* geneetikaandmebaas (ingl *Open Targets Genetics*)

pQTL – valgu koguse kvantitatiivne tunnuse lookus (ingl *protein quantitative trait locus*)

Ps – Psoriaas

PsA – Psoriaatiline artriit

RHK-10 – Rahvusvaheline haiguste klassifikatsioon (ICD-10)

SD – standardhälve (ingl *standard deviation*)

SE – standardviga (ingl *standard error*)

SNP – üksiknukleotiidne polümorfism (ingl *single nucleotide polymorphism*)

SNP2GENE – SNP-geeni kaardistamine (ingl *single nucleotide polymorphism to gene*)

TNF-α – kasvajanekroosifaktor-alfa

TÜ EGV – Tartu Ülikooli Eesti geenivaramu (EGV)

UKBB – Ühendkuningriigi biopank (ingl *UK Biobank*)

UV – ultraviolettvalgus

V2G – variandi seostamine geeniga (ingl *variant-to-gene*)

SISSEJUHATUS

Psoriaatiline artriit (PsA) on krooniline põletikuline liigesehaigus, mis kuulub spondüloartriitide rühma ja on tihedalt seotud psoriaasiga (Ps-ga). Haigus avaldub liigeste, kõõluste ja entesiitide põletikuna ning võib viia tõsiste deformatsioonide ja funktsionaalse puudeni. PsA etioloogia on keeruline ning hõlmab nii geneetilisi kui ka keskkonnategureid. Varasemad uuringud on näidanud, et ligikaudu 40% PsA patsientide perekonnas on esinenud Ps-i või teisi autoimmuunhaigusi, mis viitab tugevale geneetilisele eelsoodumusele (Rahman ja Elder, 2005). Hoolimata teaduse arengust mõistetakse PsA täpset geneetilist mehhanismi siiski alles osaliselt ning see vajab edasist uurimist.

Viimastel aastatel on genoomi ulatuslikud assotsiatsiooniuuringud (GWAS – *genome-wide association study*) võimaldanud tuvastada mitmeid geneetilisi variante, mis on seotud PsA suurenenud riskiga. Selliste uuringute käigus otsitakse kogu genoomis üksiknukleotiidsete polümorfismide (SNP-de) ja haigusfenotüüpide vahelisi statistilisi seoseid. Tänu biopankade laiendamisele ja genotüpiseerimise tehnoloogiate arengule on olnud võimalik viia läbi suuremahulisi analüüse, mis on oluliselt suurendanud meie teadmisi haiguse geneetilisest arhitektuurist. Lisaks tavalistele GWAS uuringutele on üha enam kasutatud ka GWAS-i metaanalüüse, kus mitme kohordi andmed kombineeritakse, et suurendada leidude usaldusväärsust ja statistilist jõudu (Uffelmann *et al*., 2021). Metaanalüüsid on eriti olulised heterogeensete ja komplekshaiguste, nagu PsA, geneetilise tausta uurimisel.

Magistritöö eesmärk on analüüsida PsA geneetilisi riskifaktoreid Põhja-Euroopa populatsioonis, kasutades kolme suure biopanga – Tartu Ülikooli Eesti geenivaramu (EGV), Soome FinnGeni ja Ühendkuningriigi Biopanga (UKBB) – andmeid. Töös tehakse PsA ülegenoomse assotsiatsiooni metaanalüüs, et tuvastada uusi geneetilisi seoseid, mis võiksid anda täiendavat teavet haiguse patogeneesi kohta. Lisaks GWAS-metaanalüüsile tehakse tulemuste funktsionaalne annoteerimine FUMA veebiplatvormi abil, mis võimaldab hinnata avastatud SNP-de bioloogilist tähendust ja võimalikke mõjusid. Kandidaatgeenide prioritiseerimiseks kasutatakse *Open Targets Genetics* (OTG) platvormi koos FUMA ja MGI funktsionaalsete annotatsioonidega, integreerides geneetilisi ja funktsionaalseid andmekihte, sealhulgas *variant-to-gene* (V2G) skoore. Lõpuks hinnatakse prioritiseeritud geenide fenotüüpilist mõju hiirmudelite põhjal *Mouse Genome Informatics* (MGI) andmebaasis.

Siinse töö laiem eesmärk on aidata paremini mõista PsA geneetilist tausta ning toetada tulevasi uuringuid, mis keskenduvad haiguse bioloogiliste mehhanismide ja uute ravivõimaluste

avastamisele. Töös saadud tulemused võivad pakkuda uusi sihtmärke funktsionaalseks analüüsiks ja raviarenduseks ning aitavad täpsustada teadmisi PsA patogeneesi kohta. Lisaks võivad need andmed täiendada teadmisi polügeensete riskiskooride (PRS) rakendamise potentsiaalist PsA varajaseks diagnoosimiseks ja riskihindamiseks. Magistritöö on valminud Tartu Ülikooli molekulaar- ja rakubioloogia instituudi molekulaarsete bioteaduste õppekava raames ning eksperimentaalne osa on korraldatud Tartu Ülikooli genoomika instituudis.

1 KIRJANDUSE ÜLEVAADE

1.1 Psoriaatilise artriidi klassifikatsioon ja definitsioon

Psoriaatiline artriit (PsA) ja psoriaas (Ps) on tihedalt seotud autoimmuunhaigused, mille puhul ründab immuunsüsteem ekslikult keha kudesid. Ps on nahahaigus, mille puhul tekivad nahale punased ketendavad laigud, samuti võib see mõjutada küüsi ja muid kudesid. PsA kuulub spondüloartriidi gruppi. Need on teatud tüüpi põletikulised liigeshaigused, millele on iseloomulikud liigesepõletikud, entesiit (kõõluste ja sidemete kinnitumispaikade põletik) ja sakroiliit (põletik vaagnaluude liigestes) (Odgie ja Weiss, 2015). PsA võib põhjustada liigeste kahjustusi ja deformatsioone, samuti võib see piirata liigeste funktsioone ja liikumisvõimekust. Psoriaatiline artriit on seotud geneetiliste ja keskkonnateguritega, mis mängivad olulist rolli haiguse kujunemises (Anandarajah ja Ritchlin, 2009).

PsA on tavaliselt/sageli seotud Ps-ga, kuigi see võib esineda ka ilma nahal ilmnevate Ps-i sümptomiteta. Haiguse puhul esineb nii asümmeetrilisi ja sümmeetrilisi liigesepõletikke. Samuti on PsA-l mitmesuguseid alatüüpe, mis võivad näiteks sakroiliidi ja spondüliidi tekkega ulatuda perifeersetelt liigestelt kuni aksiaalsete liigesteni (Veale ja Fearon, 2018). PsA võib olla progresseeruv, mis tähendab, et haigus võib järk-järgult kahjustada liigesestruktuure ja põhjustada pöördumatuid deformatsioone. Geneetilised tegurid, nagu *HLA-B27* (*human leukocyte antigen)* geen, on sageli seotud PsA arengu ja progresseerumisega (Odgie ja Weiss, 2015).

PsA on määratletud RHK-10 koodiga M07, mis viitab põletikulistele Ps-ga seotud liigeshaigustele. Ps on määratud RHK-10 koodiga L40[1]. Need koodid on olulised haiguste klassifitseerimiseks meditsiinilistes süsteemides ja aitavad tagada õige ravi määramise. Samuti on need koodid kasulikud teadusuuringute ja epidemioloogiliste uuringute jaoks, et jälgida ja analüüsida Ps-i ja PsA levimust ning ravi tulemusi (Odgie ja Weiss, 2015). Erinevad PsA ja Ps vormid võivad vajada spetsiifilisi raviskeeme, mis põhinevad haiguse kliinilisel pildil ja raskusastmel (Odgie *et al*., 2020).

[1] https://icd.who.int/browse10/2019/en

1.2 Psoriaatilise artriiti avaldumine

PsA on krooniline süsteemne autoimmuunhaigus, mis esineb ligikaudu 20–30% Ps-ga patsientidest (Ogdie ja Weiss, 2015). Haigus võib kujuneda välja mistahes vanuses, kuid esineb sagedamini 30–60 aastastel inimestel. Lisaks Ps-le mängivad haiguse kujunemisel rolli ka geneetilised, immuunsüsteemiga seotud ja keskkondlikud tegurid, sealhulgas stress, infektsioonid ja traumad (FizGerald *et al*., 2021; Scott *et al*., 2022).

PsA kliiniline avaldumine on väga mitmekesine ja sümptomid võivad varieeruda kergest valu ja ebamugavustundest kuni raskete põletikuliste sümptomiteni, mõjutades erinevaid kehapiirkondi. Kuna PsA sümptomid on sageli sarnased teiste põletikuliste liigeshaigustega, nagu reumatoidartriit ja anküloseeriv spondüliit, on selle haiguse eristamine ja diagnoosimine keeruline (Ogdie ja Weiss, 2015). PsA-l on mitu kliinilist alatüüpi, millel on erinevad sümptomaatilised ja anatoomilised tunnused. Haiguse alatüübid on järgmised:

- **Asümmeetriline oligoartikulaarne artriit** − see PsA vorm on tavaliselt kergem ja mõjutab korraga vähem kui viit liigest. Asümmeetriline oligoartikulaarne artriit esineb sageli sõrme-, varba- ja põlveliigestes, põhjustades paistetust ja valulikkust. Seda tüüpi PsA diagnoositakse enamasti juba varakult ning see põhjustab harvem tõsiseid liigeste deformatsioone (Poddubnyy ja Rudwaleit, 2012).
- **Sümmeetriline polüartikulaarne artriit** − sarnaneb reumatoidartriidiga ja mõjutab mõlema kehapoole samu liigeseid, mistõttu on sümptomid tavaliselt sümmeetrilised. See PsA vorm võib põhjustada randmete, põlvede ja hüppeliigeste põletikku ning on sageli seotud tugevamate põletikuliste sümptomitega, mis võivad viia liigeste struktuuriliste muutuste ja funktsionaalsete piiranguteni (Cush, 2022).
- **Distaalne interfalangeaalne artriit** − mõjutab peamiselt sõrmede ja varvaste väikeseid distaalseid liigeseid (interfalangeaalliigeseid) ning seda iseloomustab valulikkus, turse ja jäikus. See PsA vorm esineb sagedasti koos Ps-i küünemuutustega, sealhulgas küünepinna auklikkuse, värvimuutuste ja paksenemisega (McGonagle *et al*., 2024).
- **Mutileeriv artriit** − kõige raskem ja harvemini esinev PsA vorm, mille puhul hävivad liigesed pöördumatult, põhjustades luude ja liigeste deformatsioone ning funktsiooni kaotust. See vorm on kõige sagedamini seotud raske Ps-ga ning võib viia raskete funktsionaalsete piiranguteni (González-Moreno *et al*., 2021).
- **Spondüliit** − mõjutab lülisammast ja sakroiliakaalseid liigeseid (vaagna alumises osas), põhjustades seljavalu, jäikust ja liikuvuse vähenemist. Spondüliidiga patsiendil on

sageli raskusi hommikuse jäikuse leevendamisega ja pikaajalise tegevusetuse järel võib valu süveneda (Zeidler, 2022).

1.3 Psoriaatilise artriidi patogenees

PsA patogenees on keeruline protsess, mis hõlmab geneetiliste ja keskkonnategurite koostoimet. Haiguse tekke ja progresseerumise taga on düsreguleeritud immuunsüsteem, mis käivitab põletikulise protsessi, põhjustades kudede kahjustusi ja kliinilisi sümptomeid. PsA korral on immuunsüsteemi düsregulatsioon keskne patogeneetiline mehhanism. Geneetiline eelsoodumus, nagu inimese leukotsüütide antigeeni *HLA-B27* ja *HLA-Cw6* alleelid, suurendab riski põletikulise protsessi tekkeks, mille käivitavad sageli keskkonnategurid, nagu trauma või infektsioonid (Rahman ja Elder, 2005). Põletikku vahendavad mitmed rakud, sealhulgas T-abistaja 17 (Th17 *helper*) rakud, mis toodavad põletikulisi tsütokiine interleukiin-17 (*IL-17*), interleukiin-23 (*IL-23*) ja tuumorinekroosifaktor-alfa (*TNF-α*) (Gaffen *et al*., 2014). Need tsütokiinid mängivad võtmerolli nii naha kui ka liigesehaiguste kujunemisel. Lisaks aktiveerivad dendriitrakud ja neutrofiilid põletikuvastust, põhjustades kudede kahjustusi ja haiguse progresseerumist (Ritchlin *et al*., 2017).

PsA patogeneesi üks iseloomulik tunnus on entesiit − kõõluste ja sidemete kinnituskohtade põletik. Entesiidi piirkonnas võib tekkida luu erosioon ning uue luukoe moodustamine, mis viib liigeste deformatsioonideni. Lisaks võib nahas vallandunud põletik levida liigestesse, põhjustades süsteemset põletikulist vastust. Haiguse keeruline patogenees hõlmab ka soolestiku mikrobioomi rolli. Mikrobioomi tasakaalu häired võivad mõjutada immuunsüsteemi regulatsiooni ja soodustada põletikuliste reaktsioonide teket, mis võivad omakorda kaasa aidata haiguse arengule. Need tegurid loovad tugeva seose naha ja liigeste vahel, mis on PsA keskne tunnus (Barnas ja Ritchlin, 2015). Lisaks liigeste ja naha haaratusele võivad PsA puhul mängida rolli ka immuunfunktsiooniga seotud organid, nagu põrn ja lümfisõlmed. PsA-d iseloomustab süsteemne põletik, mis võib kaasa tuua põrna ja lümfisõlmede reaktiivse suurenemise ning muutused nende struktuurses ja rakulises koostises. Põrn osaleb monotsüütide ja makrofaagide aktivatsioonis, lümfisõlmedes aktiveeritakse T- ja B-rakud ning kujuneb autoimmuunreaktsioon, viidates haiguse süsteemsele olemasolule (Veale ja Fearon, 2018; Rendon ja Schäkel, 2019).

1.4 Nahabarjäär ja lipiidid

Kuna PsA-l on sageli seos Ps-ga, on naha terviklikkus ja lipiidide ainevahetus haiguse patofüsioloogias olulised tegurid. Kuigi PsA võib areneda ka enne naha Ps-i ilmingute tekkimist

või isegi nende puudumisel, on nahabarjääri kahjustused ja lipiidide ainevahetuse häired sageli seotud süsteemsete põletikuliste protsessidega, mis mõjutavad nii naha funktsionaalsust kui ka liigeseid (Nestle *et al*., 2009). Nahabarjäär kaitseb keha väliste kahjustajate, näiteks patogeenide, UV-kiirguse ja toksiinide eest. Ps-i ja PsA korral, eriti kui nahalöövet ei esine, võib nahabarjääri düsfunktsioon täita patogeneesis erinevat rolli. PsA puhul on nahabarjääri kahjustus vähem määrav, kuid nahapõletiku ja süsteemse põletiku vahel on siiski teatud seos (Ritchlin *et al*., 2017; Veale ja Fearon, 2018).

Lipiidid on nahabarjääri kriitiline osa, kuna need aitavad säilitada naha niiskust, tugevdada barjääri ja kaitsta mikroobide eest. Tavalised naha lipiidid, nagu keramiidid, kolesterool ja vabad rasvahapped, moodustavad sarvkihi vahekihis struktuuri, mis on hädavajalik nahafunktsioonide jaoks. PsA korral mõjutavad lipiidide metabolismi häired ka immuunsüsteemi tööd, süvendades põletikulisi protsesse. Üks oluline mehhanism on seotud eikosanoididega**,** mis pärinevad arahhidoonhappest. Neil lipiididel, nagu prostaglandiinid ja leukotrieenid, on oluline osa põletikuste reaktsioonide vahendamisel, suunates immuunrakkude aktiivsust ja vastust (Dennis ja Norris, 2015; Simard *et al*., 2022).

Teine tähtis rühm on lipidomediaatorid, mis on lipiidide ainevahetuse saadused ning reguleerivad põletikulisi protsesse ja immuunvastust. PsA korral on nende tasakaal sageli häiritud, mis võib soodustada kroonilise põletiku süvenemist ja haiguse progresseerumist. Lipomediaatorite dünaamika on otseselt seotud PsA sümptomite tekke ja raskusastmega, muutes need võimalikuks sihtmärgiks ravistrateegiates (Simard *et al*., 2022; Barnas ja Ritchlin, 2015).

1.5 Liigesekahjustus psoriaatilise artriidi korral

PsA korral on liigesekahjustus üks haiguse tõsisemaid tagajärgi, mis võib viia püsiva puude ja elukvaliteedi olulise halvenemiseni. Haiguse varases faasis algab põletik sünoviaalkoes − liigeskapslit vooderdavas koes, mis muutub hüperplastiliseks ja hakkab tootma põletikumediaatoreid, sealhulgas *TNF-α*, *IL-17* ja *IL-23*. Need põletikulised tsütokiinid käivitavad immuunrakkude aktivatsiooni, mis omakorda stimuleerib osteoklastide aktiivsust, põhjustades luude erosiooni ja kõhrkoe lagunemist. Kroonilise põletiku tagajärjel areneb näiteks „pannus“ – invasiivne sünoviaalne kude, mis kahjustab liigesepindu, viies struktuursete deformatsioonide ja liikumisulatuse piiranguteni. Varajane ja sihipärane sekkumine on võtmetähtsusega, et vältida pöördumatut liigesekahjustust ja säilitada funktsionaalsust (Barnas ja Ritchlin, 2015).

PsA-i puhul erineb liigesekahjustus mõnevõrra reumatoidartriidist, kuna lisaks luukaole võib esineda ka uusluu moodustumist, mis viib luuliste sildade ja jäigastumiste tekkeni. Seda protsessi nimetatakse periotaalseks luukasvuks ja see võib ilmneda kõrvuti erosioonide tekkega, tekitades unikaalse radioloogilise mustri. Uue luu teke on eriti iseloomulik perifeersetes liigestes, nagu sõrmedes ja varvastes, ning võib põhjustada liigeste jäikust isegi juhul, kui põletikuline aktiivsus on vaibunud. Luude ebanormaalne remodelleerumine muudab haiguse kulu progresseeruvaks ja raskesti ravitavaks, mistõttu on oluline jälgida haiguse arengut mitte ainult kliiniliste sümptomite, vaid ka radioloogiliste leidude kaudu. Selline kahekordne protsess – erosioon ja luuteke – muudab PsA diagnostiliselt ja terapeutiliselt keerukaks (Ritchlin *et al*., 2017).

Liigesekahjustuse raskusaste on sageli seotud haiguse alatüübiga. Näiteks daktüliit ehk „vorstisõrm“ on seisund, kus kogu sõrm või varvas tursub põletiku tõttu, kuna põletik haarab nii liigese, kõõlusekesta kui ka entesiidid ehk kõõluste kinnituspiirkonnad. Daktüliit on tugevalt seotud raskema haiguskuluga ning erosioonide tekke suurema tõenäosusega. Samuti on olemas haruldane, kuid eriti raske haigusvorm – mutileeriv artriit, mille puhul toimub ulatuslik luude hävimine, liigeste deformatsioon ja funktsiooni kaotus. Selle alatüübi korral võivad sõrmed või varbad luude sulamise ja resorptsiooni tõttu lühemaks muutuda, tuues kaasa ulatusliku töövõime kaotuse ja sõltuvuse välistest abivahenditest (González-Moreno *et al*., 2021; Kaeley *et al*., 2018).

Lisaks perifeersetele liigestele võivad PsA korral kahjustuda ka selgroolülid ja vaagnapiirkonna liigesed, eriti sakroiliakaalliigesed. Seda nimetatakse aksiaalseks PsA vormiks, mis võib sarnaneda anküloseeriva spondüliidiga, kuid erineb selle poolest, et selja jäikus võib olla asümmeetriline ning koosneda samaaegselt perifeersest liigesepõletikust. Aksiaalse PsA korral võivad selgroolülide vahel tekkida põletikulised kahjustused ja uued luusillad, mis lõpuks viivad selgroo jäigastumiseni. Sellised muutused on aeglase kuluga, kuid võivad ilma sekkumiseta patsiendi liikuvust ja igapäevast toimetulekut oluliselt piirata. Radioloogilised uuringud, nagu magnetresonantstomograafia ja röntgen, on olulised selle haigusevormi tuvastamiseks ja jälgimiseks (Tiwari ja Brent, 2024).

1.6 Psoriaatilise artriidi ravimeetodid

PsA ravi eesmärk on vähendada põletikku, parandada funktsionaalset võimekust ja ennetada tõsisemaid liigesekahjustusi. Ravi valik sõltub haiguse raskusastmest, sümptomitest ja patsiendi individuaalsest seisundist. Ravivõimalused hõlmavad farmakoloogilisi ja mittefarmakoloogilisi meetodeid. PsA valdkonnas liiguvad teadus- ja ravisuunad üha enam süsteemse ja integreeritud

käsitluse poole, kus eesmärgiks ei ole ainult sümptomite ohjeldamine, vaid ka haiguse aluseks olevate mehhanismide mõjutamine. Uued suunad hõlmavad täppismeditsiini arengut, mille abil saab analüüsida patsiendi geneetilist ja molekulaarset profiili, et prognoosida haiguse kulgu ja määrata individuaalne raviplaan. Lisaks bioloogilistele ja suunatud ravimitele, mis on juba kliinilises kasutuses, arendatakse ka uusi diagnostilisi platvorme, sealhulgas biomarkerite põhiseid teste, mis võimaldavad haiguse tuvastada varem ja täpsemalt. Oluline suundumus on ka PsA alamtüüpide eristamine molekulaarsel tasandil, mis aitaks vähendada ülemäärast ravi ja valida tõhusamaid ravikombinatsioone. Need edusammud põhinevad üha täpsematel andmetel haiguse immuunpatogeneesi kohta (Ogdie *et al*., 2020; FitGerald *et al*., 2021).

1.7 Psoriaatilise artriidi geneetika

PsA geneetiline taust on intensiivse uurimistöö keskmes, kuna tegemist on kroonilise põletikulise haigusega, mis esineb sageli geneetilise eelsoodumusega isikutel. Ligikaudu 40% PsA patsientidest on perekonnas esinenud Ps-i või sarnaseid autoimmuunhaigusi, mis viitab tugevale pärilikkusele (Rahman ja Elder, 2005). Geneetilised tegurid mõjutavad haiguse arengut mitmel tasandil, sealhulgas immuunvastuse reguleerimist, põletikuliste reaktsioonide käivitamist ja kudede kahjustust. Samuti on geneetiline mitmekesisus oluline haiguse raskusastme määramisel ja ravivastuse varieerumises. Geneetilise mitmekesisuse uurimine aitab mõista haiguse heterogeenset iseloomu ja toetab täpsemate ravimeetodite väljatöötamist. Geneetilised variatsioonid võivad esineda nii valke kodeerivates geenipiirkondades (näiteks *HLA-B, IL23R, TNFAIP3*) kui ka mittekodeerivates ja regulatiivsetes piirkondades, mis mõjutavad geenide avaldumise taset, ajastust ja koe-spetsiifilisust (Bowes *et al*., 2011; Stuart *et al*., 2015).

1.8 *HLA* geenid ja PsA risk

Peamised geneetilised riskifaktorid PsA puhul on seotud inimese leukotsüütide antigeeni (*HLA*) süsteemiga. *HLA* geenid *HLA-B* ja *HLA-C* kodeerivad klass I molekule, mis osalevad antigeeni esitamisel immuunrakkudele. *HLA-B* on tugevalt polümorfne geen, mis mängib võtmerolli immuunvastuses viiruste ja bakterite vastu ning autoimmuunhaiguste, sealhulgas Ps ja PsA kujunemises. *HLA-C* aktiveerib looduslikke tappurrakke (NK-rakke) ja on seeläbi seotud immuunregulatsiooniga, mõjutades infektsioonide ja autoimmuunsete protsesside kulgu (Chandran *et al*., 2013). PsA-ga seotud olulised *HLA* geenid on *HLA-B27*, mille esinemine on seotud entesiidi ja aksiaalse skeleti kahjustustega, ning *HLA-Cw6*, mis on seotud raskema nahapõletiku ja haiguse varasema algusega (Zalesak *et al*., 2024).

Lisaks *HLA* geenidele on PsA-ga seostatavaks peetud mitmeid mitte-*HLA* geenivariante, mis mõjutavad põletikulisi ja autoimmuunseid mehhanisme. Olulist rolli täidavad näiteks *IL12B*

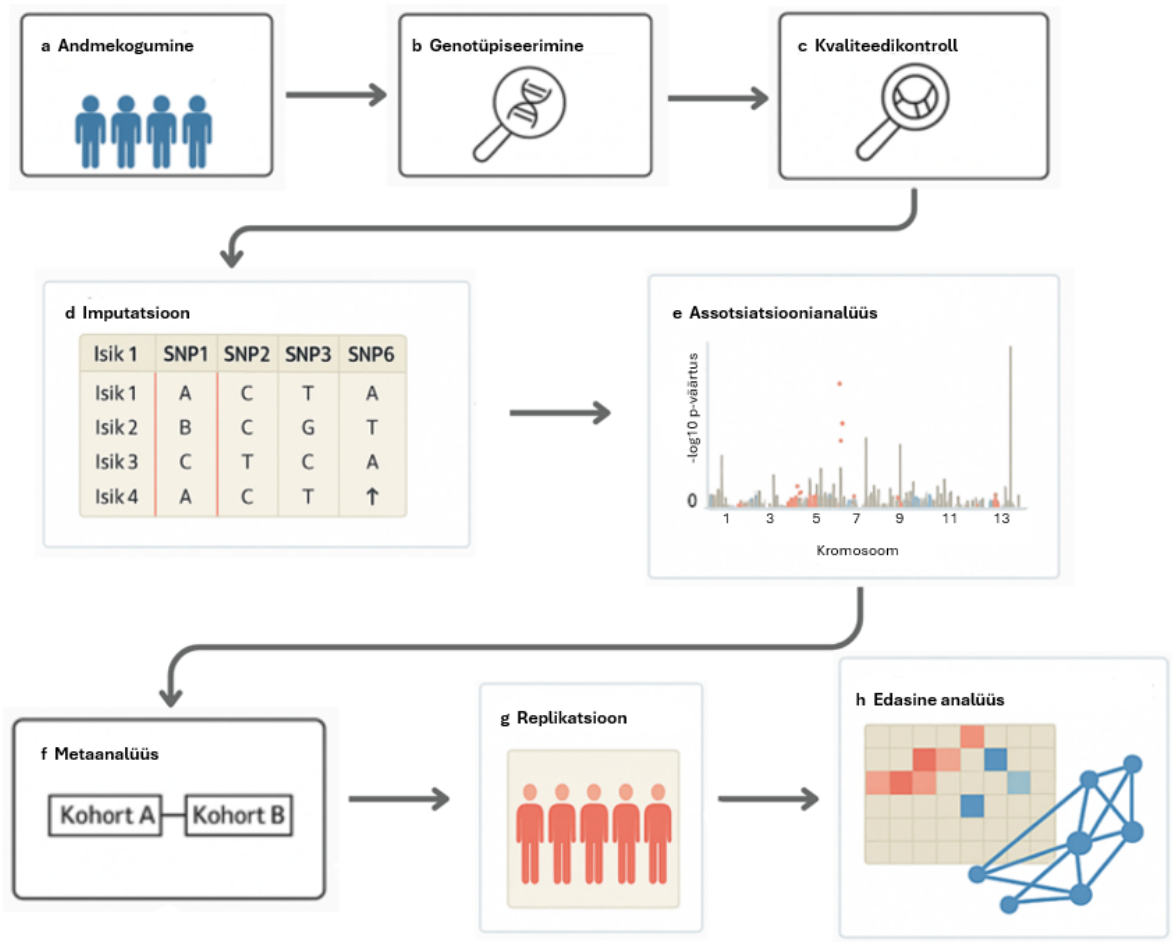
ehk interleukiin 12 subühikut p40 kodeeriv geen ja *IL23R* ehk interleukiin 23 retseptorit kodeeriv geen. Need geenid reguleerivad interleukiin-12 ja interleukiin-23 signaaliülekannet, mis on oluline *Th17*-rakkude aktivatsioonis. *Th17*-rakud toodavad põletikulisi tsütokiine, sealhulgas *IL-17* ja *IL-22*, millel on keskne roll PsA patogeneesis (Ellinghaus *et al*., 2010). Teiseks oluliseks geeniks on *TNFAIP3* ehk *Tumor Necrosis Factor Alpha-Induced Protein 3,* mis kodeerib valku A20. See valk on oluline nukleaarfaktor kappa-B (NF-kB) signaaliradade negatiivseks reguleerimiseks. Kui selles geenis esinevad mutatsioonid, võib see viia põletiku üleaktiveerimiseni, suurendades sellega PsA riski (Ciccacci *et al*., 2019). Samuti on märkimisväärne *CARD14* ehk *Caspase Recruitment Domain Family Member 14,* mis on seotud keratinotsüütide aktiveerimises ja nahapõletiku regulatsioonis. Selle geeni mutatsioonid võivad tugevdada naha põletikulisi reaktsioone, mis on omakorda seotud nii Ps-i kui ka PsA-ga (Jordan *et al*., 2012). Tähtsust on kogunud ka *SOCS1* (*Suppressor of Cytokine Signalling 1*), mis toimib negatiivse regulaatorina tsütokiinidest vahendatud signaaliradades. *SOCS1* piirab põletikuliste tsütokiinide toimet, takistades liigset immuunvastust. Selle geeni funktsioonihäired võivad soodustada kroonilise põletiku ja autoimmuunsete seisundite teket (Yoshimura *et al*., 2007). Samuti on hiljutised uuringud esile toonud ka *TRIM27* (*Tripartite Motif Containing 27*) geeni rolli T-rakkude aktivatsiooni reguleerimisel. *TRIM27* kuulub *TRIM*-valkude perekonda ja osaleb valkude ubikvitineerimisel ning immuunvastuse moduleerimisel. Selle seost autoimmuunhäiretega on kirjeldatud varasemates uuringutes (Hatakeyma, 2017). Lisaks on esile tõstetud *PDE8A* (*Phosphodiesterase 8A*), mis reguleerib cAMP taset ja on oluline T-raku funktsioonide, eriti adhesiooni ja liikumise kontrollis (Epstein *et al*., 2021). Oluliseks geeniks peetakse ka *AKAP13* (*A-kinase anchoring protein 13*), mis toimib signaalide korraldajana cAMP-radades ja reguleerib ka tsütoskeleti signaale, mis on seotud põletikuprotsessidega (Toaldo *et al*., 2015). Ühtlasi on ühe olulise geenina esile kerkinud *CAMK2G* (*Calcium/Calmodulin-Dependent Protein Kinase Type II Gamma*), kuna see vahendab signaale kaltsiumi ja kalmoduliiniga seotud radades ning mängib võtmerolli T-rakkude aktivatsioonis ja põletikuprotsesside moduleerimises. *CAMK2G* kõrgendatud aktiivsus on seotud põletikuliste seisunditega ning selle ekspressioon on suurenenud Ps kahjustuses (Yu *et al*., 2024). Samuti on tähelepanu pälvinud *PLAU* (*Plasminogen Activator, Urokinase*), mis osaleb fibrinolüüsis ning kudede remodelleerimises, mõjutades ekstratsellulaarse maatriksi lagundamist ja immuunrakkude migratsiooni põletikukoldesse. Selle geeni ülesregulatsioon võib soodustada liigesepõletiku ja koe kahjustusi (Cook *et al*., 2010). Lisaks on välja toodud ka *LEMD2* (*LEM Domain Containing 2*), mis kodeerib sisemise tuumamembraani valku ning osaleb tuuma struktuuri ja kromatiini organisatsiooni säilitamises (Szél *et al.*, 2019).

1.9 Ülegenoomsed assotsiatsiooniuuringud

Ülegenoomsed assotsiatsiooniuuringud (GWAS, *genome-wide association studies*) on tänapäeval üks olulisimaid meetodeid, et tuvastada seoseid geneetiliste variantide ning kompleksse haiguste ja tunnuste vahel (vt joonis 1, lk 17). GWAS võimaldab tuvastada ühe nukleotiidi polümorfisme (SNP-d), mille alleelid esinevad erineva sagedusega juhtude ja kontrollide rühmades. Selliste uuringute laialdane kasutusevõtt on saanud võimalikuks tänu suurtele biopankadele, genotüpiseerimistehnoloogia arengule ning järjest paranevatele arvutuslikele meetodile. Tänaseks on GWAS andmebaasi (GWAS *Catalog*) koondatud juba üle 100 000 uuringu, mis hõlmavad tuhandeid erinevaid fenotüüpe ja sadu tuhandeid geneetilisi variante. Need uuringud aitavad mõista haiguste geneetilist arhitektuuri, pakkudes alust uute diagnostiliste markerite ja ravimisihtmärkide väljatöötamiseks (Uffelmann *et al*., 2021; Lappalainen *et al*., 2024).

GWAS-i läbiviimise põhimõte tugineb regressioonianalüüsile, kus iga SNP seostatakse uuritava fenotüübiga eraldi. Kui uuritav tunnus on binaarne (näiteks haiguse olemasolu või puudumine), kasutatakse logistilist regressiooni; kvantitatiivsete tunnuste korral (näiteks vere kolesteroolitase) rakendatakse lineaarset regressiooni. Oluline on arvestada sugulussidemete, populatsioonistruktuuri ja geneetilise kihistumisega, et vältida valepositiivseid seoseid. Samuti on oluline fenotüüpide täpne määratlemine ning sobivate kontrollgruppide valik, kuna ebatäpne määratlus võib viia valepositiivsete või -negatiivsete leidudeni (Mbatchou *et al*., 2021; Boyle *et al*., 2017).

GWAS-i edasiarendusena on välja kujunenud metaanalüüsid, mis ühendavad erinevate uuringute tulemusi, et suurendada statistilist võimsust ja usaldusväärsust. Metaanalüüs võimaldab koondada andmeid mitmest sõltumatust kohordist, sageli eri riikidest või populatsioonidest, suurendades seeläbi avastatavate SNP-de hulka. Metaanalüüsi tehakse tavaliselt koondstatistikute (*summary statistics)* põhjal, mis tähendab, et indiviidi tasemel andmeid pole vaja jagada. Tulemuste kombineerimiseks kasutatakse spetsiaalseid tarkvarasid, nagu METAL või GWAMA, mis arvestavad uuringutevahelist heterogeensust ja korrigeerivad võimalike kallutatusi, mis võivad tuleneda näiteks populatsioonistruktuuri erinevustest, genotüpiseerimisvigadest või muudest tehnilistest teguritest. Metaanalüüs on muutumas järjest olulisemaks, sest see võimaldab uurida ka harvaesinevaid variante ja hinnata leitud seoste reprodutseeritavust erinevas populatsioonides (Uffelmann *et al*., 2021).



Joonis 1. Skemaatiline ülevaade ülegenoomse assotsiooniuuringu töövoost (ülevaade Uffelman *et al*., 2021, kohandatud).

1.10 Psoriaatilise artriidi senised ülegenoomsed assotsiatsiooniuuringud

Seni tehtud PsA-i ülegenoomsed assotsiatsiooniuuringud on oluliselt laiendanud arusaama haiguse geneetilisest alusest. Uuringud on keskendunud peamiselt Euroopa päritolu populatsioonidele ja hõlmanud tuhandeid PsA-ga patsiente ja kontrolle. Esimestes GWAS-ides leiti tugev seos *HLA* piirkonna, eriti *HLA-B* ja *HLA-C* geenidega (Chandran *et al*., 2013; Bowes *et al*., 2011). Hilisemad suuremad uuringud on tuvastanud ka seoseid mitte-*HLA* geenidega, näiteks *IL12B*, *IL23R* ja *TNFAIP3*. Lisaks on analüüsid näidanud, et kuigi Ps-i ja PsA geneetiline taust kattub osaliselt, leidub ka PsA-le spetsiifilisi geneetilisi riskitegureid (Stuart *et al*., 2015).

Suuremad valimid ja metaanalüüsid on võimaldanud avastada rohkem PsA-ga seotud riskilookuseid. Uuemates uuringutes on esile kerkinud geenid nagu *IL23A, TRAF3IP2* ja *REL*, millel on oluline roll immuunvastuse vahendamisel (Zalesak *et al*., 2024; Stuart *et al*., 2015). On leitud, et mõned variandid, näiteks *TRAF3IP2* geenis, seostuvad spetsiifiliselt liigesepõletikuga, viidates alamtüüpide geneetilistele erinevustele. Samuti kattuvad mõned PsA riskilookused teiste autoimmuunhaigustega, nagu reumatoidartriit ja põletikuline soolehaigus (Alesa *et al*., 2019; Zalesak *et al*., 2024). Huvipakkuvad ülegenoomsed assotsiatsiooniuuringud on kokkuvõtvalt esitatud tabelis 1 (vt lisa 1).

2 EKSPERIMENTAALOSA

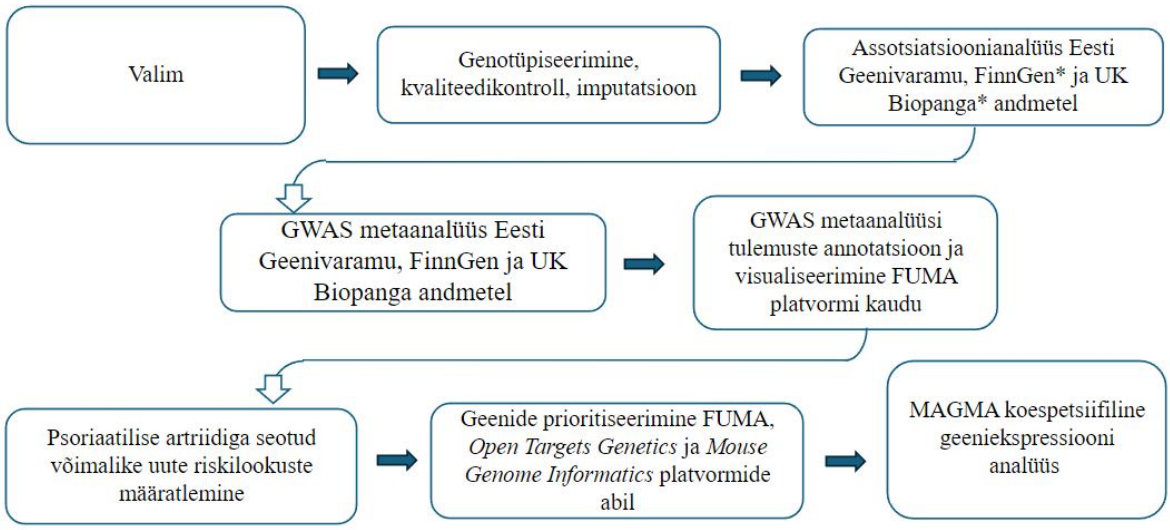
2.1 Töö eesmärgid

Uurimistöö eesmärk on analüüsida psoriaatilise artriidi geneetilisi riskifaktoreid Põhja-Euroopa populatsioonis (vt joonis 2, lk 20), et paremini mõista nimetatud haiguse patogeneesi. Selle saavutamiseks on töös järgmised alaeesmärgid:

1. Viia läbi GWAS metaanalüüs, kasutades Eesti, Soome ja UK Biopankade andmetel tehtud GWAS-ide koondstatistikuid.
2. Annoteerida GWAS metaanalüüsi tulemused FUMA veebiplatvormi abil.
3. Tuvastada võimalikke uusi PsA-ga seotud riskilookuseid, võrreldes saadud tulemusi juba GWAS-kataloogis avaldatud teadaolevate PsA riskilookustega.
4. Prioritiseerida potentsiaalseid kandidaatgeene, integreerides regulatoorseid, funktsionaalseid ja kliiniliste fenotüüpide andmeid, et määrata kindlaks kõige tõenäolisemalt haiguse mehhanismides osalevad geenid. Kasutada selleks erinevaid ressursse nagu *Open Targets Genetics* platvorm, FUMA veebiplatvorm ja *Mouse Genome Informatics* andmebaas.
5. Uurida kandidaatgeenide võimalikke fenotüübilisi mõjusid *Mouse Genome Informatics* ja *Open Targets Genetics* andmebaaside põhjal, et paremini mõista nende bioloogilist rolli.

2.2 Materjal ja metoodika

2.2.1 Valim



Joonis 2. Metoodika plaan. *FinnGen ja UKBB puhul on kasutatud saadaolevaid summaarstatistika andmeid (Eesti geenivaramu andmetel põhinev GWAS viidi käesoleva töö raames läbi koostöös Tartu Ülikooliga).

PsA geneetilise tausta uurimiseks viidi läbi ulatuslik genoomiülene assotsiatsiooniuuringu metaanalüüs, kaasates andmed kolmest Euroopa päritolu populatsioonipõhisest biopangast: Tartu Ülikooli Eesti geenivaramust (TÜ EGV), Soome FinnGen projektist (FinnGen) ja Ühendkuningriigi Biopangast (UKBB). Uuringu eesmärk oli koondada suurim võimalik andmemaht ning selgitada välja geneetilised variandid, mis võivad mõjutada PsA-i riski. Metaanalüüsi kaasatud kohortide ülevaatlik info on välja toodud tabelis 2 (lk 22).

TÜ EGV andmestiku (andmeversioon 2023v4) põhjal kaasati analüüsi 1625 PsA juhtumit ning 204 233 kontrollisikut. Juhtude seas moodustasid naised 67,3% (n = 1093) ja mehed 32,7% (n = 532); keskmine vanus oli 47,3 aastat (SD = 13,5). Kontrollide seas moodustasid naised 65,4% (n = 133671) ja mehed 34,6% (n = 70562); keskmine vanus oli 44,2 aastat (SD = 16,3). PsA-i juhtude määratlemiseks kasutati RHK-10 koode L40.5*, M07.0* ja M07.3*. Kontrollrühma määratlemisel välistati ainult juhud, kelle puhul esines mõni eelnimetatud diagnoosi koodidest. Kõik siinses töös kasutatud analüüsid viidi läbi Eesti bioeetika ja inimuuringute nõukogu eetikaloa nr. 1.1-12/624 alusel ning selleks kasutati andmeväljastust 6-7/GI/10617. Eesti geenivaramu hõlmab üle 200 000 osaleja, mis moodustab ligikaudu 20% Eesti täiskasvanud elanikkonnast. EGV üks oluline tugevus on regulaarne andmete uuendamine – geenidoonorite kohta kogutud terviseinfo täieneb igal aastal, hõlmates andmeid Eesti Tervise infosüsteemist, retseptikeskusest ja muudest elektroonilisest riiklikest andmekogudest. See võimaldab

haigusjuhtude täpset tuvastamist ning suurendab andmestiku usaldusväärsust pikaajalistes geneetilistes uuringutes (Leitsalu *et al*., 2015).

FinnGeni projektist kasutati kümnendat andmeväljalaset (*Data Freeze* 10), kus analüüsi kaasati 3537 PsA juhtu ning 262 844 kontrollisikut[2]. PsA juhud määratleti RHK-10 koodide põhjal (M07.0* kuni M07.3* koos või ilma L40.5 tähiseta), juhtudest eemaldati kõik M13 (*arthropathies)* diagnoosikoodiga indiviidid, et vältida fenotüübi kattumist teiste artriitide alatüüpidega. Kontrollgrupp koosnes isikutest, kellel puudusid kõik artriidi või psoriaasi diagnoosikoodid. FinnGeni andmestik on tuntud oma kõrge kvaliteedi, geneetilise homogeensuse ja rahvastikupõhise terviseinfosüsteemi tõttu, mis teeb sellest väärtusliku ressursi haigusassotsiatsioonide uurimisel (Kurki *et al*., 2023).

UK Biopanga andmetest kaasati 708 PsA juhtu ning 398 199 kontrolli. PsA määratleti UKBB spetsiifilise fenotüübikoodi alusel – kood 696.42 (*psoriatic arthropathy*). Kontrollideks olid isikud, kellel puudus see kood ning kellel ei esinenud Ps-i ega muid kroonilisi põletikulisi liigeshaigusi. See tähistab PsA-d, nagu see on registreeritud UK Biopanga kliinilistes andmetes, mis hõlmavad haiglaepisoodide ja perearstide andmeid. Kuigi kasutatud määratlused erinevad RHK-10 süsteemist, on see definitsioon kooskõlas varasemate UKBB GWAS uuringutega ja võimaldab sobivat metaanalüüsi tegemist. UKBB tugev külg on lai fenotüüpide ja elustiilitegurite andmestik, mis võimaldab tulemusi mitmekülgselt tõlgendada[3].

[2] http://r7.finngen.fi
[3] https://www.ukbiobank.ac.uk

Tabel 2. PsA GWAS metaanalüüsis kasutatud kohortide ülevaade. RHK-10 – Rahvusvaheline haiguste klassifikatsioon (ICD-10).

Kohort	Andme-versioon	Juhtude arv	Kontrollide arv	Juhtude definitsioon	Kontrollide definitsioon
EGV	2023v4	1625	204233	RHK-10: L40.5*, M07.0*, M07.3*	Välistati ainult PsA juhud
FinnGen R10	R10	3537	262844	RHK-10: M07.x ja L40.5, välja arvatud kõik M13 *arthropathiad*	Välistati kõik PsA ja Ps juhud
UK Biopank	-	708	398199	UKBB kood 696.42 (*psoriatic arthropathy*)	Välistati kõik PsA ja Ps juhud
Kokku		5870	865276		

2.2.2 Genotüpiseerimine, kvaliteedikontroll, imputatsioon

Magistritöös kasutati ülegenoomse assotsiatsioonianalüüsi koondstatistikuid, mille arvutamiseks kasutati genotüpiseerimise andmeid kolmest suurest Euroopa päritolu kohordist: EGV, FinnGen ja UKBB. Igas kohordis viidi genotüpiseerimine, kvaliteedikontroll ning SNP-de imputatsioon läbi vastavalt biopanga enda standardiseeritud töövoogudele.

Eesti geenivaramu puhul viidi kõik genotüpiseerimise ja kvaliteedikontrolli protseduurid läbi Tartu Ülikooli geenivaramu tuumiklaboris. Genotüpiseerimiseks kasutati Illumina GSAv1.0, GSAv2.0 ja GSAv2.0_EST kiipe[4]. Kvaliteedikontrolli käigus eemaldati andmestikust isikud, kellel esines mittevastavus genotüübilise ja fenotüübilise soo vahel, ning need, kelle *call rate* jäi alla 95%. Samuti eemaldati variandid, mille *call rate* oli < 95%, Hardy-Weinbergi tasakaalu

[4] https://www.illumina.com/

p-väärtus < 10^{-4} (autosoomide puhul) või minoorse alleeli sagedus (MAF) < 1%. Eelfaasimine toimus Eagle v2.3 tarkvara abil ja imputeerimiseks kasutati Beagle v.28Sep18.793 tarkvara koos Eesti-spetsiifilise referentspaneeliga, mis põhineb 2297 täisgenoomiga isiku andmetel (Browning ja Browning, 2007; Mitt *et al*., 2017; Pujol-Gualdo *et al*., 2022).

FinnGeni andmestikus viidi genotüpiseerimine läbi mitme Illumina ja Affymetrixi kiibi abil, kuid konkreetseid kiipide versioone pole dokumentatsioonis täpsustatud. Kvaliteedikontrolli ning imputatsiooni meetodid on kirjeldatud Kurki *et al*. (2023) teadusartiklis ning FinnGeni ametlikus dokumentatsioonis. Eelfaasimiseks kasutati Eagle v2.3.5 tarkvara ning SNP-de imputatsiooniks Beagle v.08Jun17.d8b programmi. Imputatsioon tugines *Sequencing Initiative Suomi* (SISu) v3 referentspaneelile, mis on kohandatud Soome populatsiooni geneetilistele eripäradele (Kurki *et al*.,2023)[5].

UK Biopanga andmestikus viidi genotüpiseerimine läbi kahel spetsiaalselt selle projekti jaoks välja töötatud kiibil: UK BiLEVE ja UK Biobank *Axiom arrays*. Eelfaasimine toimus Eagle2 tarkvara abil, samas kui SNP-de imputatsiooniks kasutati Minimac3 tarkvara. Imputatsiooniks kasutati kombineeritud referentspaneele: *Haplotype Reference Consortium* (HRC) ning UK10K + 1000 Genomes (Bycroft *et al*., 2018). Kogu protsess järgib UK Biopanga kvaliteedikontrolli standardeid, mis tagavad andmete kõrge usaldusväärsuse. Täpsem teave metoodikate ja andmestike kohta on kättesaadav UKBB ametlikus andmeportaalis[6].

2.2.3 Assotsiatsioonianalüüs Eesti geenivaramu, FinnGen ja UK Biopanga andmetel

Genoomiülene assotsiatsiooniuuring viidi läbi TÜ Eesti geenivaramu andmete põhjal Ene Reimanni poolt, kasutades tarkvarapaketti REGENIE v3.2., mis on loodud suurte ja keeruliste andmestike tõhusaks käsitlemiseks ja sobib hästi ka suguluses olevate isikute kaasamiseks mudelisse (Mbatchou *et al*., 2021). Kasutati logistilist regressioonimudelit koos aditiivse geneetilise mudeliga, mille eesmärgiks oli tuvastada sõltumatute SNP-de seos PsA-ga. Analüüsis võeti arvesse olulisi kovariaate: sugu, vanus ning kümme esimest geneetilist peakomponenti (PC), et kontrollida võimalikke populatsioonistruktuuride mõju tulemustele. Enne metaanalüüsi kaasamist filtreeriti andmestikust välja geneetilised markereid, mille

[5] http://www.sisuproject.fi/
[6] https://www.ukbiobank.ac.uk

imputatsiooni kvaliteedi INFO skoor oli alla 0,8, ning mille vähemuses oleva alleeli sagedus (MAF) oli alla 0.01. Pärast filtreerimist jäi EGV andmestikku alles 8 392 853 geneetilist markerit, mis kaasati edasisse GWAS metaanalüüsi. Kuna töö keskendus geneetiliste riskitegurite uurimisele laiemas Põhja-Euroopa kontekstis, siis ei käsitletud EGV andmetel tehtud GWASi tulemusi eraldi.

Soome biopanga FinnGen puhul oli võimalik kasutada kümnenda andmeväljalaske (R10) koondstatistikuid (fail *summary_stats_finngen_R10_M13_PSORIARTH* (https://r10.finngen.fi/pheno/M13_PSORIARTH)). Koondstatistikud pärinesid GWAS analüüsist, mis oli tehtud tarkvaraga REGENIE v2.0.2. Kasutatud kovariaadid olid sugu, vanus, genotüpiseerimise partii ja kümme geneetilist PC-d. Kuna FinnGeni statistikafailides puudus INFO skoori veerg, rakendati enne metaanalüüsi ainult MAF ≥ 0,01 põhine filtreerimine. Pärast andmete puhastamist jäi analüüsi kaasamiseks 8 423 453 SNP-d.

UKBB puhul kasutati samuti juba avaldatud koondstatistikuid (https://pheweb.org/UKB-SAIGE/pheno/696.42). Koondstatistikute arvutamiseks tehti assotsiatsioonianalüüs tarkvaraga SAIGE, mis on sobiv lahendus suuremahuliste ja ebavõrdselt jaotunud juht-kontroll andmestike analüüsimiseks. Genotüüpide eelfaasimine viidi läbi Eagle2 tarkvaraga ning imputatsiooniks kasutati Minimac3 tarkvara, viidates mitmele referentspaneelile: HRC, UK10K ja 1000 Genomes. Kuna UKBB koondstatistikutes puudus samuti INFO skoori veerg, rakendati ainult MAF ≥ 0,01 põhine filtreerimine. Lõpuks jäi UKBB andmestikust metaanalüüsiks sobivaks 7 647 617 geneetilist markerit. Kuna EGV ja FinnGeni GWAS-ide koondstatistikud olid joondatud GRCh38 referentsgenoomile ja UK Biopanga koondstatistikud GRCh37 referentsgenoomile, siis oli edasise metaanalüüsi jaoks vaja viimased konverteerida GRCh38-le. Selleks kasutati Reedik Mägi töörühmas välja töötatud koodi. Konversiooni aluseks oli dbSNP (*Single Nucleotide Polymorphism Database)* andmebaasi versioon 151.

2.2.4 GWAS metaanalüüs Eesti geenivaramu, FinnGen ja UK Biopanga andmetel

PsA GWAS-i metaanalüüsi tegi Galadriel Lucia Velazquez Silva (McS), kasutades tarkvara GWAMA v2.2 (Mägi ja Morris, 2010) ja rakendades genoomikontrolli parandust (-gc *option*), et korrigeerida võimalikke inflatsioonimõjusid, mis võivad tuleneda populatsioonistruktuurist või muude süsteemsete tegurite mõjust testi statistikutele. Need filtrid tagasid usaldusväärsed tulemused edasiseks funktsionaalseks tõlgendamiseks, sealhulgas FUMA analüüsiks. Metaanalüüsi kaasamiseks pidid kohordid vastama järgmistele tingimustele:

- uuringus osalesid Euroopa päritolu isikud;
- PsA juhtumid olid määratletud rahvusvahelise haiguste klassifikatsiooni (RHK-10) koodide järgi või valideeriti UK Biopanga fenotüübikoodi alusel;
- genoomi joondamine toimus ühtse referentsgenoomi GRCh38 alusel;
- GWAS-i koondstatistikud sisaldasid vähemalt efektisuurust (beeta), standardviga (SE) ja p-väärtust;
- uuringutes ei olnud indiviide, kes oleksid osalenud rohkem kui ühes kaasatud andmestikus;
- andmestikud olid filtreeritud lähtuvalt kvaliteedinäitajatest, sealhulgas MAF ≥ 0.01, variandi *call rate* $\geq$ 95% ja Hardy-Weinbergi tasakaalu p-väärtus $\geq 10^{-4}$ (kui kohaldatav).

2.2.5 GWAS metaanalüüsi tulemuste annotatsioon ja visualiseerimine FUMA platvormi kaudu

GWAS metaanalüüsi tulemused sisaldasid iga testitud SNP kohta põhjalikku teavet − sealhulgas positsiooni, efektialleeli, alternatiivset alleeli, kombineeritud efektisuurust (beeta), selle standardviga ning p-väärtust. Sageli paiknevad paljud statistiliselt olulised SNP-d (p-väärtus < 5 x 10-8) intergeensetes piirkondades, mis muudab nende bioloogilise tõlgendamise keeruliseks. Selliste variantide funktsionaalse tähenduse selgitamiseks kasutati FUMA (*Functional Mapping and Annotation*) veebiplatvormi, mis integreerib andmeid erinevatest

bioloogilistest andmebaasidest, et tuvastada tõenäolised põhjuslikud SNP-d ja geenid (Watanabe *et al*., 2017). Käesolevas töös kasutati FUMA v1.5.2 versiooni[7].

Esimeses analüüsietapis liigitab FUMA variandid sõltumatuteks olulisteks SNP-deks ja juht-SNP-deks ning määrab genoomsed riskilookused. Selleks kasutatakse genoomiülese assotsiooniuuringu summaarstatistikat ning referentspaneeli (1000 Genomes Project, faas 3), mille abil arvutatakse ahelduse tasakaalutuse (LD) seosed SNP-de vahel. Sõltumatuteks olulisteks SNP-deks loetakse need, mis saavutavad genoomiülese olulisuse taseme ($p < 5 \times 10^{-8}$) ega ole omavahel tugevas LD seoses ($r^2 < 0,6$). Nendest sõltumatutest olulistest SNP-dest valitakse juht-SNP-d, kasutades veel rangemat LD lävendit ($r^2 < 0,1$), mille eesmärgiks on tagada, et iga juht-SNP esindaks eraldiseisvat seosesignaali võrreldes teiste sõltumatute oluliste SNP-dega. Genoomsete riskilookuste määramiseks ühendatakse juht SNP-de LD plokid ($r^2 \geq 0,6$), mis asuvad üksteisest kuni 250 tuhande aluspaari kaugusel. See lähenemine vähendab andmete korduvust ja koondab SNP-d bioloogilistesse klastritesse, mis annavad sisendi edasiseks geenide kaardistamiseks ja funktsionaalseks annotatsiooniks.

Seejärel kaardistatakse iga genoomse riskilookuse kandidaat SNP-d, kasutades FUMA SNP2GENE funktsiooni. Kandidaat SNP-deks loetakse kõik teadaolevad juht-SNP-d (sõltumata sellest, kas need esinesid GWAS-i summaarstatistikas), mis on tugevas ahelduse tasakaalutuses ($r^2 \geq 0,6$) ühe sõltumatu olulise SNP-ga. See lähenemine tagab, et annotatsiooniprotsess hõlmab funktsonaalselt olulisi variante ka väljaspool neid, mida GWAS-is otseselt testiti. Kaasates ahelduse tasakaalutuse teabe referentspaneelidest, võimaldab FUMA iga lookuse ulatuses põhjalikumalt kandidaatvariante kaardistada. Need kandidaat SNP-d suunatakse edasi funktsionaalseks annotatsiooniks, sealhulgas regulatiivse potentsiaali hindamiseks, genoomse konteksti määramiseks ja geenide kaardistamiseks. Need filtreerimisetapid andsid tulemuseks usaldusväärse ja bioloogiliselt tähendusliku kandidaatvariantide komplekti. Seejärel läbisid need variandid põhjaliku funktsionaalse annoteerimise, et võimaldada hiljem PsA patogeneesis potentsiaalselt osalevaid põhjuslikke geene tuvastada.

[7] https://fuma.ctglab.nl/

FUMA viib läbi ulatusliku kandidaat juht-SNP-de funktsionaalse annotatsiooni, integreerides andmeid mitmest allikast. ANNOVAR-i abil klassifitseeritakse SNP-d vastavalt nende genoomsele paiknemisele (nt intronites, eksonites, intergeensetes piirkondades) ja potentsiaalsele mõjule geeni funktsioonile. CADD-skooridega hinnatakse üksikute variantide võimalikku kahjulikkust, samas kui RegulomeDB skoorid näitavad nende tõenäosust mõjutada transkriptsioonifaktorite seondumist ja regulatoorset aktiivsust. *Roadmap Epigenomics* projekti lisatakse epigenoomiandmed ning kromatiininteraktsiooni kaardid Hi-C (*high-throughput chromosome conformation capture*) uuringutest, et paigutada SNP-d regulatoorsesse konteksti. Lisaks kasutakse eQTL andmestikke – eriti GTEx projektist –, et seostada geneetilisi variante geenieksпressiooni muutustega haigusega seotud kudedes[8].

Kandidaat juht-SNP-de ja nendega seotud geenide vahelisi seoseid visualiseeritakse FUMA poolt genereeritud *circos*-diagrammide abil. Need ringikujulised joonised annavad integreeritud ülevaate ahelduse tasakaalutuse struktuurist, geenide kromosomaalsest paigutusest, kromatiininteraktsioonidest ning eQTL-seostest, võimaldades keerukaid genoomseid seoseid lihtsamini tõlgendada. Erinevat tüüpi funktsionaalsete seoste tähistamiseks kasutatakse eristavaid värvikoode, mis lihtsustab selliste geenide tuvastamist, mida toetavad mitmed andmekihid. Selline visualiseerimisstrateegia suurendab GWAS-tulemuste tõlgendatavust ja aitab valida geene edasiseks valideerimiseks.

[8] https://fuma.ctglab.nl

Kuigi FUMA platvorm pakub põhjalikku ülevaadet keerukate tunnustega seotud geneetilistest variantidest, aitab kandidaatgeenide täiendav tuvastamine paremini mõista, millised bioloogilised protsessid ja rajad võivad haiguse arengus osaleda. FUMA pakub kandidaatgeenide kaardistamiseks kolme erinevat meetodit:

- MAGMA V1.08 geeni-põhine kaardistamine: See meetod kasutab mitmeid lineaarseid regressioonimudeleid, et arvutada geenide p-väärtused. Kaardistamine põhineb SNP-de assotsiatsioonidel uuritava tunnusega ja nendevahelisel ahelduse tasakaalutusel (Yan *et al*., 2025).
- Geeniekspressiooniga seotud lookuste (eQTL) põhine kaardistamine: Selle meetodiga kaardistatakse need geenid, mille ekspressioon on seotud oluliste sõltumatute SNP-dega riskilookustes, samuti nende SNP-dega, mis asuvad nendega LD-s (Watanabe *et al*., 2017).
- Kromatiininteraktsioonidel põhinev kaardistamine: Selles meetodis kasutatakse kandidaatvariantide kromatiininteraktsioone, et määrata, millised geenid on seotud riskilookustega. Geen loetakse kaardistatuks, kui selle promootorpiirkond kattub SNP-ga, mis on teadaolevas kromatiininteraktsioonis kandidaatgeenivariandiga (Watanabe *et al*., 2017).

2.2.6 Psoriaatilise artriidiga seotud võimalike uute riskilookuste määratlemine

FUMA tuvastatud juht-SNP-de alusel määratleti võimalikud uued riskilookused, võrreldes nende genoomset asukohta varasemalt GWAS Kataloogis (EFO_0003778) kirjeldatud PsA-ga seotud SNP-dega. Eelnevad seosed laaditi alla kromosoomi ja aluspaaride positsiooni alusel ning võrdlus viidi läbi R-is. SNP loeti potentsiaalselt uueks seoseks, kui selle asukoht jäi rohkem kui 500 tuhat aluspaari üles- või allapoole ükskõik millisest varem kirjeldatud PsA-ga seotud SNP-st (ehk väljaspool 1 Mb raadius)[9].

Võimalike uute riskilookuste *forest*-diagrammid koostati R-is (versioon 4.3.0), kasutades pakette ggplot2, tidyr ja forcats[10]. PsA-ga seostavate juht-SNP-de mõjuhinnangud, mis metaanalüüsis tuvastati kui võimalikud uued seosed, visualiseeriti eraldi kohortide (EGV,

[9] https://www.pgscatalog.org/search/?q=EFO_0003778
[10] https://www.r-project.org/

UKBB, FinnGen) ja kogu metaanalüüsi lõikes. Šansside suhe (o*dds ratio*) ja 95% usaldusvahemikud arvutati GWAS ja metaanalüüsi summaarstatistikas esitatud beeta koefitsientide ja standardvigade põhjal, kasutades järgmisi valemeid: OR = exp(beeta); CI_lower = exp (beeta – 1.96 x SE); CI_upper = exp (beeta + 1.96 x SE).

2.2.7 Kandidaatvariantide *in silico* annotatsioon FUMA platvormi abil.

Kandidaatvariantide *in silico* annotatsioon viidi läbi FUMA platvormi SNP2GENE analüüsiprotsessi abil. Iga juht-SNP puhul koguti kõik tugevas ahelduse tasakaalutuses olevad SNP-d ($r^2 \geq 0,6$), moodustades seeläbi kandidaatsete SNP-de komplekti. Sealjuures hõlmas see nii GWAS käigus testitud kui ka imputatsiooniga saadud SNP-sid. Need kandidaatvariandid annoteeriti FUMA sisseehitatud tööriistade abil, kasutades järgmisi andmeid:

- CADD skoorid kahjulikkuse prognoosimiseks;
- RegulomeDB skoorid regulatoorse potentsiaali hindamiseks;
- kromatiini seisundi andmed Roadmap Epigenomics projektist;
- genoomne kontekst (intronis, eksonis, intergeenses piirkonnas) ANNOVAR-i kaudu.

Need annotatsioonid aitasid täpsustada tõenäoliste funktsionaalsete variantide nimekirja ja suunasid edasist geenikaardistuse ning tulemuste tõlgendamise protsessi. Integreeritud prioritiseerimisstrateegia võimaldas eristada vaid puhtalt statistiliselt olulisi signaale neist variantidest, millel võib olla tõeline bioloogiline roll PsA patogeneesis.

2.2.8 Geenide prioritiseerimine *Open Targets Genetics* ja *Mouse Genome Informatics* platvormide abil

FUMA abil saadud geenikaardistuse tulemuste täiendamiseks viidi läbi kandidaatgeenide täiendav prioritiseerimine, kasutades OTG platvormi[11]. See platvorm ühendab genoomiülesed assotsiatsiooniandmed mitmekihilise funktsionaalse tõendusmaterjaliga, pakkudes terviklikku raamistikku haiguste arengus tõenäoliselt osalevate geenide tuvastamiseks. OTG platvormi üheks oluliseks väljundiks on V2G skoorid, mis kvantifitseerivad tõenäosust, et konkreetne variant mõjutab kindlalt geeni, tuginedes kombineeritud bioloogilistele andmestikele. Nende

[11] https://genetics.opentargets.org/

hulka kuuluvad geeni lähedus, eQTL-id ja pQTL-id, kromatiini ligipääsetavus, transkriptsioonifaktorite seondumiskohad ning fenotüübilised seosed. Selle platvormi integreeriv lähenemine toetab GWAS signaalidest bioloogiliselt oluliste geenide usaldusväärsemat tuvastamist. V2G skoorid peegeldavad koondatud tõendusmaterjali, mis toetab variandi ja geeni vahelist seost, võimaldades geene prioritiseerida ka siis, kui seostav geen ei paikne juht-SNP läheduses. Paljudel juhtudel võib variant reguleerida kaugemal asuvat geeni kromatiini kaardumise (*chromatin looping*) kaudu või toimida eQTL-ina haigusega seotud koes. Selline lähenemine ületab pelgalt läheduskaalutlustel põhinevaid eeldusi ja toob esile geenide funktsionaalse olulisuse[12].

Siinses analüüsis otsiti OTG portaalist andmeid iga FUMA-s tuvastatud ja varasemate andmetega võrreldes potentsiaalselt uudsena klassifitseeritud juhtsignaali kohta. Iga genoomse lookuse puhul valiti välja kolm kõrgeima V2G skoori geeni, mida käsitleti kandidaatgeenidena edasiseks tõlgendamiseks. Valitud kandidaatgeenide edasiseks hindamiseks uuriti viit toetava tõendusmaterjali tüüpi. Esiteks hinnati kandidaatgeenide *missense*- või muid mittesünonüümsed variante, kasutades FUMA poolt esitatud ANNOVAR annotatsioone. Need variandid paiknesid kas juht-SNP-s või variantides, mis olid juht-SNP-ga tugevas ahelduse tasakaalus ($r^2 > 0,6$). Teiseks hinnati füüsilist kaugust juht-SNP ja iga geeni transkriptsiooni alguskoha vahel; kui kaugus oli alla 500 tuhande aluspaari, käsitleti seda lisatõendusena regulatoorsest potentsiaalist (Pujol Gualdo *et al*., 2023). Kolmandaks tehti kindlaks, kas variant toimis kandidaatgeenide eQTL-ina haigusega seotud kudedes, kasutades selleks OTG platvormi ja/või FUMA-st pärit eQTL-andmeid ning füüsilist kaugust juhtvariandist. Asjakohaste kudede hulka kuulusid immuunsüsteemi funktsioneerimisega seotud veri, põrn, lümfisõlmed ning nahk ja sünoviaalkude — kõik süsteemid, mis on seotud PsA-i patofüsioloogiaga.

Neljandaks hinnati FUMA integreeritud kromatiininteraktsioonide andmestikke kasutades, kas kromatiininteraktsioonide andmed seostasid genoomse lookuse kandidaatgeeniga. See oli oluline ka juhul, kui eQTL tõendus oli olemas, kuna kromatiininteraktsioonid annavad lisateavet ruumiliste regulatsioonimehhanismide kohta, mida ainult järjestuse lähedus ei suuda edasi anda. Need regulatoorsed seosed on sageli olulised kaugmõjuga geenide kontrolli

[12] https://genetics.opentargets.org/

mõistmisel, eriti immuunsüsteemi ja liigeseid puudutavates kudedes. Kromatiinil põhinevat tõendusmaterjali hinnati seetõttu süstemaatiliselt kõikide lookuste puhul, mitte ainult siis, kui eQTL andmed puudusid. Sellise lähenemise eesmärk oli, et ükski potentsiaalne regulatoorne seos ei jääks märkamata ning lisas seega usaldusväärsust kandidaatgeenide prioritiseerimise tulemustesse.

Viimaks hinnati MGI andmebaasi, kasutades geenide inaktiveerimise fenotüübilisi tagajärgi. See ressurss võimaldab uurida geenide funktsiooni, analüüsides fenotüüpe hiiremudelites, kus geen on välja lülitatud (*knockout,* KO). MGI on eriti väärtuslik, kuna paljud hiire geenid on homoloogsed inimese geenidega ning nende funktsioonikadu võib peegeldada haigusmehhanisme, mis on olulised ka inimestel. Siinses uuringus otsiti hiirte fenotüübilisi muutusi, mis sarnanevad PsA-i põdevatel patsientidel täheldatavate sümptomitega, nagu liigesepõletik, naha kõrvakalded või immuunsüsteemi talitlushäired[13].

Kandidaatgeenide prioritiseerimiseks kasutati järgmisi kriteeriume: esmalt valiti geenid, mille seotust toetas rohkem kui üks sõltumatu tõenduskiht ning mille KO-hiiremudelites täheldati PsA-ga seotud fenotüübilisi muutusi. Sellest grupist kitsendati valikut edasi, lähtudes kriteeriumide olulisuse järjekorrast: 1) geenid, millel oli täiendav regulatoorne tõendusmaterjal või mida märgistasid kodeerivad GWAS signaalid; 2) geenid, mis paiknesid GWAS signaalile lähedal või millel esines kromatiininteraktsioone. Selline lähenemine võimaldas tuvastada geenid, mille seotust PsA-ga toetavad mitmed sõltumatud tõendusallikad, sealhulgas KO-hiirte fenotüübid ja regulatoorsed või funktsionaalsed mehhanismid.

[13] https://www.informatics.jax.org/

Kokkuvõttes võimaldas see mitmeetapiline prioritiseerimisprotsess paremini mõista, millised geenid on PsA-ga tõenäoliselt funktsionaalselt seotud. Integreerides geneetilisi assotsiatsioonisignaale erinevat tüüpi bioloogiliste andmetega, tuvastati geenid, mida toetasid mitmekesised ja sõltumatud tõendusallikad. See strateegia võimaldas eristada üksnes seostunud SNP-de läheduses paiknevaid geene neist, millel on tegelik funktsionaalne roll haiguse arengus. FUMA, OTG ja MGI platvormide kasutamine lisas geenitõlgendusele sügavust ja aitas esile tuua mehhanistlikke seoseid geneetilise varieeruvuse ja haiguse fenotüübi vahel. Prioritiseeritud geenid pakuvad tugeva aluse edasiseks funktsionaalseks valideerimiseks ja võimalike ravisihtmärkide uurimiseks PsA-i kontekstis.

2.2.9 MAGMA koespetsiifiline geenikspressiooni analüüs

Koespetsiifilise geenikspressiooni analüüsi tulemusi kujutav diagramm annab ülevaate sellest, millistes kudedes avalduvad PsA-ga seotud geenid statistiliselt oluliselt, viidates võimalikule seosele nende kudede ja haiguse vahel. Analüüs põhineb GTEx v8 andmestikul ja on tehtud FUMA platvormil, kasutades MAGMA statistilist raamistikku, mis võimaldab hinnata geeni tasemel seoseid erinevates koetüüpides. Diagrammi tulemused aitavad mõista, millistes kudedes on need geenid, millel on tugevam geneetiline seos PsA-ga, ka keskmiselt kõrgemalt ekspresseeritud, viidates nende kudede võimalikule rollile haiguse patogeneesis. See teave on kasulik mitte ainult haiguse patofüsioloogia paremal mõistmisel, vaid ka potentsiaalsete biomarkerite tuvastamisel ja sihitud raviviiside arendamisel[14].

[14] https://fuma.ctglab.nl/

2.3 Tulemused

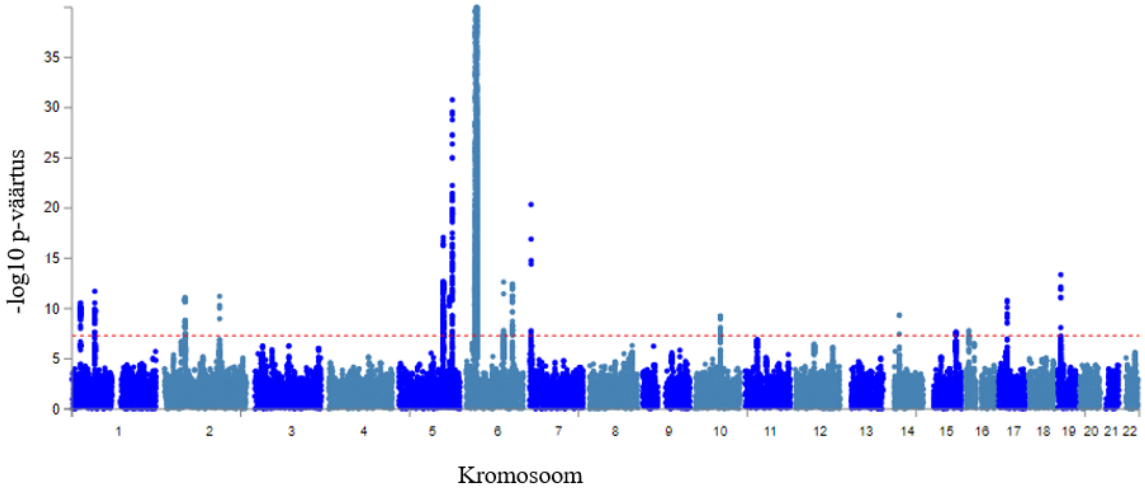
2.3.1 Genoomsete riskilookuste iseloomustamine

PsA-ga seotud genoomsete piirkondade tuvastamiseks analüüsiti esmalt kolme kohordi GWAS-tulemusi: EGV, FinnGen ja UKBB. Iga kohordi summaarstatistika läbis eraldi kvaliteedikontrolli ning viidi vastavusse GRCh38 referentsgenoomiga. Kuigi üksikutes kohortides tuvastati mõned genoomiüleselt olulised variandid, piiras väiksem valimimaht statistilist võimekust uute võimalike PsA seoste tuvastamisel. Seetõttu kombineeriti tulemused metaanalüüsis, et suurendada tundlikkust ja tuvastada potentsiaalselt uusi seoseid.

Lõplik GWAS metaanalüüs, mis hõlmas 5870 PsA juhtu ja 865 276 kontrollisikut Euroopa päritolu populatsioonidest, andis aluse edasiseks geneetiliste seoste kaardistamiseks FUMA SNP2GENE metoodika abil. Kõigepealt tuvastati metaanalüüsi tulemuste põhjal 120 sõltumatut olulist SNP-d, mis ületasid genoomiülese olulisuse lävendi ($p < 5 \times 10^{-8}$) ega olnud omavahel tugevas ahelduses ($r^2 < 0,6$). Nende hulgast määrati 43 juht-SNP-d (*lead* SNPs) − need on kõige tugevama statistilise seosega variandid, mis olid üksteise suhtes veelgi rangemalt sõltumatud ($r^2 < 0,1$). Seejärel nende 43 juht-SNP-de põhjal määratleti 20 eraldiseisvat genoomset riskilookust, tuginedes SNP-de füüsilisele kaugusele (kuni 250 tuhat aluspaari) ja ahelduse struktuurile ($r^2 \geq 0,6$). Kõik tuvastatud lookused ületasid Bonferroni korrektsiooni alusel kehtestatud genoomiülese olulise lävendi ($p < 5 \times 10^{-8}$) ning paiknesid mitmel kromosoomil. Lisaks määratleti FUMA kaudu 2554 kandidaatvarianti, millest 1982 olid olemas ka senises GWAS metaanalüüsi andmestikus. Selline samm-sammuline lähenemine võimaldas tuvastada kõige olulisemad genoomipiirkonnad, mis võivad olla seotud PsA geneetilise eelsoodumusega.

Metaanalüüsi tulemuste visualiseerimiseks koostati FUMA platvormi abil Manhattani diagramm (joonis 3, lk 34), mis kajastab kogu genoomi lõikes SNP-de statistilise olulisuse jaotust. PsA-ga seostunud riskilookused ning neid esindavad juht-SNP-d koos oma kromosomaalsete asukohtadega on kokku võetud tabelis 3 (lk 35). Iga riskilookuse puhul on esitatud peamine juht-SNP – variant, millel oli vastavas piirkonnas tugevaim statistiline seos ning mille alusel vastav lookus määratleti. Tuleb siiski arvestada, et üks riskilookus võib sisaldada lisaks ka teisi juht-SNP-sid ning hulgaliselt muid seotud variante, mida tabelis ei kuvata. Kuigi tabelis 3 on näha ka tugev signaal *HLA* piirkonnas (kromosoom 6), mis on varasemates uuringutes koorduvalt seotud PsA-ga, ei käsitleta seda piirkonda käesolevas töös põhjalikumalt. Töö keskendub mitte-*HLA* piirkondadele, et uurida uusi potentsiaalseid

geneetilisi seoseid, kuna *HLA* piirkonna keerukas ahelduse struktuur raskendab põhjuslike variantide eristamist.



Joonis 3. PsA metaanalüüsi tulemusi kajastav Manhattani joonis (N= 871146). X-teljel on geneetiliste markerite asukohad kogu genoomis, samas kui Y-teljel on vastavad p-väärtuste negatiivsed kümnendlogaritmid. Punane katkendlik joon tähistab korrigeeritud olulisuse piiri mitmese testi puhul (p-väärtus $\leq 5 \times 10^{-8}$). Joonise kohandamisel eemaldati väga kõrged p-väärtused (-log10PVAL > 40), et visuaalne kujutus oleks selgem ja arusaadavam. Joonisel on esitatud iga genoomse riskilookuse üks peamine juht-SNP (*lead* SNP), mis tähistab kõige tugevama seosega varianti konkreetsetes sõltumatus piirkonnas ning erineb teistest juht-SNP-dest selle poolest, et esindab kogu lookust. Manhattani joonis on loodud kasutades FUMA veebiplatvormi

Tabel 3. PsA-ga seostunud genoomsed riskilookused ja neid esindavad juhtvariandid. Tabelis on esitatud 20 eraldiseisvat genoomset riskilookust, mis tuvastati GWAS metaanalüüsi põhjal FUMA SNP2GENE analüüsi abil. Iga lookuse kohta on toodud kromosoom (*Chr*), kromosoomalne asukoht ja juhtvariant – see on kõige tugevama statistilise seosega SNP vastavas piirkonnas, mille alusel riskilookus defineeriti. Tuleb arvestada, ei iga lookus võib sisalda ka muid olulisi SNP-sid, mida siin tabelis ei kuvata.

Nr	Genoomne lookus (kromosoom:asukohad)	Juht-SNP (esindajad)	Chr	Asukoht (aluspaari)	p-väärtus
1	1:25294607:C:T	rs7542079	1	25294607	2.83×10^{-11}
2	1:67617594:C:G	rs12079083	1	67617594	1.95×10^{-12}
3	2:61072567:C:T	rs35741374	2	61072567	7.95×10^{-12}
4	2:62551472:A:G	rs10865331	2	62551472	3.03×10^{-08}
5	2:163167746:C:G	rs17715343	2	1,63E+08	6.11×10^{-12}
6	5:131996500:A:C	rs848	5	1,32E+08	8.21×10^{-18}
7	5:150469973:A:G	rs74817271	5	1,5E+08	7.14×10^{-12}
8	5:158826493:A:G	rs918518	5	1,59E+08	1.69×10^{-31}
9	6:29611885:C:T	rs3131854	6	29611885	2.78×10^{-27}
10	6:33564296:A:T	rs12206050	6	33564296	1.02×10^{-11}
11	6:33973941:C:G	rs10947473	6	33973941	1.83×10^{-09}
12	6:111913262:C:T	rs33980500	6	1,12E+08	2.34×10^{-13}
13	6:138239434:C:T	rs6933987	6	1,38E+08	3.72×10^{-13}
14	7:5500455:C:T	rs149256681	7	5500455	1.24×10^{-17}
15	10:75601596:C:T	rs2459446	10	75601596	5.76×10^{-10}
16	14:35871093:C:T	rs696	14	35871093	4.46×10^{-10}
17	15:85671755:A:G	rs7177110	15	85671755	2.14×10^{-08}
18	16:11443183:A:G	rs11643024	16	11443183	1.59×10^{-08}
19	17:26127518:C:G	rs10459953	17	26127518	1.59×10^{-11}
20	19:10463118:C:G	rs34536443	19	10463118	4.33×10^{-14}

2.3.2 Võimalike uute riskilookuste määratlemine

Kaheksa potentsiaalselt uut juht-SNP-d valiti nende seast, mis kuulusid 43 genoomiüleselt olulise ja omavahel sõltumatu ($r^2 < 0,1$) SNP hulka. Need SNP-d tuvastati varasema analüüsi käigus ning osutusid uudseteks, kuna need ei paiknenud teadaolevate PsA-ga seotud SNP-de (EFO_0003778) läheduses 1 miljoni aluspaari (Mb) raadiuses (lähtudes GWAS kataloogi andmetest). Kokku tuvastati kaheksa juht-SNP-d, mida võib potentsiaalselt pidada uuteks seosteks PsA-ga, lähtudes asjaolust, et neid ei ole varasemates PsA GWAS-uuringutes kirjeldatud. Kuna mõned neist SNP-dest paiknesid samades genoomsetes riskilookustes, esindavad need kokku viit unikaalset võimalikku uut riskilookust (lookused 9, 11, 15, 17 ja 18). Need lookused paiknesid kromosoomidel 6, 10, 15 ja 16 ning nende kõigi juhtvariandid vastasid nii genoomiülese olulisuse lävendi nõudele ($p < 5 \times 10^{-8}$) kui ka sõltumatuse kriteeriumile ($r^2 < 0,1$) üksteise suhtes. Oluline on märkida, et mitte kõik kaheksa juht-SNP-d ei esine FUMA poolt määratletud genoomsete riskilooksute tabelis (vt tabel 3, lk 35), sest FUMA esitab iga lookuse kohta vaid ühe juhtvariandi – ehk tugevaima seosega juht-SNP. Samas võib iga riskilookus sisaldada mitmeid muid statistiliselt olulisi või funktsionaalselt huvipakkuvaid SNP-sid, mida visualiseeritakse FUMA väljundites, kuid mis ei ole kaasatud kokkuvõtvasse tabelisse. Näiteks SNP rs62402015 ei ole tabelis 3, kuigi see oli üks kaheksast funktsionaalselt tugevast variandist, kuna vastavas lookuses oli veelgi tugevaima seosega SNP, mis FUMA algoritmi alusel määrati lookust esindavaks.

Nende 8 lookuste kokkuvõte – sealhulgas kromosomaalsed asukohad, juht-SNP-d ja muud statistilised näitajad – on esitatud tabelis 4 (lk 37). Et hinnata potentsiaalselt uute juht-SNP-de efektide järjepidavust erinevates kohortides, koostati R-is (versioon 4.3.0) *forest*-diagramm (vt lisa 2). Diagramm kujutab iga juht-SNP puhul šansside suhet ja 95% usaldusvahemikke EGV, FinnGen, UKBB ja kogu metaanalüüsi lõikes ning näitab, et enamik variante avaldas mõju sarnases suunas ja ulatuses, mis tugevdab nende võimalikku olulisust. Mitmed juhtvariandid näitasid tugevat statistilist seost PsA-ga, näiteks rs145245084 (p-väärtus = 4.95×10^{-10}), rs10947473 (p-väärtus = 1.83×10^{-9}) ja rs2459446 (p-väärtus = 5.76×10^{-10}). Teiste variantide puhul, nagu rs62402015 ja rs9461362, olid p-väärtused lähedased genoomiülese olulise lävendi väärtusele (vastavalt 3.38×10^{-8} ja 3.44×10^{-8}), kuid siiski ületasid selle, viidates võimalikule rollile haiguse kujunemisel.

Tabel 4. PsA GWAS-i metaanalüüsi tulemused valitud 8 juht-SNP kohta, mida peetakse võimalikeks uuteks seosteks. Need SNP-id pärinevad 43 juht-SNP seast ja valiti välja, kuna neid ei ole varasemates PsA GWAS-ides kirjeldatud või neil oli tugev funktsionaalne tõendusmaterjal (näiteks eQTL, kromatiininteraktsioonid, KO-hiiremudelite andmed). EAF – efektialleeli sagedus *(effect allele frequency)*; OR – šansside suhe (*odds ratio*); 95% CI − šansside suhte 95%-usaldusintervall (*confidence interval*).

Juht-SNP	Genoomne riskilookus	Kromosoom: positsioon	Efektialleel/ teine alleel	EAF	OR (95% CI)	p-väärtus
rs62402015	9	6:29611885	T/C	0.785	0.87 (0.83-0.92)	3.38×10^{-8}
rs9461362	9	6:27303927	T/C	0.148	1.16 (1.10-1.22)	3.44×10^{-8}
rs145245084	9	6:28867034	A/G	0.012	1.65 (1.41-1.94)	4.95×10^{-10}
rs58243442	11	6:33964907	G/A	0.027	1.36 (1.22-1.52)	2.73×10^{-8}
rs10947473	11	6:33973941	C/G	0.016	1.55 (1.34-1.79)	1.83×10^{-9}
rs2459446	15	10:75601596	T/C	0.523	0.89 (0.85-0.92)	5.76×10^{-10}
rs7177110	17	15:85671755	A/G	0.761	1.13 (1.08-1.18)	2.14×10^{-8}
rs11643024	18	16:11443183	G/A	0.703	1.13 (1.08-1.18)	1.59×10^{-8}

2.3.3 Kandidaatgeenide prioritiseerimine FUMA, *Open Target Genetics* ja *Mouse Genome Informatics* abil.

PsA-ga seotud geneetiliste riskitegurite tuvastamiseks viidi läbi põhjalik kandidaatgeenide prioritiseerimine, integreerides mitmekihilisi andmestikke FUMA platvormi, OTG ja MGI abil. Iga potentsiaalselt uueks määratletud riskilookuse puhul otsiti selles paikneva juht-SNP andmed OTG platvormist ning valiti välja kolm kõrgeima V2G skooriga geeni iga juht-SNP kohta. Seejärel hinnati neid geene täiendavalt, kasutades funktsionaalseid andmekihte FUMA-st, OTG ja MGI andmebaasist. Eriti pöörati tähelepanu juht-SNP-de vahetus läheduses asuvatele geenidele, kuna nende paiknemine signaalipiirkonnas võib anda esialgsed vihjeid haiguse mehhanismide kohta. Lisaks analüüsiti riskilookustes asuvate variantide mõju geenieksпressioonile (eQTL-andmed), mis aitab mõista, kuidas geneetiline varieeruvus võib mõjutada vastavate geenide aktiivsust. Geenide prioritiseerimise täiendamiseks kasutati MGI andmebaasi, et uurida kandidaatgeenide võimalikke fenotüübilisi toimeid hiiremudelites. Samuti võeti arvesse FUMA-s esitatud kromatiininteraktsioonide andmeid, et tuvastada kaugmõjusid geenide regulatsioonile. Need visualiseeriti *circos*-diagrammides (vt joonis 5, lk 40 ja lisa 3). Kõiki neid tegureid arvesse võttes kaardistati potentsiaalselt haigusega seotud kandidaatgeenid ning tulemused on kokku võetud tabelis 5 (lk 41). Geenid, mida toetasid mitmed sõltumatud andmekihid ning mille puhul ilmnes eeldatav funktsionaalne seos haigusega, loeti potentsiaalselt põhjuslikeks geenideks.

2.3.4 *Variant-to-gene* skoori tõendusmaterjal

Geenide prioritiseerimiseks nende potentsiaalse funktsionaalse olulisuse alusel kasutati OTG platvormi V2G skoori. Genoomses lookuses 9 kromosoomil 6 (rs62402015, rs9461362 ja rs145245084) valiti välja järgmised geenid: *HMGN4, ZNF322, H2BC11, ZNF391, ABT1, POM121L2, TRIM27* ja *H2BC121*. Lookuses 11 kromosoomil 10 (rs58243442 ja rs10947473) valiti geenid: *LEMD2, UHRF1BP1, MLN* ja *IP6K3*. Lookuses 15 (rs2459446) prioritiseeriti geenid: *CAMK2G, PLAU* ja *NDST2*. Lookuses 17 kromosoomil 15 (rs7177110) valiti: *PDE8A, AKAP13* ja *SLC28A1*. Viimaks valiti lookuses 18 kromosoomil 16 (rs11643024) geenid: *SOCS1, RMI2* ja *DEXI*.

2.3.5 Tõendusmaterjal geeni paiknemise kohta juht-SNP läheduses

Funktsionaalse olulisuse toetamiseks uuriti, kas eelnevalt valitud kandidaatgeenid paiknesid iga uue genoomse riskilookuse juht-SNP-de läheduses. Geene käsitleti positsiooniliselt asjakohastena, kui nende transkriptsiooni alguskoht (TSS) asus juht-SNP-st kuni 500 tuhande

aluspaari kaugusel. Sellele kriteeriumile vastas mitu geeni, sealhulgas *ZNF322, HMGN4, H2BC11, ZNF391, POM121L2* ja *TRIM27* (kromosoom 6, lookuses 9); *LEMD2, MLN* ja *IP6K3* (kromosoom 6, lookus 11); *CAMK2G, PLAU* ja *NDST2* (kromosoom 10, lookus 15); *AKAP13, SLC28A1* ja *PDE8A* (kromosoom 15, lookus 17) ning *RMI2, SOCS1* ja *DEXI* (kromosoom 16, lookus 18). Neid ruumilisi seoseid kasutati täiendava tõendusmaterjalina edasistes analüüsides geenide ja PsA vaheliste seoste hindamisel.

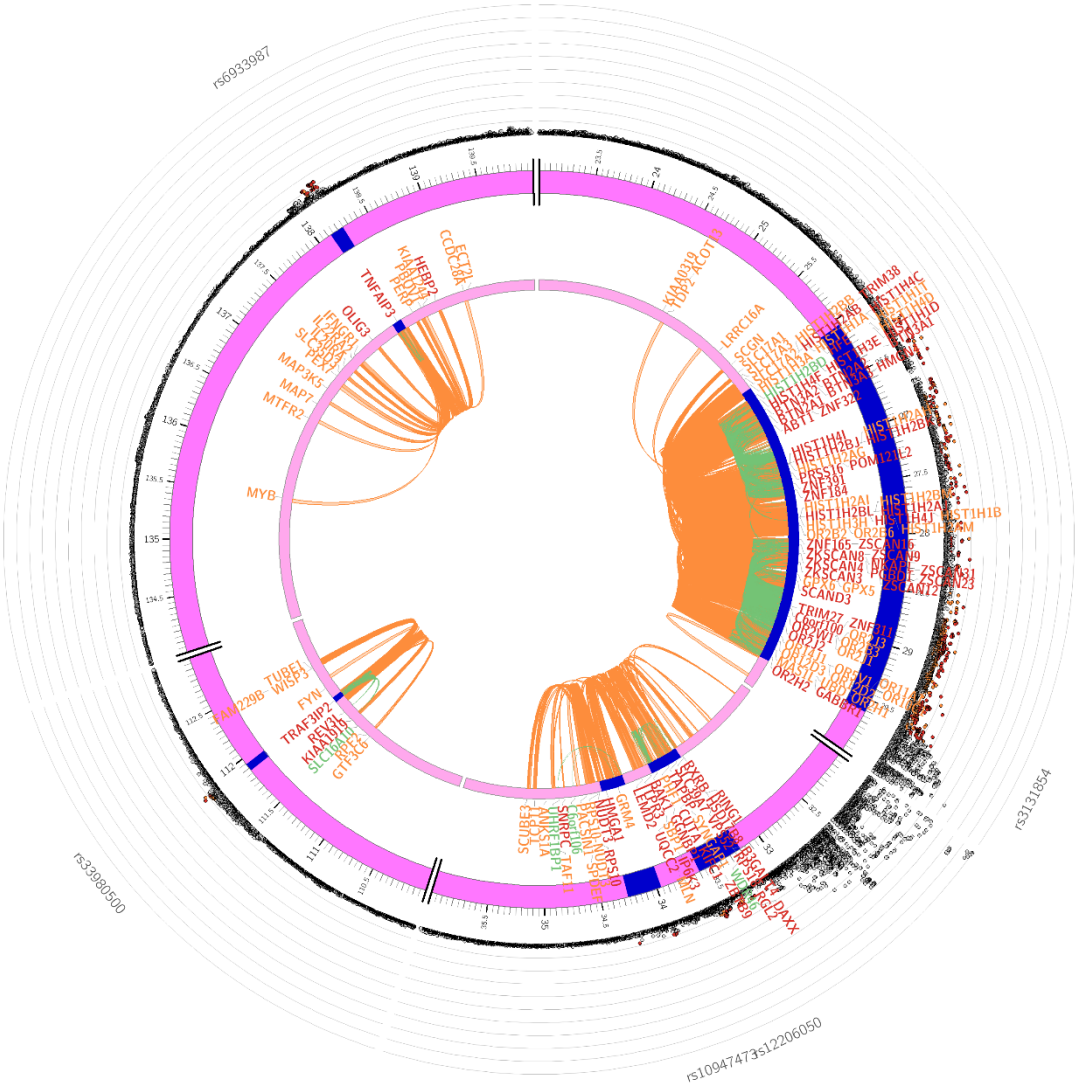
2.3.6 Tõendusmaterjal eQTL-seoste kohta

Kandidaatvariantide võimalikku mõju geeniekspressioonile hinnati eQTL andmete alusel, kasutades OTG platvormi. Iga genoomse riskilookuse puhul uuriti, kas juht-SNP-d või nende asevariandid ($r^2 > 0,6$) toimivad eQTL-idena asjakohastes kudedes, sealhulgas veres, lümfikoes, nahas ja sünoviaalkoes – kõik olulised süsteemid, mis on seotud PsA patogeneesiga. Kandidaatvariantide vahendatud eQTL-efekte täheldati mitmete geenide puhul: *HMGN4, ZNF322, ZNF391* ja *ABT1* kromosoomi 6 lookuses 9; *UHRF1BP1* kromosoomi 6 lookuses 11; *CAMK2G* kromosoomi 10 lookuses 15; *PDE8A* kromosoomi 15 lookuses 17 ning *RMI2, SOCS1* ja *DEXI* kromosoomi 16 lookuses 18. Need regulatoorsed seosed viitavad sellele, et variandid võivad mõjutada haiguse riski, moduleerides geeniekspressiooni bioloogiliselt olulistes kudedes. Geenid, millel esines eQTL-tõendus, loeti funktsionaalselt oluliseks ning kaasati täiendava tõendusmaterjalina edasisse analüüsi.

2.3.7 Tõendusmaterjal kromatiininteraktsioonide kohta

Kuna kromatiininteraktsioonid (visualiseeritud oranžide ja punaste joontega FUMA genereeritud *circos*-diagrammides) näitavad DNA elementide ruumilist lähedust rakutuumas, kajastavad need tänapäevast arusaama geeniregulatsioonist kui ruumilisest ja dünaamilisest protsessist, mitte lihtsalt järjestusepõhisest nähtusest. Selleks et teha kindlaks, kas eelvalitud kandidaatgeenid interakteeruvad PsA-ga seotud riskilookustega kromatiini struktuuri kaudu, uuriti FUMA-s integreeritud kromatiininteraktsioonide andmeid.

Kokku näitasid mitmed eelvalitud geenid kromatiininteraktsioonide olemasolu juht-SNP-dega seotud genoomipiirkondades. Need olid *HMGN4, ZNF322, ZNF391, ABT1, POM121L2* ja *TRIM27* genoomses lookuses 9; *UHRF1BP1, LEMD2, MLN* ja *IP6K3* lookuses 11; *CAMK2G, PLAU* ja *NDST2* kromosoomi 10 lookuses; *PDE8A, AKAP13* ja *SLC28A1* kromosoomi 15 lookuses ning *RMI2, SOCS1* ja *DEXI* kromosoomi 15 lookuses. Konkreetseid geenidevahelisi interaktsioone on illustreeritud joonisel 5 (lk 40; vt ka lisa 3).



Joonis 5. PsA-ga seotud 6. kromosoomi geneetilised seosed. GWAS metaanalüüsi põhjal loodud *circos*-diagramm kujutab 6. kromosoomi riskilookusi. Välimisel kihil on Manhattani graafik, kus SNP-d p-väärtusega ≤ 0,05 on tähistatud värvidega vastavalt nende ahelduse tasakaalutusele (r^2): punane (>0,8), oranž (>0,6), roheline (>0,4), sinine (>0,2) ja hall (≤0,2). Keskmises kihis on kromosoomi struktuur ja riskilookused (sinine). Sisemises kihis on kandidaatgeenid: rohelised (eQTL), oranžid (kromatiininteraktsioonid) ja punased (mõlemad). Rohelised ja oranžid kaared näitavad vastavalt eQTL ja kromatiininteraktsiooni seoseid. Diagramm loodi FUMA platvormil, mis aitab tuvastada haigusega seotud geneetilisi piirkondi. Mitmete SNP-ide füüsilise lähestikuse paiknemisetõttu ei pruugi kõik rsID-d *circos*-diagrammil nähtavale ilmuda. See on tingitud osaliselt visualiseerimispiirangutest (nähtavuse ja loetavuse tagamiseks) ning osaliselt FUMA platvormi seadistusest, mille kohaselt kuvatakse igas riskilookuses vaid üks juhtvariant.

Tabel 5. PsA GWAS-i metaanalüüsis tuvastatud potentsiaalsed uued riskilookused ning nendega seotud geenide funktsionaalne prioritiseerimine erinevate kriteeriumite alusel. Tabelis on esitatud kaheksa juht-SNP, mis esindavad viit unikaalset uut genoomset riskilookust. Geenide prioritiseerimine põhineb kombinatsioonil erinevatest funktsionaalsetest näitajatest, sealhulgas kaugus variandist, eQTL, kromatiininteraktsioonide olemasolu, seosed ennustava funktsionaalse tähtsusega (nt regulatoorne mõju) ning fenotüübiline seos hiiremudelites. Mustad täpid (●) aitavad visuaalselt eristada, millised geenid ja variandid on võimaliku bioloogilise mõju alusel prioriteetsemad.

	Genoomilookus	Kolm suurima V2G väärtusega geeni (Open Targets)	Kaugus <500kb	Kaugus <100kb	eQTL kaardistamine vastavates kudedes	Kromatiininteraktsioonid	Funktsionaalne ennustamine (kodeeriva järjestuse tagajärjed)	Asjakohane fenotüüp knockout hiirtel	
rs62402015	9	*HMGN4* *ZNF322* *H2BC11*	● ● ●		● ●	● ●			chr6:29611885
rs9461362	9	*ZNF391* *ABT1* *POM121L2*	● ●	● ●	● ●	● ● ●		●	chr6:27303927
rs14524508	9	*TRIM27* *H2BC12* *H2BC11*	●	●		●		●	chr6:28867034
rs58243442	11	*LEMD2* *UHRF1BP* *MLN*	● ●		●	●		●	chr6:33964907
rs10947473	11	*LEMD2* *MLN* *IP6K3*	● ● ●			● ● ●		●	chr6:33973941
rs2459446	15	*CAMK2G* *PLAU* *NDST2*	● ● ●	● ● ●	●	● ● ●		● ● ●	chr10:75601596
rs7177110	17	*PDE8A* *AKAP13* *SLC28A1*	● ● ●		●	● ● ●		●	chr15:85671755
rs11643024	18	*RMI2* *SOCS1* *DEXI*	● ● ●	● ●	● ● ●	● ● ●		●	chr16:11443183

2.3.8 *Missense*-variantide tõendusmaterjal

Missense-variantidel, mis põhjustavad aminohappe muutusi valku kodeerivates piirkondades, võib olla otsene funktsionaalne mõju geeniproduktidele ning seetõttu peetakse neid tugevateks kandidaatideks haigusega seotud mehhanismide puhul. Osana variantide annotatsiooniprotsessist hinnati kõiki juht-SNP-sid ja tugevas ahelduse tasakaalutuses olevaid SNP-sid ($r^2 \geq 0,6$) kodeerivate tagajärgede osas, kasutades FUMA platvormi integreeritud ANNOVAR-i annotatsioone.

See analüüs keskendus spetsiaalselt mittesünonüümsete kodeerivate variantide – eriti *missense*-mutatsioonide – tuvastamisele kandidaatgeenides või nende vahetus läheduses igas riskilookuses. Siiski ei olnud juht-SNP-sid ega nende asendusvariante üheski tuvastatud genoomses riskilookuses *missense*-variantidena annoteeritud. *Missense*-variantide puudumine viitab sellele, et enamik selles uuringus tuvastatud geneetilisi seoseid avaldab tõenäoliselt mõju regulatoorsete mehhanismide kaudu − näiteks mõjutades geenikspressiooni läbi tugevdajaelementide (*enhancer element*), kromatiinistruktuuri muutuste või transkriptsioonifaktorite seondumiskohtade. See rõhutab vajadust integreerida muu funktsionaalse tõendusmaterjali, näiteks eQTL-seoste ja kromatiininteraktsioonide kasutamist kandidaatgeenide prioritiseerimisel PsA kontekstis.

2.3.9 Fenotüübiline tõendusmaterjal hiire geeni väljalülitamise (*knockout*) mudelitest

MGI andmebaasi põhjal näitasid mitmed kandidaatgeenid kromosoomidel 6, 10, 15 ja 16 PsA-le sarnaseid fenotüüpe hiire geeni väljalülitamise (*knockout*) mudelites. Riskilookustes 9 ja 11, mis mõlemad asuvad kromosoomil 6, näitasid geenid *POM121L2, TRIM27* ja *LEMD2* fenotüübilisi toimeid, mis olid seotud immuunsüsteemi talitlushäirete, skeleti kõrvalekallete või põletikuliste protsessidega. Kromosoomil 10 olid geenid *CAMK2G*, *PLAU* ja *NDST2* seotud immuunvastuse reguleerimise ja lihasluukonna arengu regulatoorsete mehhanismidega. Kromosoomil 15 paiknev geen *AKAP13* näitas potentsiaalset osalust PsA-ga seotud signaaliradades. Kromosoomil 16 tuvastati, et geenil *SOCS1* esinevad fenotüübid viitavad selle võimalikule seotusele autoimmuunsete protsesside ja liigesefunktsioonidega.

Need leiud toetavad GWAS metaanalüüsis tuvastatud eelvalitud geenide funktsionaalset olulisust. Hiiremudelite ja inimese haigusilmingute fenotüübiline kattuvus pakub bioloogilist tõendusmaterjali ning tõhustab geenide prioritiseerimist PsA-i patogeneesi kontekstis (vt tabel 5, lk 41).

2.3.10 Prioritiseerimise tõendusmaterjali kokkuvõte oluliste lookuse kohta

Eelnevalt käsitletud erinevad tõendusastmed PsA patogeneesis osalevate geenide kohta on koondatud tabelis 5 (lk 41). Analüüsis tuvastati kolm juht-SNP-d: rs62402015, rs9461362 ja rs145245084, mis paiknevad kromosoomi 6 riskilookuses 9. Mitmed geenid kromosoomi 6 riskilookuses (lookus 9) olid toetatud erinevate sõltumatute funktsionaalsete annotatsioonidega. *HMGN4, ZNF322* ja *ZNF391* paiknesid juht-SNP läheduses ja näitasid kromatiininteraktsioone ning ekspressioonimõjusid asjakohastes kudedes, nagu veri ja immuunorganid. *TRIM27* ja *POM121L2* olid samuti toetatud eQTL ja kromatiinipõhiste andmetega ning lisaks ilmnes nende geenide seos skeleti- ja immuunsüsteemi muutustega hiiremudelis vastavalt MGI andmetele. Geenid *H2BC11, H2BC12* ja *ABT1* esinesid samas piirkonnas, kuid neile tuginev funktsionaalne tõendusmaterjal oli piiratum – *H2BC11* puhul peamiselt asukohaline seos juht-SNP-ga ning *ABT1* puhul regulatoorsed annotatsioonid.

Genoomses riskilookuses 11 (kromosoomil 6) tuvastati 2 juhtvarianti – rs58243442 ja rs10947473 –, mis vastasid genoomiülese olulisuse kriteeriumile ja näitasid tugevat sõltumatust. Need SNP-d kaardistusid geenide *LEMD2, UHRF1BP1, MLN* ja *IP6K3* lähedusse. *LEMD2, MLN* ja *IP6K3* paiknesid juhtvariantide vahetus läheduses ja näitasid kromatiininteraktsioone, viidates potentsiaalsetele regulatoorsetele seostele. *UHRF1BP1* seost toetasid nii eQTL-andmed asjakohastes kudedes kui ka kromatiini interaktsioonid, saades OTG platvormil selles lookuses ühe kõrgeima V2G skoori. Lisaks on *LEMD2* geeni osas täheldatud hiiremudelis fenotüübilisi muutusi, mis on seotud skeleti ja immuunsüsteemi funktsioonidega. Kuigi mitmetel geenidel selles piirkonnas on teatud määral toetav funktsionaalne tõendusmaterjal, on hetkel ebaselge, milline neist kannab peamist bioloogilist rolli PsA kujunemisel. Seetõttu vajab see piirkond edasist funktsionaalset uurimist, et selgitada välja geenide vaheline hierarhia ja täpsustada haiguse mehhanisme. Oluline on märkida, et rs10947473 paikneb väljaspool klassikalist *HLA* piirkonda, mis viitab võimalikule *HLA*-st sõltumatule mehhanismile (vt tabel 5, lk 41).

Kromosoomil 10 paiknev genoomne riskilookus 15 hõlmas juhtvariant rs2459446, mille läheduses asuvad geenid *CAMK2G, PLAU* ja *NDST2.* Kõik kolm geeni paiknesid juht-SNP lähedal ja näitasid kromatiininteraktsioone, viidates võimalikele kaugregulatoorsetele mõjudele (vt lisa 3). *CAMK2G* ja *PLAU* puhul ilmnesid ka eQTL seosed, mis viitavad nende geenide ekspressiooni mõjutamisele haigusega seotud kudedes. Lisaks *CAMK2G* geenile tuvastati ka *PLAU* ja *NDST2* selles lookuses, mille kohta leiti MGI andmetes fenotüübilisi muutusi hiiremudelites, mis on seotud immuunfunktsiooni ja skeletiarenguga. *CAMK2G* sai OTG

platvormil ühe kõrgema V2G skoori selles piirkonnas, millele lisandus tugev toetus eQTL, kromatiininteraktsiooni ja hiiremudeli andmetest. Need leiud rõhutavad *CAM2G* potentsiaalselt olulist rolli PsA patogeneesis, eriti rakusisese signaaliedastuse ja põletikulise vastuse regulatsiooni kaudu (vt tabel 5, lk 41).

SNP rs7177110, mis asub kromosoomil 15 genoomses riskilookuses 17, osutus analüüsis potentsiaalselt kõige olulisemaks funktsionaalseks variandiks. See SNP paikneb tihedas ahelduse tasakaalutuse plokis ($r^2 > 0,8$) koos mitme sõltumatu olulise SNP-ga ning on seotud mitmete kromatiininteraktsioonidega. FUMA *circos*-diagrammi (vt lisa 3) kohaselt sisaldab rs7177110 ümbritsev piirkond mitmeid tugevalt seotud variante (märgitud punaselt) ning rikkalikku regulatoorset annotatsiooni, sealhulgas kromatiinisilmuseid (oranžid ja punased kaared). See variant seostub geenidega *PDE8A, AKAP13* ja *SLC28A1*, mis kõik osalevad signaaliedastuses ja immuunregulatsioonis. Nendest geenidest paistis eriti silma *PDE8A,* mida toetasid mitmed sõltumatud funktsionaalsed andmekihid, sealhulgas kromatiinistruktuuri seosed ja eQTL-analüüsid. *PDE8A* osaleb tsüklilise AMP (cAMP) metabolismis, mille häired võivad mõjutada immuunfunktsioone ja põletikuliste protsesside regulatsiooni. *AKAP13* puhul täheldati funktsionaalseid muutusi hiiremudelites − selle geeni väljalülitamine põhjustas häireid immuunsüsteemi töös, luustruktuuris ja põletikulise tasakaalu reguleerimises. Kuigi *SLC28A1* kohta esines vähem funktsionaalset tuge, näitas ka see geen kromatiinseoseid juhtvariandiga. Need tulemused viitavad sellele, et rs7177110 ei ole mitte ainult statistiliselt oluline, vaid ka funktsionaalselt tähenduslik ning võib oma mitmekihiliste regulatoorsete toimetega aidata kaasa PsA-i kujunemisele (vt tabel 5, lk 41).

Kromosoomi 16 genoomsest piirkonnast (lookus 18) tuvastati statistiliselt oluline SNP rs11643024, mis on tugevas ahelduse tasakaalus ($r^2 > 0,8$) teiste sama piirkonna sõltumatute tippvariantidega. Selle variandi esiletõus FUMA *circos*-diagrammi välisringis (vt lisa 3) viitab tugevale seosele fenotüübilise varieeruvusega. Rs11643024 on seotud kolme geeniga – *SOCS1, RMI2* ja *DEXI* (vt tabel 5, lk 41) –, millel on kas tõendatud või oletatav roll immuunregulatsioonis ja genoomse stabiilsuse säilitamises. Kõiki kolme geeni toetas eQTL-seos haigusega seotud kudedes, sealhulgas veres, lümfoidkoes ja nahas, viidates sellele, et rs11643024 võib mõjutada nende geenide ekspressiooni bioloogiliselt asjakohastes kontekstides. *SOCS1* on üks peamisi tsütokiinidest vahendatud signaaliradade negatiivseid regulaatoreid ning selle funktsioonihäired võivad põhjustada mitmesuguseid põletikulisi ja autoimmuunseid seisundeid. Kromatiininteraktsioonid rs11643024 ja *SOCS1* vahel viitavad sellele, et see variant võib mõjutada *SOCS1* geenikespressiooni, millel on potentsiaalset mõju immuunvastuse dünaamikale. *DEXI* on samuti immuunmodulatsiooniga ja on varem seostatud

mitmete autoimmuunhaigustega, sealhulgas I tüüpi diabeediga. *RMI2* on seotud DNA replikatsiooni ja parandamise mehhanismidega ning mängib võtmerolli genoomse terviklikkuse säilitamises, osutades sellele, et rs11643024 mõju võib ulatuda immuunvastuse regulatsioonist kaugemale. Lisaks on *SOCS1* funktsionaalne tähtsus kinnitatud hiiremudelite abil – geeni väljalülitamine põhjustab häired nii immuunfunktsioonides kui ka skeleti arengus. Arvestades mitme geeni (näiteks *POM121L2, TRIM27, LEMD2, CAMK2G, AKAP13* ja *SOCS1*) puhul täheldatud mitmekihilist regulatiivsed tuge, võib neid pidada tugevateks kandidaatideks.

2.3.11 MAGMA koespetsiifiline geeniekspressiooni analüüs

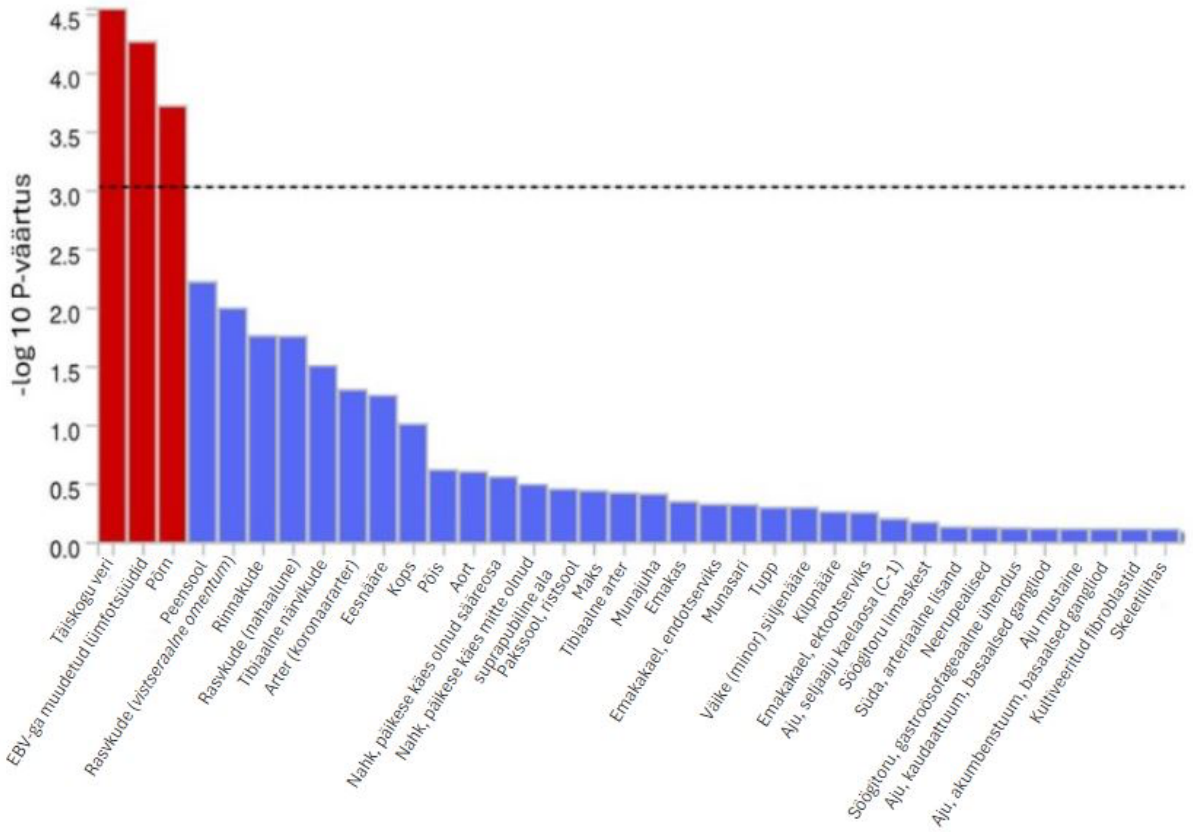
MAGMA analüüs keskendus kudedele, millel võib olla bioloogiline roll uuritava fenotüübi kujunemisel. Geeniekspressiooni mustreid hinnati 53 erinevas koes, kasutades GTEx v8 andmestikku. Analüüsis tõsteti esile näiteks perifeerne veri, põrn, EBV-transformeeritud lümfotsüüdid, tibiaalne närvikude, nahk (nii päikesele eksponeeritud sääreosa kui ka mitte eksponeeritud suprapubiline piirkond), tibiaalne arter, seljaaju kaelaosa (C1) ja skeletilihas (vt joonis 9, lk 47). Olulise tulemusena selgus, et PsA seotud geenide ekspressioon oli statistiliselt oluliselt rikastunud immuunseotud kudedes, sealhulgas perifeerses veres, põrnas ja EBV-transformeeritud lümfotsüütides. Suuremaid -log10(p) väärtusi täheldati just neis kudedes, mis viitab tugevale seosele immuunsüsteemi ja PsA patogeneesi vahel.

Nõrgema signaaliga kudede osas tuleb märkida, et ükski neist ei ületanud statistilise olulisuse piiri (-log10 $p > 3$). Siiski näitasid tibiaalne närvikude ja päikesele eksponeeritud nahk mitmetest teistest kudedest suuremaid väärtusi. Selline muster viitab võimalikule, kuid mitte statistiliselt olulisele geeniekspressiooni ja fenotüübi vahelisele seosele, mistõttu ei saa välistada nende kudede rolli PsA-i patogeneesis.

Närvikoes täheldatud mõõdukas signaal ei pruukinud osutuda statistiliselt oluliseks koe keeruka rakulise koostise tõttu. Geeniekspressiooni muutused võivad piirduda spetsiifiliste neuronitüüpidega, mida koe keskmine ekspressioonitase ei kajasta. Kuna MAGMA analüüs ei erista rakkude alatüüpe, võivad lokaalsed mõjud jääda märkamatuks. Samuti võivad mõned geenid olla aktiivsed ainult teatud haiguse staadiumites, mida olemasolev andmestik ei hõlma. Sellest tulenevalt ei saa järeldada, et bioloogiline seos närvikoes puudub.

Tibiaalne arter ja skeletilihas paigutusid analüüsi tulemuste alumisse ossa, viidates vähesele seosele uuritavate geneetiliste variantidega. Samuti ei näidanud seljaaju kaelaosa (C1) − hoolimata selle potentsiaalsest neuroloogilisest tähtsusest − statistiliselt olulist ega mõõdukat signaali. Kuigi nende kudede otsene seos PsA patogeneesiga võib olla ebatõenäoline, ei saa

välistada kaudseid mõjusid või osalust muude regulatsioonitasandite kaudu. Edasised funktsionaalsed uuringud võiksid aidata selgitada, milline roll võib neil kudedel haiguse kujunemises olla.



Joonis 9. MAGMA koespetsiifiline geenieksрessiooni analüüs PsA riskigeenide seose alusel. Katkendlik joon tähistab statistilise olulise piiri, millest kõrgemad väärtused viitavad kudedele, kus geenide seos PsA-ga on tõenäoliselt usaldusväärne.

2.4 Arutelu

Magistritöö eesmärk oli tuvastada PsA-ga seotud geneetilisi riskitegureid, kasutades ulatuslikku genoomiüleste assotsiatsiooniuuringute metaanalüüsi, et saadud teadmistest tulenevalt mõista paremini PsA patogeneesi. Suurima võimaliku statistilise võimsuse saavutamiseks kasutati analüüsi sisendina kolme Põhja-Euroopa päritolu kohorti: EGV, FinnGeni ja UKBB. Selline mitme populatsiooni andmestike ühendamine suurendas uurimuse usaldusväärtust ning võimaldas tuvastada geneetilisi seoseid, mida väiksemates kohortides ei pruugi piisavalt sageli esineda. Metaanalüüsiga tuvastati 43 juht-SNP-d, millest kaheksa osutusid potentsiaalselt uuteks ning paiknesid neljal kromosoomil (6, 10, 15, 16). Need leiud loovad aluse edasisteks uuringuteks, mis võiksid aidata paremini mõista PsA geneetilist arhitektuuri ja selle aluseks olevaid mehhanisme.

Tugevaimad seosed leiti kromosoomil 6, kus paiknesid kaks eraldi genoomset riskilookust – lookus 9 ja 11. Lookuses 9 kaardistati genoomiüleselt olulised juht-SNP-d geenide *TRIM27, ZNF322, HMGN4, ZNF391* ja *H2BC11* läheduses, samas kui lookuses 11 seostusid SNP-d geenidega: *LEMD2, UHRF1BP1, MLN* ja *IP6K3*. Need leiud viitavad sellele, et kromosoomil 6 on mitu sõltumatut piirkonda, mis võivad mängida olulist rolli PsA geneetilises etioloogias. Geenide prioritiseerimise käigus mitmete sõltumatute tõenduskihtide põhjal tõusis lookuses 9 esile eriti *TRIM27* kui tugev kandidaatgeen. Seda toetasid nii kromatiininteraktsioonid kui ka hiiremudelite andmed, mis näitasid skeletti ja immuunsüsteemi mõjutavaid fenotüübilisi muutusi. Lisaks on *TRIM27* teadaolevalt seotud T-rakkude aktivatsiooni regulatsiooniga ning varasemates uuringutes on seda seostatud ka autoimmuunsete häiretega (Hatakeyama, 2017). Lookuses 11 paistis *LEMD2* silma kui tugev kandidaat, kuna selle geeni puhul täheldati kromatiininteraktsioone, lähiasetust juht-SNP-dega ning PsA-ga seotud fenotüüpe hiire *knockout*-mudelis. Ülejäänud geenidest oli *UHRF1BP1* toetatud peamiselt regulatiivsete andmetega, *MLN* ja *IP6K3* pigem läheduskaardistuse kaudu, ilma ulatusliku funktsionaalse toetuseta. Seetõttu võib *LEMD2* pidada selles piirkonnas kõige tõenäolisemaks funktsionaalselt tähenduslikuks geeniks. Sarnast lähenemist kandidaatgeenide prioritiseerimisel, kus toetakse mitmele sõltumatutele andmekihtidele, on rakendatud ka varasemat PsA GWAS-analüüsides (Soomro *et al*., 2022).

Nimetatud geenide funktsionaalne olulisus viitab nende võimalikule rollile immuunradade moduleerimisel, mis on olulised PsA kujunemisel. Tulemused toetavad hüpoteesi, et MHC piirkonna immuunregulatoorsetel geenidel võib olla keskne roll PsA-i arengus ning et haigus hõlmab süsteemset immuundüseregulatsiooni, mitte ei mõjuta ainult nahka või liigeseid.

SNP-d, mis paiknesid kromosoomidel 10, 15 ja 16, kaardistusid samuti geenidele, millel on seos immuun- ja põletikuliste protsessidega. Näiteks rs2459446 (kromosoom 10) asus geenide *CAMK2G* ja *PLAU* läheduses, mis osalevad rakusiseses signaaliedastuses ja fibrinolüüsis (Yu *et al*., 2024; Cook *et al*., 2010). Geneetiliste ja funktsionaalsete tõendite integreerimise tulemusena prioritiseeriti selles lookus *CAMK2G*, kuna see geen oli seotud eQTL-seostega asjakohastes kudedes, näitas kromatiininteraktsioone ning selle häiritus põhjustas hiiremudelites põletiku ja skeletistruktuuride häired. *CAMK2G* kodeerib seriin/treoniin kinaasi, mis vahendab signaaliedastust Ca^{2+}-sõltuvates radades ja omab keskset rolli T-rakkude aktivatsioonis ja põletikuprotsesside moduleerimises (Yu *et al*., 2024). Kromosoomil 15 paiknes rs7177110 geenide *PDE8A* ja *AKAP13* lähedal, millest mõlemad reguleerivad cAMP signaalirada – rada, mis on seotud tsütokiinitasakaalu säilitamisega (Epstein *et al*., 2021; Toaldo *et al*., 2015). Selles piirkonnas prioritiseeriti *PDE8A,* kuna see geen oli ainus, millel oli toetus nii eQTL-analüüsist kui ka kromatiinpõhistest interaktsioonidest. *PDE8A* lagundab cAMP-i, mille tase mõjutab põletikutsütokiinide tootmist (Epstein *et al*., 2021). Selle geeniga seostuvad mehhanismid võivad olla PsA põletikuprotsessides kriitilise tähtsusega. *AKAP13* puhul täheldati samuti kromatiininteraktsioone ning lisaks ilmnesid hiiremudelite põhjal PsA-ga seonduvad fenotüübid, mis viitavad selle geeni potentsiaalsele rollile haiguse arengus. Võttes arvesse mitmekihilist funktsionaalset tõendusmaterjali, võib mõlemat geeni − *PDE8A* ja *AKAP13* – käsitleda selle piirkonna tugevaimate kandidaatidena. Kromosoomil 16 asus rs11643024 geeni *SOCS1* läheduses, mis on tsütokiinide signaalide ja immuunrakkude aktivatsiooni negatiivne regulaator. *SOCS1* mängib olulist rolli põletikuliste signaaliradade pärssimisel ning selle geen oli toetatud nii eQTL-seostega, kromatiininteraktsioonidega kui ka hiiremudelitega, kus selle häiritus põhjustas immuunvastuse düsregulatsiooni ja skeletisüsteemi muutusi (Yoshimura *et al*., 2007). Seetõttu prioritiseeriti *SOCS1* kui tugevaima funktsionaalse tõendusmaterjaliga kandidaat selles piirkonnas.

Need seosed viitavad sellele, et PsA kujunemist võivad käivitada kattuvad põletikulised signaalirajad, mis hõlmavad nii omandatud kui ka kaasasündinud immuunsust. Sellised mehhanismide kattuvused toovad esile haiguse patogeneesi keerukuse ja rõhutavad vajadust mitmekesistele sihtmärkidele suunatud lähenemiste järele edasistes funktsionaalsetes uuringutes. Kuigi mitmes riskipiirkonnas paikneb arvukalt huvipakkuvaid geene, ei ole praegu veel võimalik kindlalt öelda, kas kõik need geenid või vaid osa neist on tõepoolest seotud PsA arenguga.

MAGMA tarkvara abil tehtud koespetsiifilise ekspressiooni analüüs näitas, et PsA-ga seotud geenid, mis kaardistati kogu GWAS-i geenipõhise analüüsi käigus FUMA platvormi abil,

avalduvad kõige tugevamalt immuunsüsteemiga seotud kudedes, nagu põrn, lümfisõlmed ja luuüdi. See toetab hüpoteesi, et PsA geneetilised signaalid mõjutavad olulisel immuunrakkude funktsiooni ning rõhutab haiguse süsteemset ja põletikulist iseloomu (Veale ja Fearon, 2018; Rendon ja Schäkel, 2019). Lisaks täheldati mõõdukat geenikespressiooni ka nahas ja sidekoes, mis võib viidata sellele, et PsA geneetiline alus mõjutab erinevaid koetüüpe, kajastudes haiguse mitmekesistes ilmingutes – alates liigesepõletikust kuni dermatoloogiliste sümptomiteni. Kuigi need koed ei ületanud statistilise olulise piiri, ei saa välistada nende rolli haiguse patogeneesis, eriti arvestades kliinilisi tunnuseid. Kokkuvõttes toetab see koetasandil saadud tõendusmaterjal geenide prioritiseerimist edasistes funktsionaalsetes valideerimisuuringutes.

Võrdlus varasemate GWAS-uuringute tulemustega aitab paremini mõtestada siinse uuringu tulemusi. Mitmed varem kirjeldatud PsA-ga seostuvad lookused leidsid taas kinnitust, hõlmates sealhulgas geene *TRAF3IP2*, *IL12B* ja *TNFAIP3* (Liu *et al*., 2008; Soomro *et al*., 2022; Aterido *et al*., 2018). Samas tuvastati siinses uuringus ka uusi lookusi – näiteks geenidega *HMGN4, CANK2G, PDE8A* ja *SOCS1* seotud piirkonnad – mida varasemates PsA GWAS-uuringutes ei ole kirjeldatud. Need võivad viidata kas populatsioonispetsiifilistele riskivariantidele või varem tuvastamata signaalidele, mis eelnevates uuringutes madala statistilise võimsuse tõttu ei ilmnenud. Kui varasemates uuringutes näitasid mõned *HLA* piirkonna SNP-d suurt mõju (OR > 2), siis paljud selles uuringus tuvastatud uued seosed näitasid mõõdukat mõju (OR ~ 1,2-1,4). See kajastab PsA polügeenset olemust, kus haiguse kujunemist ei mõjuta vaid üksikud suure efektiga variandid, vaid pigem arvukad väiksema efektiga geneetilised tegurid.

Need võrdlused varasemate andmetega kinnitavad tulemuste usaldusväärsust ning pakuvad samal ajal uusi teadmisi seni uurimata bioloogiliste radade kohta. Nii tuntud kui ka uute lookuste tuvastamine parandab meie arusaama PsA geneetilistest alusest ning rõhutab metaanalüüside väärtust varjatud signaalide avastamisel. Tulevased uuringud peaksid neid uusi lookusi sõltumatutes populatsioonides täiendavalt uurima. Kuigi siinne uuring tuvastas edukalt mitmeid paljutõotavaid genoomseid riskilookusi ja funktsionaalselt toetatud kandidaatgeene, on oluline märkida, et paljudes neis lookustes – eriti kromosoomi 6 *HLA* piirkonnas – paikneb mitu huvipakkuvat geeni. Praegu ei ole veel selge, kas PsA patogeneesiga on tõepoolest seotud kõik need geenid või ainult osa neist. Põhjuslike geenide täpne tuvastamine mõnes genoomipiirkonnas on raskendatud kättesaadavate andmekogumite analüüsimeetodite piirangute tõttu. Seetõttu on vajalik teha täiendavaid põhjalikke analüüse, sealhulgas peenkaartimist ja funktsionaalset valideerimist (näiteks laboripõhiste mudelite abil), et täpsustada geneetilised mehhanismid, mis on PsA kujunemisega tegelikult seotud. Siiski

kujutavad selles uuringus prioritiseeritud geenid endast väärtuslikku bioloogiliselt tõenäoliste kandidaatgeenide kogumit ning pakuvad tugeva lähtekoha edasisteks uuringuteks.

Vaatamata tugevustele tuleb tunnistada ka mitmeid piiranguid. Fenotüübi määratlemine tugines RHK-10 koodidele, mis ei pruugi alati peegeldada haiguse kliinilist alatüüpi või aktiivsust. PsA alatüübid (nt aksiaalne, perifeerne, entesiitiline) võivad geneetiliselt erineda, kuid siinses töös neid ei eristatud. See võib tähendada, et osa signaalidest jäi seetõttu tuvastamata või sai väiksema kaalu. Samuti ei võetud arvesse geneetiliste leidude mõju erinevatele soorühmadele ega potentsiaalseid keskkonnategureid, mis võivad genotüübi-fenotüübi seoseid moduleerida. Need aspektid väärivad tulevastes uuringutes eraldi käsitlemist.

Kokkuvõttes kinnitavad tulemused, et PsA kujunemisel mängivad olulist rolli immuunvastust ja põletikku reguleerivad geneetilised mehhanismid. Töös tuvastatud geneetilised variandid ja kanditaatgeenid võivad aidata suunata tulevasi funktsionaalseid uuringuid ning arendada geneetilisel informatsioonil põhinevaid ravivõimalusi. Eraldi tähelepanu tasuks pöörata mitme signaaliraja ristumiskohtadele, mis võivad olla võtmelülid haiguse arengus. Samuti võiks edasistes töödes süvendatult uurida geneetilise riski interaktsioone elustiiliteguritega, nagu kehakaal või suitsetamine. Selline interaktiivne lähenemine aitaks paremini mõista PsA-keerukat geneetilist ja bioloogilist tausta.

KOKKUVÕTE

Magistritöö eesmärk oli tuvastada psoriaatilise artriidiga (PsA-ga) seotud geneetilisi riskitegureid Põhja-Euroopa päritolu populatsioonides, et mõista paremini nimetatud haiguse patogeneesi. Selleks kasutati genoomiülese assotsiatsiooniuuringu (GWAS) summaarstatistika andmeid kolmest populatsioonipõhisest biopangast: Eesti geenivaramu, Soome FinnGen ja UK Biopank. Seejärel tehti GWAS metaanalüüs, kaasates kokku 5870 PsA juhtu ja 865 276 kontrollisikut, muutes selle üheks suurimaks Põhja-Euroopa kohortidel põhinevaks PsA GWAS metaanalüüsiks. Erinevaid veebiplatvorme ja nende andmebaase, nagu FUMA, *Open Targets Genetics ja Mouse Genome Informatics,* kasutati GWAS metaanalüüsi geneetiliste variantide funktsionaalseks annotatsiooniks ja tulemuste tõlgendamiseks

Metaanalüüsis käigus tuvastati 20 genoomiüleselt olulist riskilookust, millest 15 kattusid varasemates PsA GWAS-uuringutes kirjeldatud seostega (näiteks Liu *et al*., 2008; Soomro *et al*., 2022; Aterido *et al*., 2018). Lisaks avastati viis potentsiaalselt uut riskilookust, millel puudus kattuvus varasemate PsA-ga seostatud piirkondadega GWAS kataloogis (1 Mb ulatuses), viidates nende võimalikule uudsele iseloomule. Need potentsiaalsed uued riskilookused paiknesid kromosoomidel 6, 10, 15 ja 16. Geenide prioritiseerimine potentsiaalselt uutes PsA-ga seotud genoomsetes piirkondades toimus mitmekihilise lähenemisega, mis kombineeris GWAS-signaalid funktsionaalse genoomika andmetega. Eesmärgiks oli tuvastada bioloogiliselt kõige olulisemad kandidaatgeenid. Kõigepealt valiti iga riskilookuse kõrgeima V2G-skooriga geenid *Open Targets Genetics* platvormi alusel. Seejärel täiendati tõlgendust järgmiste kriteeriumide abil: (1) mittesünonüümsete kodeerivate variantide esinemine geeni piirkonnas, (2) juht-SNP-le lähimad geenid (≤500 kb), (3) eQTL-seosed PsA-ga seotud kudedes, (4) kromatiininteraktsioonid geenide promootoritega ja (5) fenotüübid hiire *knockout*-mudelites. Mitmekihilise tõendusmaterjali põhjal tuvastati viis uut geeni, millel on teadaolev roll immuunsüsteemis − *TRIM27, LEMD2, CAMK2G, AKAP13* ja *SOCS1*.

Kokkuvõttes pakub käesolev uuring täiendavat tõendusmaterjali selle kohta, et PsA arengut mõjutavad geneetilised variandid, mis põhjustavad häired immuun- ja põletikuradade regulatsioonis. GWAS metaanalüüsi ja funktsionaalse tõlgenduse tulemusena tuvastati mitu bioloogiliselt ja kliiniliselt olulist kandidaatgeeni, mis võivad olla põhjuslik roll PsA kujunemisel. Need leiud loovad aluse edasistele uuringutele ja toetavad personaalmeditsiini arengut PsA ennetuses, diagnoosimises ja ravis.

Genome-wide association study of psoriatic arthritis

Aleksandra Judin

SUMMARY

Psoriatic arthritis (PsA) is a chronic inflammatory musculoskeletal disease associated with psoriasis, characterized by heterogeneous clinical manifestations and a significant genetic component (Rahman ja Elder, 2005). Genome-wide association studies (GWAS) have previously identified numerous risk loci, particularly in immune-related genes. However, the genetic architecture of PsA in Northern European populations remains incompletely understood. Identifying novel genomic regions associated with PsA could provide deeper insight into disease mechanisms and support the development of precision medicine approaches.

The aim of the Master´s thesis was to identify genetic risk factors associated with PsA in populations of Northern European origin. To achieve this, GWAS summary statistics were collected from three large-scale population-based biobanks: the Estonian Biobank (EGV), the Finnish FinnGen database, and the UK Biobank (UKBB). A fixed-effect meta-analysis was conducted using GWAMA v2.2 with genomic control correction to mitigate potential test statistic inflation, which can result from factors such as population stratification, cryptic relatedness, or technical artifact. The final analysis included 5870 PsA cases and 865 276 controls, making it the largest PsA GWAS meta-analysis carried out to date in Northern European cohorts.

The meta-analysis identified 20 genome-wide significant risk loci, of which 15 overlapped with previously reported associations either in the GWAS Catalog or PsA GWAS literature (e.g. Liu *et al*., 2008; Soomro *et al*., 2022; Aterido *et al*., 2018). Additionally, five potentially novel risk loci were identified that showed no overlap with previously reported PsA-associated regions in the GWAS Catalog (within a 1 Mb window), indicating their possible novel nature. These potential new risk loci were located on chromosomes 6, 10, 15, and 16. Functional annotation of candidate SNPs and mapping of associated genes were performed using the FUMA platform, integrating data from several repositories including gene proximity data, eQTL datasets and chromatin interactions datasets. Circos plots generated via FUMA visualized the complex spatial and regulatory relationship between SNPs and gene targets, reinforcing the importance of long-range interactions in immune regulation.

Gene prioritization in this study was carried out through a systematic multi-layered approach that integrated statistical association signals with functional genomic evidence and phenotypic data from knockout mouse models. For each genomic risk locus identified in the GWAS meta-analysis as potential novel, genes were first prioritized using variant-to-gene (V2G) scores from the Open Targets Genetics platform, which combines data from chromatin accessibility, eQTLs, pQTLs, transcription factor binding, and proximity-based annotations. The three highest-scoring genes per each lead SNP were selected for further interpretation. To refine these results, additional biological evidence was assessed: first, genes were evaluated for the presence of missense or nonsynonymous coding variants based on ANNOVAR annotations, although none of the selected genes carried missense mutations. Second, spatial proximity was considered and genes located within 500 kb of the lead SNP were flagged as positionally relevant, including *ZNF322, TRIM27*, and *SOCS1,* among others. Third, we examined whether the lead SNPs or their proxies ($r^2 > 0.6$) functioned as expression quantitative trait loci (eQTLs) in disease-relevant tissues (blood, lymphoid tissue, skin, and synovial tissue), identifying several candidates genes, such as *HMGN4, CAMK2G, PDE8A, SOCS1* and *DEXI,* among others. Fourth, chromatin interaction data from FUMA was used to detect spatial regulatory links between PsA-associated genomic regions and distal gene promoters; this layer further supported genes like *TRIM27, CAMK2G*, and *SOCS1,* among others. Lastly, functional validation was explored through the Mouse Genome Informatics database, where several genes — including *POM121L2, TRIM27*, *LEMD2, CAMK2G, AKAP13* and *SOCS1,* among others — were associated with phenotypes relevant to psoriatic arthritis, such as skeletal or immune system abnormalities in knockout models.

Based on the multi-step gene prioritization strategy, the novel genes with the strongest evidence supporting their involvement in psoriatic arthritis in this study and with known functions in the immune system are:

- *TRIM27* (locus 9, chromosome 6):
    - Encodes a transcriptional regulator that modulates T cell activation and innate immune responses through ubiquitination pathways, contributing to PsA pathogenesis (Hatakeyma, 2017). Its prioritization is supported by eQTLs, chromatin interactions, and PsA-relevant phenotypes in mouse models.
- *LEMD2* (locus 11, chromosome 6):
    - Encodes a protein of the inner nuclear membrane involved in maintaining nuclear architecture and chromatin organization (Szél et al., 2019). Although less studied in

PsA, its association with immune-related phenotypes in knockout mouse and chromatin interactions data suggests a potential role in immune regulation.

- *CAMK2G* (locus 15, chromosome 10):
    - Encodes a calcium/calmodulin-dependent kinase involved in intracellular signalling, including pathways related to T-cell activation (Yu *et al*., 2024). Its functional importance is supported by high V2G score, chromatin interactions, and skeletal/immune phenotypes in mouse models.

- *AKAP13* (locus 17, chromosome 15):
    - Encodes a A-kinase anchoring protein 13, which coordinates cAMP and cytoskeletal signalling and regulates inflammation (Toaldo *et al*., 2015). Mouse model data and chromatin interaction highlight its relevance to PsA.

- *SOCS1* (locus 18, chromosome 16):
    - Acts as a negative regulator in cytokine-mediated signaling pathways. Its role in dampening inflammatory responses makes it highly for PsA (Yoshimura *et al*., 2007). Strong support comes from chromatin interactions, eQTLs in immune tissues, spatial proximity, and *knockout* mouse phenotypes.

In conclusion, this study provides strong evidence that psoriatic arthritis is influenced by genetic variants involved in immune and inflammatory pathways. By integrating GWAS meta-analysis with functional annotations, and animal phenotype data, this work highlights several potential candidate genes with biological and clinical relevance in PsA. These findings lay the groundwork for future studies focused on replication, functional validation, PsA subtype stratification, and integration of environmental and clinical data. Ultimately, the results contribute to advancing precision medicine approaches for better prevention, diagnosis, and treatment of psoriatic arthritis.

TÄNUSÕNAD

Esmalt soovin südamest tänada oma juhendajaid Ene Reimann, Galadriel Lucia Velazquez Silva ja Tanel Traks nende asjatundliku juhendamise, väärtuslike soovituste eest, suure kannatlikkuse ja pideva toetuse eest kogu tööprotsessi vältel.

Soojad tänud kuuluvad ka mu perele, kes on olnud mulle alati toeks ning julgustanud mind ka kõige keerulisematel hetkedel. Suur tänu mu kallitele sõpradele, kes aitasid hoida motivatsiooni ja vaimujõudu lõpuni.

KASUTATUD KIRJANDUS

Alesa, D. I., Alshamrani, H. M., Alzahrani, Y. A., Alamssi, D. N., Alzahrani, N. S., and Almohammadi, M. E. (2019). The role of gut microbiome in the pathogenesis of psoriasis and the therapeutic effects of probiotics. Journal of Family Medicine and Primary Care, 8(11), 3496–3503.

Anandarajah, A. P. and Ritchlin, C. T. (2009). The diagnosis and treatment of early psoriatic arthritis. Nature Reviews Rheumatology, 5(11), 634–641.

Aterido, A., Cañete, J. D., Tornero, J., … Julià, A. (2018). Genetic variation at the glycosaminoglycan metabolism pathway contributes to the risk of psoriatic arthritis but not psoriasis. Annals of the Rheumatic Diseases, 78(3), e214158.

Barnas, J. L. and Ritchlin, C. T. (2015). Etiology and pathogenesis of psoriatic arthritis. Rheumatic Disease Clinics of North America, 41(4), 643-663.

Bowes, J., Eyre, S., Flynn, E., … Barton, A. (2011). Evidence to support IL-13 as a risk locus for psoriatic arthritis but not psoriasis vulgaris. Annals of the Rheumatic Diseases, 70(6), 1016–1019.

Boyle, E. A., Li, Y. I. and Pritchard, J. K. (2017). An expanded view of complex traits: From polygenic to omnigenic. Cell, 169(7), 1177–1186.

Browning, S. R. and Browning, B. L. (2007). Rapid and Accurate Haplotype Phasing and Missing-Data Inference for Whole-Genome Association Studies By Use of Localized Haplotype Clustering. Am J Hum Genet, 81(5): 1084–1097.

Bycroft, C., Freeman, C., Petkova, D., … Marchini, J. (2018). The UK Biobank resource with deep phenotyping and genomic data. Nature, 562(7726), 203–209.

Chandran, V., Bull, S. B., Pellett, F. J., Ayearst, R., Rahman, P. and Gladman, D. D. (2013). Human leukocyte antigen alleles and susceptibility to psoriatic arthritis. Human Immunology, 74(10), 1333–1338.

Ciccacci, C., Latini, A., Perricone, C., ... Borgiani, P. (2019). TNFAIP3 gene polymorphisms in three common autoimmune diseases: Systemic lupus erythematosus, rheumatoid arthritis, and primary Sjogren syndrome—Association with disease susceptibility and clinical phenotypes in Italian patients. Journal of Immunology Research, 2019, 6728694.

Cook, A. D., De Nardo, C. M., Braine, E. L., Turner, A. L., Vlahos, R., Way, K. J., Beckman, S. K., Lenzo, J. C., Hamilton, J. A. (2010). Urokinase-type plasminogen activator and arthritis progression: role in systemic disease with immune complex involvement. Arthritis Research & Therapy, 12(2), R37.

Cush, J. J. (2022). Rheumatoid arthritis: Early diagnosis and treatment. Rheumatic Disease Clinics of North America, 48(2), 537–547.

Dennis, E. A. and Norris, P. C. (2015). Eicosanoid storm in infection and inflammation. Nature Reviews Immunology, 15(8): 511–523.

Ellinghaus, E., Ellinghaus, D., Stuart, P. E., … Franke, A. (2010). Genome-wide association study identifies a psoriasis susceptibility locus at TRAF3IP2. Nature Genetics, 42(11), 991–995.

Epstein, P. M., Basole, C. and Brocke, S. (2021). The role of PDE8 in T cell recruitment and function in inflammation. Frontiers in Cell and Developmental Biology, 9, 636778.

FitzGerald, O., Ogdie, A., Chandran, V., … Mease, P. J. (2021). Psoriatic arthritis. Nature Reviews Disease Primers, 7(1), 59.

Gaffen, S. L., Jain, R., Garg, A. V. and Cua, D. J. (2014). IL-23–IL-17 immune axis: Discovery, mechanistic understanding, and clinical testing. Nature Reviews Immunology, 14(9), 585–600.

González-Moreno, I. M., Martínez-Fernández, M., García-Gerónimo, A., Suazo-Aravena, M. F., Márquez-Argente Del Castillo, B. (2021). Mutilating arthritis: A typical manifestation of a rare disease. Journal of Clinical Rheumatology, 27(8S), S849–S850.

Hatakeyama, S. (2017). TRIM family proteins: roles in autophagy, immunity, and carcinogenesis. Trends in Biochemical Sciences, 42(4), 297–311.

Jordan, C. T., Cao, L., Roberson, E. D. O., … Bowcock, A. M. (2012). PSORS2 is due to mutations in CARD14. American Journal of Human Genetics, 90(5), 784–795.

Kaeley, G. S., Eder, L., Aydin, S. Z., Gutierrez, M., and Bakewell, C. (2018). Dactylitis: A hallmark of psoriatic arthritis. Seminars in Arthritis and Rheumatism, 48(2), 263–273.

Kurki, M. I., Karjalainen, J., Palta, P., … Palotie, A. (2023). FinnGen provides genetic insights from a well-phenotyped isolated population. Nature, 613(7944), 508–518.

Lappalainen, T. and MacArthur, D. G. (2021). From variant to function in human disease genetics. Science, 373(6562), 1464–1468.

Leitsalu, L., Haller, T., Esko, T., … Metspalu, A. (2015). Cohort Profile: Estonian Biobank of the Estonian Genome Center, University of Tartu. International Journal of Epidemiology, 44(4): 1137–1147.

Liu, Y., Helms, C., Liao, W., … Bowcock, A. M. (2008). A genome-wide association study of psoriasis and psoriatic arthritis identifies new disease loci. PLoS Genetics, 4(3), e1000041.

Mägi, R. and Morris, A. P. (2010). GWAMA: software for genome-wide association meta-analysis. BMC Bioinformatics, 11(1): 288.

Mbatchou, J., Barnard, L., Backman, J., ... Marchini, J. (2021). Computationally efficient whole-genome regression for quantitative and binary traits. Nat Genet, 53(7): 1097—1103.

McGonagle, D., Kavanaugh, A., McInnes, I. B., ... Ritchlin, C. (2024). Association of the clinical components in the distal interphalangeal joint synovio-entheseal complex and subsequent response to ixekizumab or adalimumab in psoriatic arthritis. Rheumatology (Oxford), 63(11), 3115–3123.

Mitt, M., Kals, M., Pärn, K., … Palta, P. (2017). Improved imputation accuracy of rare and low-frequency variants using population-specific high-coverage WGS-based imputation reference panel. Eur J Hum Genet, 25(7): 869–876.

Nestle, F. O., Kaplan, D. H. and Barker, J. (2009). Psoriasis. New England Journal of Medicine, 361(5), 496–509.

Ogdie, A. and Weiss, P. (2015). The Epidemiology of Psoriatic Arthritis. Rheumatic Disease Clinics of North America, 41(4), 545–568.

Ogdie, A., Coates, L. C. and Gladman, D. D. (2020). Treatment guidelines in psoriatic arthritis. Rheumatology (Oxford), 59(Suppl 1), i37–i46.

Poddubnyy, D. and Rudwaleit, M. (2012). Early spondyloarthritis. Rheumatic Disease Clinics of North America, 38(2), 387–403.

Pujol Gualdo, N., Estonian Biobank Research Team, Mägi, R., and Laisk, T. (2023). Genome-wide association study meta-analysis supports association between MUC1 and ectopic pregnancy. Human Reproduction, 38(12), 2516–2525.

Pujol-Gualdo, N., Läll, K., Lepamets, M., Rossi, H.-R., Arffman, R. K., Piltonen, T. T., Mägi, R. and Laisk, T. (2022). Advancing our understanding of genetic risk factors and potential personalized strategies for pelvic organ prolapse. Nat Commun, 13: 3584.

Rahman, P. and Elder, J. T. (2005). Genetic epidemiology of psoriasis and psoriatic arthritis. Annals of the Rheumatic Diseases, 64(Suppl 2), ii37–ii39.

Rendon, A. and Schäkel, K. (2019). Psoriasis pathogenesis and treatment. International Journal of Molecular Sciences, 20(6), 1475.

Ritchlin, C. T., Colbert, R. A. and Gladman, D. D. (2017). Psoriatic arthritis. New England Journal of Medicine, 376(10), 957–970.

Scott, I. C., Whittle, R., Bailey, J., … Jordan, K. P. (2022). Rheumatoid arthritis, psoriatic arthritis, and axial spondyloarthritis epidemiology in England from 2004 to 2020: An observational study using primary care electronic health record data. The Lancet Regional Health – Europe, 23, 100519.

Simard, M., Morin, S., Ridha, Z., and Pouliot, R. (2022). Current knowledge of the implication of lipid mediators in psoriasis. Frontiers in Immunology, 13, 961107.

Soomro, M., Stadler, M., Dand, N., ... Barton, A. (2022). Comparative genetic analysis of psoriatic arthritis and psoriasis for the discovery of genetic risk factors and risk prediction modeling. Arthritis & Rheumatology, 74(9), 1535–1543.

Stuart, P. E., Nair, R. P., Tsoi, L. C., … Elder, J. T. (2015). Genome-wide association analysis of psoriatic arthritis and cutaneous psoriasis reveals differences in their genetic architecture. The American Journal of Human Genetics, 97(6), 816–836.

Szél, E., Bozó, R., Hunyadi-Gulyás, É., Manczinger, M., Szabó, K., Kemény, L., Bata-Csörgő, Z., & Groma, G. (2019). Comprehensive proteomic analysis reveals intermediate stage of non-lesional psoriatic skin and points out the importance of proteins outside this trend. Scientific Reports, 9, 11382.

Tiwari, V. and Brent, L. H. (2024). Psoriatic Arthritis. In StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing.

Toaldo, C. B., Alexi, X., Beelen, K., … Zwart, W. (2015). Protein Kinase A-induced tamoxifen resistance is mediated by anchoring protein AKAP13. BMC Cancer, 15, 588.

Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H.C., Lappalainen, T. and Posthuma, D. (2021). Genome-wide association studies. Nat Rev Methods Primers, 1(1): 1–21.

Veale, D. J. and Fearon, U. (2018). The pathogenesis of psoriatic arthritis. The Lancet, 391(10136), 2273–2284.

Watanabe, K., Taskesen, E., van Bochoven, A. and Posthuma, D. (2017). Functional mapping and annotation of genetic associations with FUMA. Nat Commun, 8(1): 1826.

Yan, F., Tao, J., Liu, J., Chen, Y. ja Huang, Z. (2025). Cross-tissue transcriptome-wide association study reveals novel psoriasis susceptibility genes. Journal of Translational Autoimmunity, 10, 100286.

Yoshimura, A., Naka, T., and Kubo, M. (2007). SOCS proteins, cytokine signalling and immune regulation. Nature Reviews Immunology, 7(6), 454–465.

Yu, Y., Chen, W., Li, B., … Sun, L. (2024). Cutaneous Calcium/Calmodulin-Dependent Protein Kinase II-γ–Positive Sympathetic Nerves Secreting Norepinephrine Dictate Psoriasis. Advanced Science, 11(23), 2306772.

Zalesak, M., Danisovic, L. and Harsanyi, S. (2024). Psoriasis and psoriatic arthritis—Associated genes, cytokines, and human leukocyte antigens. Medicina, 60(5), 815.

Zeidler, H. K. (2022). Psoriatic spondylitis: A disease manifestation in debate: Evidence to know for the clinical rheumatologist. Journal of Clinical Rheumatology, 28(1), 38–43.

KASUTATUD VEEBIAADRESSID

FinnGen R7 andmebaas: http://r7.finngen.fi (22.03.2024)

FUMA veebiplatvorm: https://fuma.ctglab.nl/ (22.03.2024)

Illumina koduleht: https://www.illumina.com/ (05.06.2024)

Mouse Genome Informatics (MGI), (2025): https://www.informatics.jax.org/ (24.02.2025)

Open Targets Genetics (OTG), (2025): https://genetics.opentargets.org/ (13.01.2025)

PGS Catalog. (2025). Polygenic Score Catalog – Psoriatic Arthritis. Veebiallikas: https://www.pgscatalog.org/search/?q=EFO_0003778 (16.12.2024)

R tarkvara: https://www.r-project.org/ (15.01.2025)

Rahvusvahelise haiguste klassifikatsiooni 10. väljaanne: https://icd.who.int/browse10/2019/en (13.09.2024)

Soome-spetsiifiline referentspaneel SISu: http://www.sisuproject.fi/ (23.09.2024)

UK Biobank. (n.d.). Genotyping and imputation : https://www.ukbiobank.ac.uk (15.10.2024)

LISAD

LISA 1.

Tabel 1. Kokkuvõte varasemate PsA-i ülegenoomse assotsiatsiooniuuringute tulemustest. EA − efektialleel (*effect allele*); AA − alternatiivne alleel (*alternative allele*); OR – šansside suhe (*odds ratio*); 95% CI – šansside suhe 95%-usaldusintervall (*confidence intervall*).

Viide	**Marker (EA/AA)**	**Geen**	**OR (95% CI)**	**p-väärtus**	**Juhud/ kontrollid**
(Liu *et al*., 2008)	rs7993214 (C/T)	*COG6*	1.41	2×10^{-6}	576/ 480
	rs2395029 (C/T)	*HCP5*	4.1	2×10^{-26}	
	rs10484554 (T/C)	*LINC02571, HLA-B*	2.8	2×10^{-39}	
(Soomro *et al*., 2022)	rs1395621 (C/T)	*RUNX3*	1.17	6×10^{-11}	5065/ 21286
	rs12044149 (T/G)	*C1orf141*	1.27	4×10^{-20}	
	rs2111485 (G/A)	*FAP, IFIH1*	1.25	1×10^{-20}	
	rs17622208 (A/G)	*SLC22A5*	1.19	6×10^{-14}	
	rs848 (C/A)	*TH2LCRR, IL13*	1.28	9×10^{-16}	

	rs76956521 (C/A)	*TNIP1*	1.49	3×10^{-16}	
	rs62377586 (A/G)	*LINC01845, IL12B*	1.36	8×10^{-35}	
	rs33980500 (T/C)	*TRAF3IP2-AS1, TRAF3IP2*	1.66	1×10^{-36}	
	rs610604 (T/G)	*TNFAIP3*	1.15	8×10^{-9}	
	rs2020854 (C/T)	*STAT2*	1.43	1×10^{-13}	
	rs3794767 (C/T)	*NOS2*	1.19	5×10^{-13}	
	rs34536443 (C/G)	*TYK2*	1.71	1×10^{-14}	
	rs5754467 (G/A)	*CCDC116, YDJC*	1.19	2×10^{-9}	
	rs17194140 (C/T)	*CNTN4*	1.2	3×10^{-7}	
	rs11665266 (G/A)	*APCDD1, LINC01254*	1.34	2×10^{-6}	
	rs76800961 (A/C)	*LINC02329, RNU6-976P*	1.39	3×10^{-6}	
	rs306281 (G/A)	*PAXIP1*	1.17	4×10^{-6}	

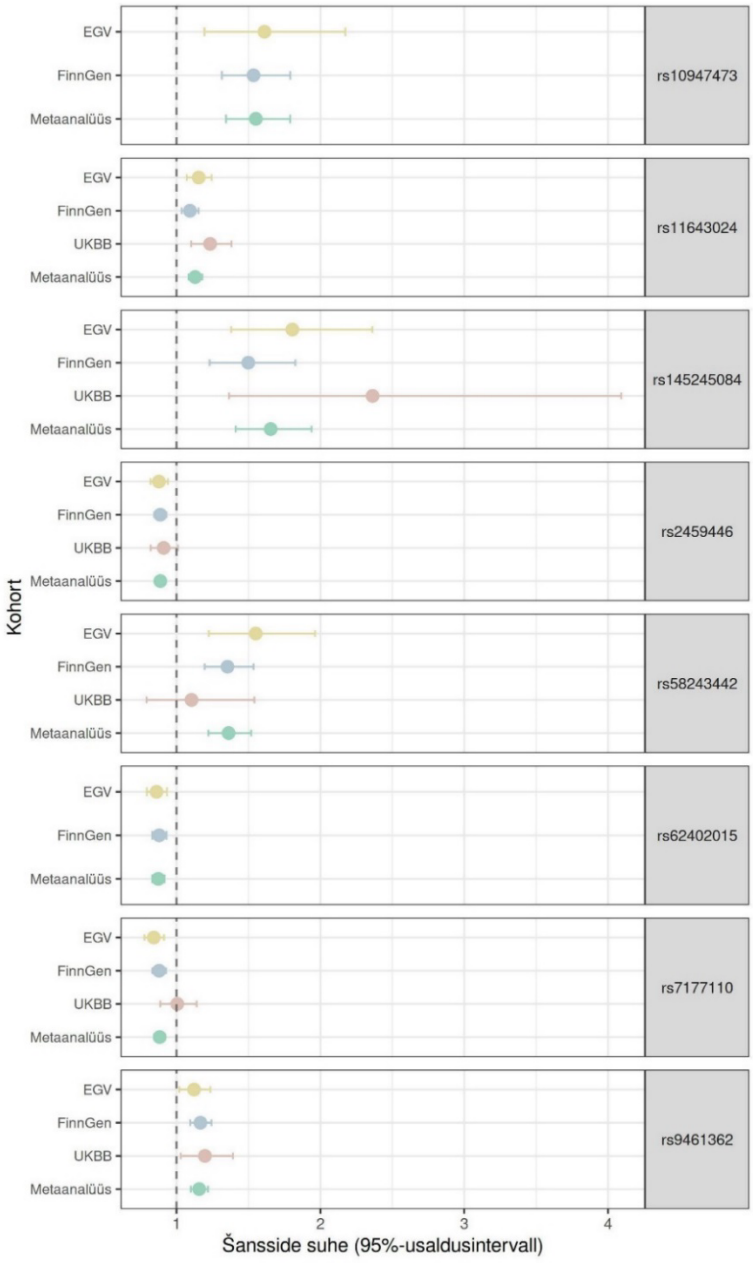
(Aterido *et al*., 2018)	rs11249215 (A/G)	*MIR4425, RUNX3*	-	9 x 10^{-6}	1430/ 1417
	rs4655683 (A/G)	*C1orf141, IL23R*	-	8 x 10^{-7}	
	rs10865331 (G/A)	*RN7SL51P, RN7SL18P*	-	1 x 10^{-8}	
	rs6736203 (A/G)	*GYPC*	-	7 x 10^{-6}	
	rs16858224 (A/G)	*TNS1, RPL7L1P9*	-	3 x 10-6	
	rs74817271 (A/G)	*TNIP1*	-	2 x 10^{-8}	
	rs4921483 (G/A)	*LINC01845, IL12B*	-	9 x 10^{-10}	
	rs115645848 (T/ C või G)	*GABBR1*	-	6 x 10^{-10}	
	rs1610601 (A/C)	*HLA-F-AS1, HLA-F*	-	6 x 10^{-9}	
	rs115529844 (C/ A või T)	*POLR1HASP, HLA-G*	-	5 x 10^{-9}	
	rs115928623 (A/ C või T)	*POLR1HASP, MICD*	-	5 x 10^{-10}	

	rs115618145 (G/C)	*TRIM31-AS1*	-	1 x 10^{-10}	
	rs114265779 (C/ T với A)	*TRIM26, HCG18*	-	4 x 10^{-7}	
	rs116199901 (C/G)	*HCG18*	-	5 x 10^{-7}	
	rs115429849 (A/G)	*MICC*	-	2 x 10^{-17}	
	rs114757170 (A/G)	*MRPS18B, PPP1R10*	-	9 x 10^{-10}	
	rs28780111 (T/C)	*HCG20*	-	7 x 10^{-11}	
	rs149997688 (A/G)	*DDR1*	-	8 x 10^{-18}	
	rs114383947 (T/C)	*MUC22*	-	6 x 10^{-21}	
	rs116805908 (T/C)	*POU5F1, PSORS1C3*	-	3 x 10-31	
	rs115429782 (C/T)	*LINC02571*	-	1 x 10^{-36}	
	rs114598684 (T/C)	*MICA*	-	5 x 10^{-28}	
	rs116340302 (T/A)	*RPL15P4*	-	5 x 10^{-16}	

	rs115872194 (T/C với A)	*BAG6*	-	9×10^{-10}	
	rs115295573 (T/C)	*SNHG32*	-	7×10^{-6}	
	rs114212579 (-)	*SKIC2*	-	4×10^{-14}	
	rs13199524 (T/C)	*TNXB*	-	6×10^{-10}	
	rs115145466 (T/C với A)	*TSBP1-AS1, NOTCH4*	-	2×10^{-11}	
	rs115112280 (A/T với C)	*TSBP1-AS1*	-	5×10^{-10}	
	rs28724231 (G/A)	*HLA-DQA1, HLA-DQB1*	-	4×10^{-11}	
	rs2621322 (G/T)	*HLA-DOB, TAP2*	-	2×10^{-8}	
	rs9378127 (A/G)	*HLA-DMA*	-	4×10^{-7}	
	rs354528 (T/C)	*SLC16A10*	-	1×10^{-7}	
	rs458017 (C/T)	*MFSD4B, REV3L*	-	3×10^{-11}	

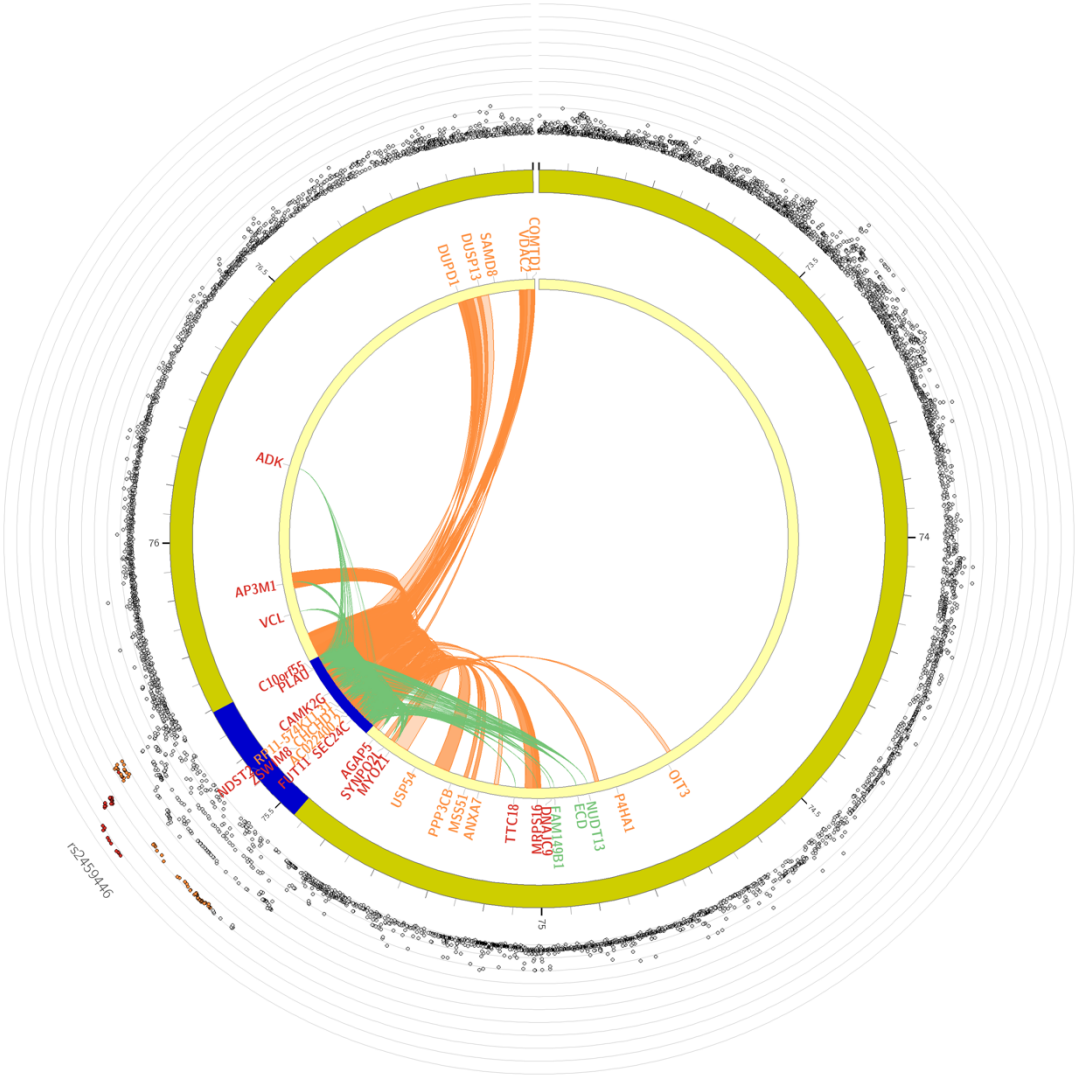
	rs71562294 (G/A)	*TRAF3IP2-AS1*	-	6×10^{-11}	
	rs13190932 (A/G)	*TRAF3IP2-AS1, TRAF3IP2*	-	1×10^{-10}	
	rs10138228 (A/G)	*ATP5MC2P2, RPL18P1*	-	9×10^{-6}	
	rs10459953 (G/C)	*NOS2*	-	2×10^{-6}	

LISA 2

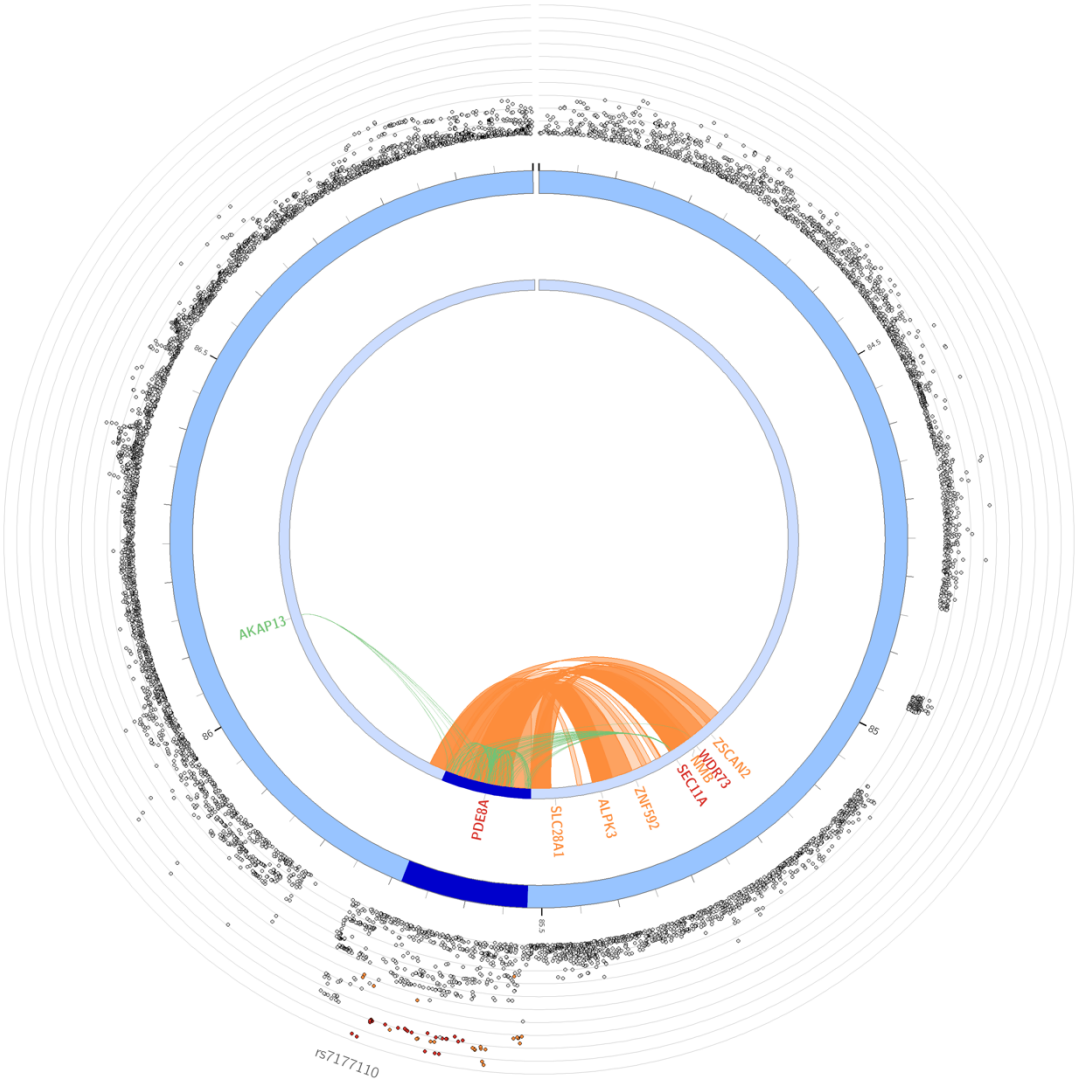


Joonis 4. *Forest*-diagramm kujutab šansside suhet (*odds ratio*) koos 95% usaldusvahemikega kaheksa juht-SNP kohta, mis tuvastati kui potentsiaalselt uued seosed PsA genoomiüleses assotsiatsiooni metaanalüüsis. Tulemused on esitatud iga kohordi kohta eraldi – Eesti geenivaramu (EGV, efektiivne valimi suurus ≈ 6275), FinnGen (Neff ≈ 13320) ja UK Biopank (Neff ≈ 2831) – ning ka kogu metaanalüüsi kohta (Neff ≈ 22427). Mõjuhinnangud põhinevad GWAS-I summaarstatistikal ning efektiivsed valimisuurused arvutati valemiga: 4 / [(1 / N_juhtumid) + (1 / N_kontrollid)]. Visualiseerimine võimaldab võrrelda mõjuhinnanguid ja usaldusvahemikke erinevates kohortides ning kombineeritud analüüsis, aidates hinnata geneetiliste seoste järjepidavust ja tugevust.

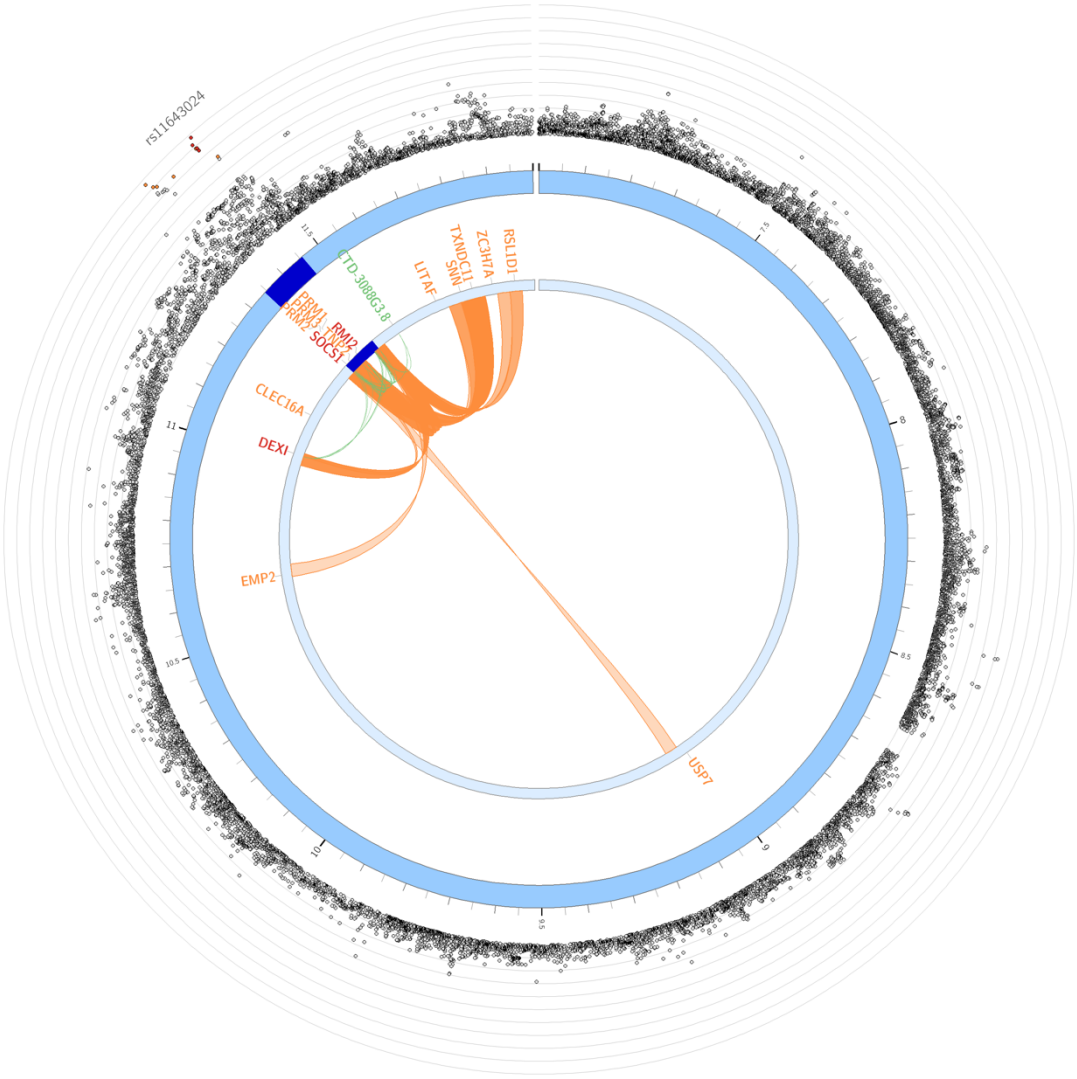
LISA 3.



Joonis 6. PsA-ga seotud 10. kromosoomi geneetilised seosed. GWAS metaanalüüsi põhjal loodud *circos*-diagramm kujutab 10. kromosoomi riskilookusi. Välimisel kihil on Manhattani graafik, kus p-väärtusega ≤ 0,05 SNP-d on tähistatud värvidega vastavalt nende ahelduse tasakaalutusele (r2): punane (>0,8), oranž (>0,6), roheline (>0,4), sinine (>0,2) ja hall (≤0,2). Keskmises kihis on kromosoomi struktuur ja riskilookused (sinine). Sisemises kihis on kandidaatgeenid: rohelised (eQTL), oranžid (kromatiininteraktsioonid) ja punased (mõlemad). Rohelised ja oranžid kaared näitavad vastavalt eQTL ja kromatiininteraktsiooni seoseid. Diagramm loodi FUMA platvormil, mis aitab tuvastada haigusega seotud geneetilisi piirkondi. Mitmete SNP-ide füüsilise lähestikuse paiknemise tõttu ei pruugi kõik rsID-d *circos*-diagrammil nähtavale ilmuda. See on tingitud osaliselt visualiseerimispiirangutest (nähtavuse ja loetavuse tagamiseks) ning osaliselt FUMA platvormi seadistusest, mille kohaselt kuvatakse igas riskilookuses vaid üks juhtvariant.



Joonis 7. PsA-ga seotud 15. kromosoomi geneetilised seosed. GWAS metaanalüüsi põhjal loodud *circos*-diagramm kujutab 15. kromosoomi riskilookusi. Välimisel kihil on Manhattani graafik, kus p-väärtusega ≤ 0,05 SNP-d on tähistatud värvidega vastavalt nende ahelduse tasakaalutusele (r2): punane (>0,8), oranž (>0,6), roheline (>0,4), sinine (>0,2) ja hall (≤0,2). Keskmises kihis on kromosoomi struktuur ja riskilookused (sinine). Sisemises kihis on kandidaatgeenid: rohelised (eQTL), oranžid (kromatiininteraktsioonid) ja punased (mõlemad). Rohelised ja oranžid kaared näitavad vastavalt eQTL ja kromatiininteraktsiooni seoseid. Diagramm loodi FUMA platvormil, mis aitab tuvastada haigusega seotud geneetilisi piirkondi. Mitmete SNP-ide füüsilise lähestikuse paiknemise tõttu ei pruugi kõik rsID-d *circos*-diagrammil nähtavale ilmuda. See on tingitud osaliselt visualiseerimispiirangutest (nähtavuse ja loetavuse tagamiseks) ning osaliselt FUMA platvormi seadistusest, mille kohaselt kuvatakse igas riskilookuses vaid üks juhtvariant.



Joonis 8. PsA-ga seotud 16. kromosoomi geneetilised seosed. GWAS metaanalüüsi põhjal loodud circos-diagramm kujutab 16. kromosoomi riskilookusi. Välimisel kihil on Manhattani graafik, kus p-väärtusega ≤ 0,05 SNP-d on tähistatud värvidega vastavalt nende ahelduse tasakaalutusele (r2): punane (>0,8), oranž (>0,6), roheline (>0,4), sinine (>0,2) ja hall (≤0,2). Keskmises kihis on kromosoomi struktuur ja riskilookused (sinine). Sisemises kihis on kandidaatgeenid: rohelised (eQTL), oranžid (kromatiininteraktsioonid) ja punased (mõlemad). Rohelised ja oranžid kaared näitavad vastavalt eQTL ja kromatiininteraktsiooni seoseid. Diagramm loodi FUMA platvormil, mis aitab tuvastada haigusega seotud geneetilisi piirkondi. Mitmete SNP-ide füüsilise lähestikuse paiknemise tõttu ei pruugi kõik rsID-d *circos*-diagrammil nähtavale ilmuda. See on tingitud osaliselt visualiseerimispiirangutest (nähtavuse ja loetavuse tagamiseks) ning osaliselt FUMA platvormi seadistusest, mille kohaselt kuvatakse igas riskilookuses vaid üks juhtvariant.

LIHTLITSENTS

Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Aleksandra Judin,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose
   " Psoriaatilise artriidi ülegenoomne assotsiatsiooniuuring ",

   mille juhendajad on Ene Reimann (PhD), Tanel Traks (McS) ja Galadriel Lucia Velazquez Silva (McS),

   reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni.

2. Annan Tartu Ülikoolile loa teha punktis 1 nimetatud teos üldsusele kättesaadavaks Tartu Ülikooli veebikeskkonna, sealhulgas digitaalarhiivi kaudu Creative Commonsi litsentsiga CC BY NC ND 4.0, mis lubab autorile viidates teost reprodutseerida, levitada ja üldsusele suunata ning keelab luua tuletatud teost ja kasutada teost ärieesmärgil, kuni autoriõiguse kehtivuse lõppemiseni.

3. Olen teadlik, et punktides 1 ja 2 nimetatud õigused jäävad alles ka autorile.

4. Kinnitan, et lihtlitsentsi andmisega ei riku ma teiste isikute intellektuaalomandi ega isikuandmete kaitse õigusaktidest tulenevaid õigusi.

Aleksandra Judin

29.05.2025