

TARTU ÜLIKOOL
LOODUS- JA TÄPPISTEADUSTE VALDKOND
MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT
GENOOMIKA INSTITUUT

Inimese papilloomiviirusega seotud diagnooside geneetilised riskifaktorid

Bakalaureusetöö

12 EAP

Steffi Salu

Juhendajad
PhD Triin Laisk
PhD Mariann Võsa

TARTU 2025

INFOLEHT

Inimese papilloomiviirusega seotud diagnooside geneetilised riskifaktorid

Kokkuvõte. Inimese papilloomiviirus (HPV) on seotud mitmete haiguste, sealhulgas viiruslike tüügaste, genitaalkondüloomide, kurgu- ja genitaalpiirkonna pahaloomuliste vähivormidega. Lisaks erinevatele riskifaktoritele, nagu HPV, rohke seksuaalpartnerite arv, nõrgenenud immuunsüsteem, tubakatoodete kasutamine ja rohke alkoholi tarvitamine, on nendel haigustel ka geneetilised riskifaktorid, mida on senini vähe uuritud. Käesolevas bakalaureusetöös annoteeriti viiruslike tüügaste, genitaalkondüloomide, ning kurgu- ja genitaalpiirkonna pahaloomuliste vähivormidega seotud geneetilisi variante, mis tuvastati Eesti geenivaramu andmete ülegenoomsel seoseuuringul. Töö tulemused toetavad teooriat, mille kohaselt papilloomiviiruse infektsiooni geneetilised riskifaktorid on seotud eelkõige koosobivuse antigeeni geenidega.

Märksõnad: inimese papilloomiviirus, geneetilised riskifaktorid, GWAS

CERCS kood: B110 Bioinformaatika, meditsiiniinformaatika, biomatemaatika, biomeetrika; B220 Geneetika, tsütogeneetika

The genetic risk factors for human papillomavirus-related diagnoses

Abstract. Human papillomavirus (HPV) is associated with various diseases, including viral warts, genital condylomas, and malignancies in the oropharyngeal and genital regions. While there are known risk factors such as HPV exposure, a high number of sexual partners, weakened immune system, tobacco use, and heavy alcohol consumption, the biological mechanisms and genetic susceptibility factors underlying these diagnoses remain poorly understood. In this bachelor's thesis, I annotated the genetic variants associated with the abovementioned HPV-related diseases in genome-wide association studies conducted on Estonian Biobank data. The results of this work support the idea that genetic susceptibility factors for HPV-related conditions are primarily located within immune system genes, particularly in the human leukocyte antigen region.

Keywords: human papillomavirus, genetic susceptibility factors, GWAS

CERCS code: B110 Bioinformatics, medical informatics, biomathematics, biometrics; B220 Genetics, cytogenetics

SISUKORD

INFOLEHT.....	2
KASUTATUD LÜHENDID.....	4
SISSEJUHATUS.....	6
1. KIRJANDUSE ÜLEVAADE.....	7
1.1. Inimese papilloomiviirus.....	7
1.2. Papilloomiviiruste mõju inimese tervisele.....	7
1.2.1. Viirustüükad ehk soolatüükad.....	9
1.2.2. Genitaalkonduloomid.....	11
1.2.3. Genitaal- ja pärakuvähk.....	14
1.2.4. Peenisevähk.....	17
1.2.5. Suu- ja neeluvähk.....	19
1.3. Inimese papilloomiviirusega seotud diagnooside geneetilised riskifaktorid.....	21
1.3.1. Geneetiliste seoseuringute ülevaade.....	21
1.3.2. Ülevaade inimese papilloomiviiruse põhjustatud diagnooside geneetilistest riskifaktoritest.....	23
2. EKSPERIMENTAALOSA.....	28
2.1. Töö eesmärgid.....	28
2.2. Materjal ja meetodika.....	28
2.2.1. TÜ Eesti geenivaramu valim ja uuritavad tunnused.....	28
2.2.2. Genotüüpiseerimine.....	29
2.2.3. Ülegenoomne seoseanalüüs TÜ Eesti geenivaramu andmetel.....	30
2.2.4. Seoseanalüüsi tõlgendamine.....	30
2.2.5. Funktsionaalne kaardistamine ja geneetiliste variantide prioritseerimine.....	31
2.2.6. Varasemate assotsiatsioonide leidmine ja tulemuste replitseerumine.....	32
2.3. Tulemused.....	33
2.3.1. Teravad genitaalkonduloomid.....	33
2.3.2. Viirustüükad ehk soolatüükad.....	35
2.3.3. Teised analüüsitud vähid.....	37
2.4. Arutelu.....	40
KOKKUVÕTE.....	44
SUMMARY.....	45
TÄNUSÕNAD.....	47
KIRJANDUSE LOETELU.....	48
KASUTATUD VEEBIAADRESSID.....	57
LISAD.....	62
Lisa 1.....	62
LIHTLITSENTS.....	63

KASUTATUD LÜHENDID

A63.0	RHK-10 tähtnumbriline kood, genitaalkondüloomid
B07	RHK-10 tähtnumbriline kood, viirustüükad
CI	usaldusvahemik (ingl <i>confidence interval</i>)
CR	edukalt genotüpiseeritud markerite protsentuaalne osakaal (ingl <i>call rate</i>)
C01	RHK-10 tähtnumbriline kood, keelepära pahaloomuline kasvaja
C09	RHK-10 tähtnumbriline kood, mandli pahaloomuline kasvaja
C10	RHK-10 tähtnumbriline kood, suuneelu pahaloomuline kasvaja
C21	RHK-10 tähtnumbriline kood, päraku ja pärakukanali pahaloomuline kasvaja
C51	RHK-10 tähtnumbriline kood, häbemevähk
C52	RHK-10 tähtnumbriline kood, tupevähk
C60	RHK-10 tähtnumbriline kood, peenisevähk
EAF	efektialleeli sagedus (ingl <i>effect allele frequency</i>)
eQTL	ekspressiooni taset mõjutavad lookused (ingl <i>expression quantitative trait loci</i>)
FUMA	funktsionaalne kaardistamine ja annoteerimine (ingl <i>Functional Mapping and Annotation</i>)
GWAS	ülegenoomne seoseuuring (ingl <i>genome-wide association study</i>)
HLA	inimese leukotsüütide antigeen (ingl <i>human leukocyte antigen</i>)
HPV	inimese papilloomiviirus (ingl. <i>human papillomavirus</i>)
HWE	Hardy-Weinbergi tasakaalustatus (ingl <i>Hardy-Weinberg equilibrium</i>)
kb	tuhat aluspaari (ingl <i>kilobase</i>)
LD	ahelduse tasakaalutus (ingl <i>linkage disequilibrium</i>)
MAF	harvema alleeli sagedus (ingl <i>minor allele frequency</i>)
MAGMA	geenipõhise analüüsi meetod (ingl <i>Multi-marker Analysis of GenoMic</i>

Annotation)

MHC	peamine koesobivuskompleks (ingl <i>major histocompatibility complex</i>)
OR	šansside suhe (ingl <i>odds ratio</i>)
RHK-10	Rahvusvaheline Haiguste Klassifikatsioon, 10. väljaanne (ingl <i>International Classification of Diseases 10th Revision</i>)
SNP	üksiknukleotiidne polümorfism (ingl <i>single nucleotide polymorphism</i>)
TÜ EGV	Tartu Ülikooli Eesti Geenivaramu

SISSEJUHATUS

Inimese papilloomiviirus (HPV) on üks levinumaid viirusnakkusi maailmas, millega puutub elu jooksul kokku suurem osa inimestest (Chesson *et al.*, 2014). Papilloomiviirused on väikesed, kaheaheelised DNA viirused, mis nakatavad limaskestast ja nahaepiteeli ning liigitatakse sõltuvalt nende vähi tekitamise potentsiaalst kõrge ja madala riskiga tüüpideks (International Agency for Research on Cancer, 2007).

Papilloomiviirusinfektsioon on enamasti ajutine, väljendudes viirustüügaste või genitaalkonduloomide näol (Kidd *et al.*, 2017), kuid kõrge riskiga tüvedega kokkupuutel suureneb tõenäosus pahaloomuliste vähivormide nagu emakakaelavähk või genitaal- ning pea- ja kaelapiirkonna vähkide tekkele (American Cancer Society, i.a; American Cancer Society, i.a; Ramqvist ja Dalianis, 2010). Seda, kas nakkus püsib organismis ning kujuneb raskemateks haigusseisunditeks, mõjutavad mitmed tegurid, sealhulgas nakatumise viis ja sagedus, seksuaalkäitumine, immuunsüsteemi seisund, kaasuvad haigused ning elustiilitegurid nagu suitsetamine ja alkoholi tarvitamine (DermNet, i.a; BHC, i.a; Mayo Clinic Staff, 2023; Cancer Research UK, 2023a; Cleveland Clinic, i.a; b). Haigusrisiki võivad mõjutada ka geneetilised riskifaktorid, mis reguleerivad organismi vastust HPV nakkusele.

Varasemad uuringud on näidanud, et HPV-ga seotud haiguste, eriti emakakaelavähi, kujunemisel on oluline roll inimese leukotsüütide antigeeni (HLA) geenide variatsioonidel, mis mõjutavad antigeeni esitlust immuunrakkudele (Shi *et al.*, 2013). HLA piirkonna geenid on osutunud seotuks ka teiste HPV-st tingitud haigustega, nagu pea- ja kaelapiirkonna vähkidega (Dendrou *et al.*, 2018; Yuan *et al.*, 2012), kuid senised uuringud on valdavalt kasutanud nn kandidaatgeenide lähenemist (Amos *et al.*, 2010), mis eeldab eelnevat teadmist vaadeldavate geenide võimalikust rollist haiguse kujunemisel. Ülegenoomsed seoseuuringud (GWAS) võimaldavad hüpoteesivabalt analüüsida geneetilisi variante üle kogu genoomi, et leida statistilisi seoseid haigusriskide ja geneetiliste variantide vahel, tuvastades seeläbi uusi riskilookusi, mida varasemalt pole uuritavate seisunditega seostatud (Cooke *et al.*, 2008).

Käesolevas töös uuriti inimese papilloomiviirusega seotud viiruslike tüügaste, genitaalkonduloomide ning kurgu- ja genitaalpiirkonna vähkide geneetilist tausta Tartu Ülikooli Eesti Geenivaramu andmete põhjal. Eksperimentaalse osa eesmärgiks oli analüüsida nende diagnooside kujunemisega seotud geneetilisi riskifaktoreid, viies läbi ülegenoomse seoseanalüüsi tulemuste annotatsiooni FUMA platvormi abil. Antud bakalaureusetöö on valminud Tartu Ülikooli molekulaar- ja rakubioloogia instituudi geenitehnoloogia õppekava raames. Töö eksperimentaalosa on teostatud TÜ genoomika instituudis.

1. KIRJANDUSE ÜLEVAADE

1.1. Inimese papilloomiviirus

Papilloomiviirused on väikesed, epiteliotroopsed, kaheahelalised DNA viirused, mis nakatavad limaskestast ja nahaepiteeli (International Agency for Research on Cancer, 2007).

Inimese papilloomiviirusel (HPV) on kirjeldatud rohkem kui 200 eri tüüpi, millest mõned on tuvastatud genitaalpiirkonnast (National Cancer Institute, 2023). HPV liigitatakse madala ja kõrge riskiga tüüpideks (vt. allpoolt) sõltuvalt nende võimest põhjustada vähki – kõrge riskiga tüved on seotud pahaloomuliste kasvajate, madala riskiga tüved aga healoomuliste kergete muutustega (International Agency for Research on Cancer, 2007). Papilloomiviirustel on kõigil tüüpidel sarnane geneetiline struktuur ning genoom, mis sisaldab umbes kaheksat avatud lugemisraami (International Agency for Research on Cancer, 2007). HPV klassifitseerimine põhineb L1 geeni nukleotiidjärjestuse homoloogial, mis on kohati ka viiruse genoomi kõige konserveeritum piirkond. Taksonoomiline jaotus põhineb nukleotiidide sarnasusel eri tasemetel, kus 60% ulatuses nukleotiidide erinevusel klassifitseeritakse HPV-d eri perekonda. Perekonna sees jagavad erinevad liigid omavahel 60–70% nukleotiidide sarnasust, mis näitab, et kuigi nad kuuluvad samasse perekonda, on neil siiski olulised geneetilised erinevused (Mühr *et al.*, 2021). Kõrge riskiga inimese papilloomiviiruste alla kuuluvad tüübid 16, 18, 31, 33, 35, 39, 45, 51, 52, 56, 58 ja 59, mis võivad põhjustada erinevaid vähivorme sh. häbeme-, tupe-, peenise-, suu- ja neeluvähki. Enim seostatakse erinevate papilloomiviirusest tingitud vähkidega HPV tüvesid 16 ja 18 (National Cancer Institute, 2023).

Madala riskiga inimese papilloomiviirused põhjustavad harva vähki, kuid mõned neist, nagu tüübid 6 ja 11, võivad tekitada tüükaid erinevates keha piirkondades (National Cancer Institute, 2023).

1.2. Papilloomiviiruste mõju inimese tervisele

HP-viirusinfektsioon on enamasti ajutine ning mõjutab lameepiteeli kas viirusnakkusena või vähieelse kahjustusena. Kui lameepiteel võimaldab viirusosakeste tootmist, areneb infektsioon madala astme morfoloogilisteks kahjustusteks, nagu kondüloomid ja kerge düsplaasia (Kidd *et al.*, 2017). Vähieelsete kahjustuste tagajärjel integreerub viiruse genoom peremeesrakku, mistõttu viiruse onkogeenid ekspresseeruvad, stimuleerivad rakkude paljunemist ning hiljem diferentseerumata rakud koloniseeruvad, suurendades pahaloomulise

transformatsiooni riski (Spiess *et al.*, 2016). Krooniliseks muutunud HPV, eriti kõrge riskiga tüvede puhul, võib areneda erinevateks vähivormideks (American Cancer Society, i.a). Eestis diagnoositakse üle 300 HPV-st põhjustatud vähijahu aastas. Emakakaelavähk on senini olnud enim diagnoositud, kuid suu-, neelu- ja pärakuvähi diagnoosimine on Eestis alates 2013. aastast hakanud suurenema (Terviseamet, 2024).

Elu jooksul puutub enamik inimesi vähemalt ühe inimese papilloomiviiruse tüvega kokku (Chesson *et al.*, 2014). Nakatumise tõenäosus sõltub mitmetest teguritest, sealhulgas vanusest (Yang *et al.*, 2022), immuunsüsteemi seisundist (Mayo Clinic Staff, 2023), naha või limaskestast mikrotraumadest (BHC, i.a), viiruse tüübist, erinevate inimestega kokkupuutest ja seksuaalpartnerite arvust (DermNet, i.a; Mayo Clinic Staff, 2023).

Inimese papilloomiviiruse vastu on välja töötatud vaktsiinid, mis pakuvad tõhusat kaitset viirustüvede vastu, mis on kõige sagedamini seotud pahaloormuliste kasvaja tekkega. Kuigi on saadaval erinevaid HPV vaktsiine, sisaldavad kõik neist kaitset HPV tüüpide 16 ja 18 vastu, kuna need kaks tüve põhjustavad valdava osa papilloomiviirusega seotud vähkidest. Vaktsiin Gardasil-9 kaitseb lisaks ka HPV tüvede 6, 11, 31, 33, 45, 52 ja 58 vastu (CDC, 2024), aidates ennetada emakakaela-, häbeme-, tupe- ja päraku vähieelseid seisundeid ning vähki, samuti genitaalkonduloomide teket (vaktsineeri.ee, i.a).

Papilloomiviiruse infektsiooniga seotud haiguste diagnoosimiseks kasutatakse rahvusvaheliselt tunnustatud klassifikatsioonisüsteemi. Üks selline süsteem on rahvusvaheline haiguste klassifikatsioon ehk RHK, mis on standardiseeritud vahend erinevate haigusseisundite, sealhulgas viirusnakkuste, kirjeldamiseks ja kodeerimiseks. Samuti aitab see hinnata tervise seisundite trende nii ülemaailmsel kui ka riiklikul tasandil. Iga RHK jaotis määrab haigustele, häiretele, vigastustele ja seisunditele unikaalse koodi, mis koosneb ühest tähest ja kahest numbrist (TAI, i.a).

Inimese papilloomiviiruse roll emakakaelavähi kujunemisel on üldteada, kuid HPV olemasolu iseenesest ei tähenda veel vähi teket. HPV on laialt levinud viirus, samas kui emakakaelavähk esineb oluliselt harvem, ehkki kuni 99% juhtudest on selle haiguse põhjuseks just HPV-infektsioon (Mount Sinai Tisch Cancer Center, i.a). Vähem teatakse HPV rolli viirustüügaste, genitaalkonduloomide, kurgu- ja genitaalpiirkonna muude vähkide tekkel, mistõttu käsitletakse neid käesolevas bakalaureusetöös põhjalikumalt. Käesolevas bakalaureusetöös põhjalikumalt käsitletud inimese papilloomiviirusega seotud diagnoosid ning seotud tüvi ja riskitase on näidatud kokkuvõtlikult Tabelis 1.

Tabel 1. Papilloomiviirusest tingitud haigusseisundid ja nendega seotud HPV tüved.

Haigusseisund	Seotud HPV tüvi	Riskitase (madal/ kõrge)	Viide allikale
Viirustüükad	2, 4, 7, 27, 57	madal	(Cubie, 2013; Redzic <i>et al.</i> , 2023)
Genitaalkondüloomid	6, 11, 42 ja 81	madal	(Clifford <i>et al.</i> , 2005; Petry <i>et al.</i> , 2012)
Häbeme- ja tupevähk	16, 18 ja 33	kõrge	(MOFFITT Cancer Center, i.a; a)
Genitaal- ja pärakuvähk	16 ja 18	kõrge	(Daling <i>et al.</i> , 2004)
Peenisevähk	16 ja 18	kõrge	(Cancer Research UK, 2023a)
Suu- ja neeluvähk	16	kõrge	(Cleveland Clinic, i.a; b)

1.2.1. Viirustüükad ehk soolatüükad

Viirustüükad ehk soolatüükad on HPV infektsiooni enimlevinud tagajärg ning neid liigitatakse asukoha järgi naha- ja limaskestade tüügasteks (DermNet, i.a). RHK-10 ehk rahvusvahelise haiguste klassifikatsiooni 10. versioonis on need kirjeldatud B07 koodiga.

Viirustüükad on oma olemuselt väikesed kõrgenenud tüükad, mille suurus jääb tavaliselt vahemikku 1–10 mm. Tüükad võivad olla nii kareda kui ka sileda pinnaga ning esineda üksikult või kobaratena (BHC, i.a). Kätel ja näol esinevad viirustüükad on valdavalt nahapinnast kõrgemad, mitmekordsed ja ebaregulaarse kujuga karedad sõlmed (Cubie, 2013).

Suurtes populatsioonipõhistes uuringutes on mittesugulisel teel levivate viirustüügaste levimus olnud väga erinev, ulatudes 0,84% Ameerika Ühendriikides kuni 12,9% Venemaal (Loo ja Tang, 2014). Viiruslikud tüükad on sagedased eeskätt lastel ja noorukitel, esinedes algkooliealistel lastel hinnanguliselt 22–33% ulatuses (Kilkenny *et al.*, 1998; Haalen *et al.*, 2009).

Viirus levib peamiselt nahk-nahk kontaktil või autoinokulatsiooni teel, nakatades epidermise basaalkihi rakke ja soodustades keratinotsüütide proliferatsiooni ning hüperkeratoosi, mille tulemusel tekivad nakkuslikud viirusosakesed ja tüügaste moodustumine (DermNet, i.a).

Viirustüügaste tekkeriski suurendavad naha pindmise kihi traumad ja infektsioonid, mis loovad viirusele soodsa sisenemistee ning soodustavad viiruse levikut. Samuti suurendab riski sage kokkupuude veega ning käte või jalgade liigne higistamine (hüperhidroos), mis pehmen­dab nahka ja muudab selle vastuvõtlikumaks nakkusele. Nakatumise tõenäosust suurendab ka avalikes ujumisbasseinides viibimine, kus viirus võib levida niiskete pindade kaudu. Küünte närimise harjumus ja otsene kontakt teiste inimeste tüügastega võivad samuti viiruse levikut soodustada. Lisaks võib sügelusest tingitud tüügaste kratsimine või raseerimine viia nakkuse levikuni teistesse kehapiirkondadesse (BHC, i.a). Viirustüügaste kujunemises võivad olulised olla ka geneetilised riskifaktorid, millest on pikemalt juttu peatükis 1.3.2.

Kõige levinumad nahka nakatavad HPV tüved on 1, 2, 3, 4, 10, 27, 29 ja 57 (DermNet, i.a; Redzic *et al.*, 2023). Lisaks näidati Redzic *et al.* (2023) uuringus, et 2% proovidest olid HPV negatiivsed, samas kui HPV 27 esines 73% juhtudest. HPV tüüp 7 antud uuringus ei tuvastatud, kuna selle esinemissagedus on väga väike (0,5%) (Redzic *et al.*, 2023), kuid on näidatud, et see tüvi on seotud suuremate ja püsivamate viirustüügastega (Cubie, 2013). Kätel ja näol esinevad viirustüükad on enamasti põhjustatud HPV tüübist 2 (vt. Joonis 1) (Williams *et al.*, 1993) ja sellega lähedastest tüüpidest 27 ja 57 (Rübben *et al.*, 1997). Peopesadel esinevaid väiksemaid, endofüütilisi, täpilisi kahjustusi on kõige sagedamini seostatud HPV 4-ga (Williams *et al.*, 1993). HPV 7 (vt. Joonis 2) on enimlevinud isikutel, kelle nahk on pidevas kokkupuutes vee ja kül­maga, nagu kalakaupmehed ja lihakäitlejad (Cubie, 2013).

Viirustüükad suuremat ohtu tervisele ei kujuta (DermNet, i.a), kuid need võivad põhjustada sügelusest tingitud ebamugavustunnet (Dudley, 1995) ning psühholoogilist mõju sh. narrimise kogemist (DermNet, i.a).

Tõhusaimaks viirustüügaste raviks peetakse vedela lämmastikuga tüügaste külmutamist (krüoteraapia). Isikud, kellel on urtikariaalne reaktsioon külmale (ehk külmaallergia) või Raynaud' sündroom, kus sõrmed ja varbad temperatuuri languse tõttu muutuvad tuimaks ning muudavad värvust, on krüoteraapia vältimiseks apteekide käsimüügis saadaval ka mitmeid salitsüülhapet sisaldavaid lahuseid ja plaastreid. Vedela lämmastikuga tüügaste töötlemise käigus võib tekkida kerge ebamugavustunne, valulikkus ning villide moodustumine ravitud piirkonnas. Täielikuks eemaldamiseks võib osutada vajalikuks korduv külmutamine iga 1–3 nädala järel, tavaliselt 3–6 korda, kuid mõnel juhul ka rohkem (Confido, 2022).



Joonis 1. HPV 2 tüübiga seostatud viirustüükad kätel (Cubie, 2013, kohandatud). (a) Viirustüükad käeseljal. (b) Küünetüükad. (c) Immuunsupressiooniga patsiendi käte tüükad.



Joonis 2. Suuremad ja püsivad viirustüükad (Cubie, 2013). (a) HPV 7-ga tuvastatud tüükad kalakaupmehel. (b) Tüükad lihakaitlejal.

1.2.2. Genitaalkondüloomid

Genitaaltüükad on üks levinumaid sugulisel teel levivaid nakkusi, mille esinemissagedus on viimastel aastatel märkimisväärselt kasvanud (Woodhall *et al.*, 2008). RHK-10 klassifikatsioonis on need kirjeldatud A63.0 koodiga. Kahjustused paiknevad suguelundite piirkonnas, sh. peenisel, skrootumil, vulva ja tupepiirkonnas ja päraku sisepinnal (vt. Joonis 3) (Leslie *et al.*, 2023).

Suguelundite tüükad võivad esineda üksikute väikeste papulaarsete või suuremate kahjustustena niisketil pindadel. Kuivadel nahapiirkondadel nagu häbememokkadel, võivad

need meenutada nahatüükaid. Kondüloomide värvus varieerub valgest, roosast punaseni ning pruunist pigmenteerunud toonideni (Cubie, 2013).

Nakkus levib peamiselt seksuaalvahekorral, genitaalide nahk-nahk kokkupuutel ning oraalseksi ajal. Edasikandumine võib toimuda ka asümptomaatilisel kandjalt ehk inimeselt, kellel tüükaid pole, kuid kes on vastava HPV tüve kandja (Cleveland Clinic, i.a; a).

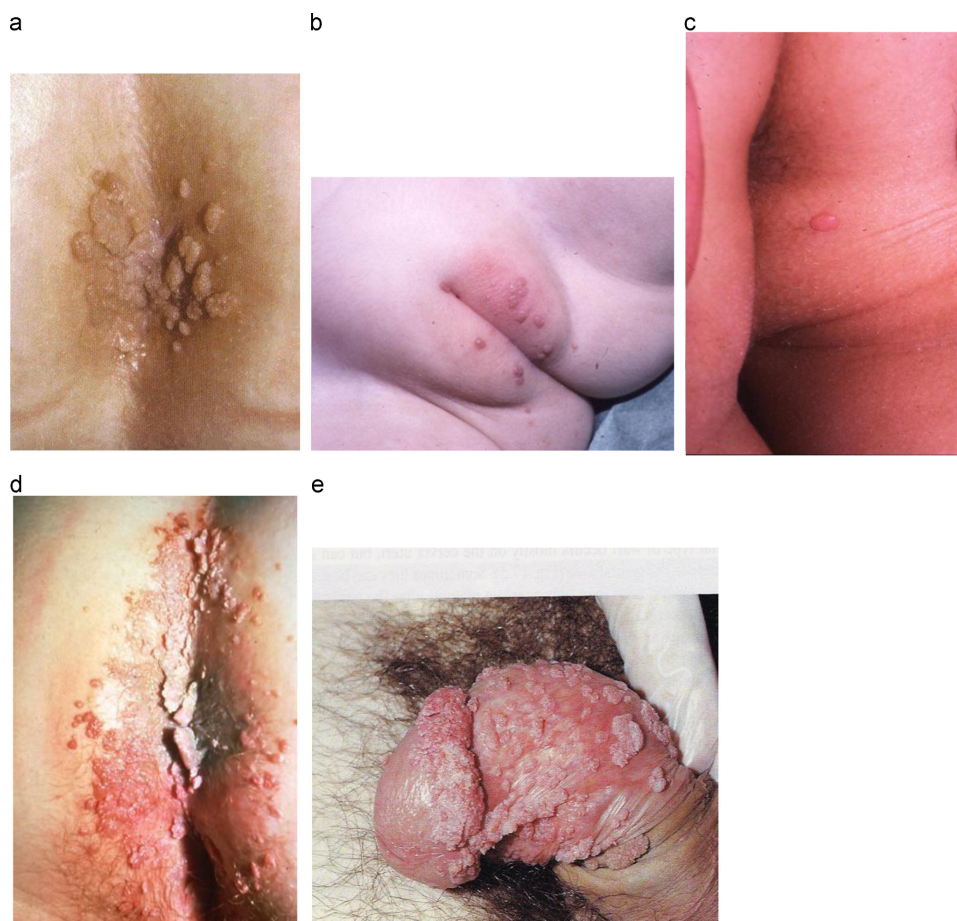
Lisaks füüsilistele sümptomitele võivad genitaalkondüloomid põhjustada märkimisväärset psühhosotsiaalset stressi ning mõjutada üldist heaolu, sealhulgas tekitada süütunnet, ärevust, madalat enesehinnangut kui ka ravi edasilükkamist või vältimist (Woodhall *et al.*, 2008). Sageli kaasnevad ka sügelus, veritsus, lõhenemine ning valulikkus seksuaalvahekorra ajal (Cubie, 2013).

Kondüloomide tekkega seostatakse enamasti HPV tüüp 6, 11, 42 ja 81 (Clifford *et al.*, 2005; Petry *et al.*, 2012). Genitaalpiirkonnas on need põhjustatud madala riskiga HPV tüvedest, millele pole ravi ning tüügaste eemaldamine ilmtingimata ei vähenda HPV edasikandumist, kuid 80% nakatunutel taandub viirus spontaanselt 18–24 kuu jooksul (Leslie *et al.*, 2023). HPV-6 ja HPV-11 tüvedega kokkupuutel tekivad genitaaltüükad 10%-il nakatunutest (Leslie *et al.*, 2023).

Viimase kahe aastakümnega on genitaalkondüloomide esinemissagedus suurenenud peaaegu neljakordselt. Ameerika Ühendriikides diagnoositakse neid igal aastal ligikaudu 1% elanikkonnast. Andmed Inglismaalt, Panamast, Itaaliast, Hollandist ja mitmest teisest riigist viitavad sellele, et genitaaltüükad on seal sama laialt levinud kui Ameerika Ühendriikides (Kalczynski, 2024). Hinnanguliselt levib nakkus partnerite vahel umbes 60% tõenäosusega (Woodhall *et al.*, 2008). Esinemissagedus ja levimus on kõige kõrgemad hilisteismelistel ja 20ndate alguses, mida mõjutavad nii käitumuslikud kui ka bioloogilised tegurid. Levimus langeb vananedes, kuid 50.–55. eluaastates täheldatakse paljudes riikides teist tõusulainet, mis võib olla tingitud varjatud viiruse taasaktiveerumisest või uue seksuaalpartneri kaudu saadud nakkusest (Cubie, 2013). Uusküla *et al.* (2015) uuringu alusel 2011–2013 aastatel genitaalkondüloomide levimus oli Eestis naiste seas 2,9% ning meestel 2,8% (Uusküla *et al.*, 2015).

Ameerika Ühendriikide Toidu- ja Raviamet (ingl *United States Food and Drug Administration*) on heaks kiitnud Condylori (0,5% podofüloksi lahuse) genitaalkondüloomide paikseks raviks. Lahust kantakse kahjustatud alale ning seda ei tohi maha pesta. Mõned arstid määravad ka Podocon-25 või Podofini (20% podofüllüüni sisaldava antimitootilise lahuse),

mida tuleb peale kanda ja hiljem maha pesta, või 5% 5-fluorouratsiili kreemi. Enamik Condyluxi kasutajaid kogeb ravipiirkonnas põletustunnet, valu, põletikku, sügelust või erosiooni. Rasedatele need ravivõimalused ei sobi, kuna võivad põhjustada sünnidefekte (HealthCentral, 2019). Euroopa Komisjon on heaks kiitnud otse tüükale kantava imikvimoodi sisaldava Aldara kreemi, mis stimuleerib immuunvastust (European Commission, i.a). Arst võib soovitada ka trikloroäädikhapet (TCA) kahjustatud aladele. Nähtavaid kondüloome saab eemaldada füüsiliste meetoditega, nagu krüoteraapia (külmutamine), termoteraapia (kuumutamine), skalpelliga lõikamine või laseriga eemaldamine, kuid nende meetoditega võivad kondüloomid uuesti tekkida, kuna HPV jääb ümbritsevatesse rakkudesse alles (HealthCentral, 2019). Genitaalkondüloome tekitavate HPV tüvede vastu aitab vaktsineerimine. Lisaks aitab nakkust ennetada kondoomide kasutamine, suguhaiguste rutiinsed testid ja väike seksuaalpartnerite arv (Cleveland Clinic, i.a; a).



Joonis 3. Genitaalkondüloomid (Cubie, 2013). (a) Genitaaltüükad täiskasvanul. b) Genitaaltüükad lapse häbememokkadel. c) Üksik genitaaltüügas lapse tuharal. d) Ulatuslikud perianaalsed tüükad neeru allotransplantaadi retsiipiendil. e) Genitaaltüükad peenisel ja eesnahal.

1.2.3. Genitaal- ja päarakuvähk

Madala riskiga inimese papilloomiviiruse tüved tekitavad nii genitaal- kui päraku tüükaid, mis hiljem suurendavad tõenäosust nakatuda kõrge riskiga HPV tüvedesse, mis võivad tekitada genitaal- ja päarakuvähki (American Cancer Society, i.a). Rahvusvahelise haiguste klassifikatsiooni 10. versiooni järgi on häbeme- ja tupevähk koodidega vastavalt C51 ja C52 ning päarakuvähk koodiga C21.

Tupest alguse saanud vähid on väga haruldased. Enamasti algab rakkude peatamatu kasv mujal ning levib edasi tuppe. Vähk võib edasi levida kopsudesse, maksa ja luudesse, misjuhul nimetatakse seda metastaatiliseks vähiks (Mayo Clinic Staff, 2025b). Päarakuvähk tekib päarakukanali rakkude kontrollimatu kasvu tagajärjel. Üldjuhul vähk ei levi, kuid kui see jõuab kopsudesse või maksa, muutub selle ravi oluliselt keerulisemaks (Mayo Clinic Staff, 2023).

Vaginaalne lamerakk-kartsinoom on enimlevinud tupevähi tüüp, mis saab alguse tupe lameepiteelist (Mayo Clinic Staff, 2025b). Häbemevähk avaldub tavaliselt häbemepiirkonnas mügarike või haavandite (vt. Joonis 4) kujul, mis võivad põhjustada sügelust. Kuigi see võib esineda igas vanuses, diagnoositakse haigus kõige sagedamini vanematel naistel. Lisaks võivad esineda nahamuutused, nagu häbeme naha värvuse muutus või paksenemine, samuti suguelundite piirkonnas verejooks, mis ei ole seotud menstruatsiooniga (Mayo Clinic Staff, 2025a).

Tupevähil (vt. Joonis 4) ei pruugi algselt sümptomeid esineda, kuid vähi kasvamisel võib tekkida verejooks ning tupevoolus. Samuti tihe valulik urineerimine, kõhukinnisus ja vaagnapiirkonna valud. Kindlaks vähi olemasolu sümptomiks võib pidada tüki või massi leidmist tupes (Mayo Clinic Staff, 2025b).

Päarakvähi sümptomid võivad sarnaneda hemorroididega, kuna mõlemal juhul võib esineda rektaalset verejooksu, verd väljaheites ja valu päraku piirkonnas. Lisaks võivad päarakuvähile viidata ka päarakukanali suurenemine, sügelus pärakupiirkonnas ning sagedasem sooletegevus (Mayo Clinic Staff, 2023).

Häbeme- ja tupevähi riskifaktoriteks peetakse eakamat vanust, kuna keskmine diagnoosimise vanus on 65 aastat, samuti suitsetamist ja kokkupuudet HP-viirusega, mis võib muuta rakud kartsinogeenseteks. Häbemevähi riski võivad suurendada ka teatud haigused, näiteks nahahaigus sklerootiline lihhen (ingl *lichen sclerosis*), nõrgem immuunsüsteem ja eelnevad häbeme vähieelsete seisundite esinemised. Tupevähi riski võib suurendada raseduse katkemise ennetamiseks kasutatavate ravimite nagu dietüülstilbestrool (DES) tarbimine

(Mayo Clinic Staff, 2025a; 2025b). Pärakuvähi riskifaktoriteks on kokkupuude inimese papilloomiviirusega, suur hulk seksuaalpartnereid ja anaalvahekord. Lisaks võivad riski suurendada ka suitsetamine, nõrgenenud immuunsüsteem ning varasem emakakaela-, häbeme- või tupevähi diagnoos (Mayo Clinic Staff, 2023). Genitaalvähkide kujunemisel võivad olulised olla ka geneetilised riskifaktorid, millest on pikemalt juttu peatükis 1.3.2.

Häbeme- ja tupevähi enimlevinud tüüp on lamerakk-kartsinoom (ingl *squamous cell carcinoma*), mis on pooltel juhtudel seotud inimese papilloomiviiruse kõrge riskiga tüüpidega 16, 33 ja 18 ning teine pool juhtudest on põhjustatud p53 supressorgeeni mutatsioonidest (MOFFITT Cancer Center, i.a; a). HPV tüvedest on pärakuvähi kõige sagedasemaks põhjustajaks on inimese papilloomiviiruse tüüp 16 (Daling *et al.*, 2004).

Genitaalvähid, v.a emakakaelavähk, on küllaltki haruldased. Ameerika Ühendriikides on vaid mõnes piirkonnas häbemevähi esinemissagedus naistel või pärakuvähi esinemissagedus meestel suurem kui 2 juhtu 100 000 inimese kohta. Vaginaalvähk on veelgi haruldasem, esinedes kogu maailmas vähem kui 1 juhtu 100 000 inimese kohta (Curado *et al.*, 2008). Ülemaailmse Vähivaatluskeskuse (ingl *Global Cancer Observatory*) ja Rahvusvahelise Vähiuuringute Agentuuri (ingl *International Agency for Research on Cancer*) kohaselt oli Euroopa 2022. aastal vaginaalvähi esinemissageduselt viiendal kohal. Sel perioodil diagnoositi nimetatud vähivorm ligikaudu kolmel tuhandel naisel (WorldGODay; i.a).

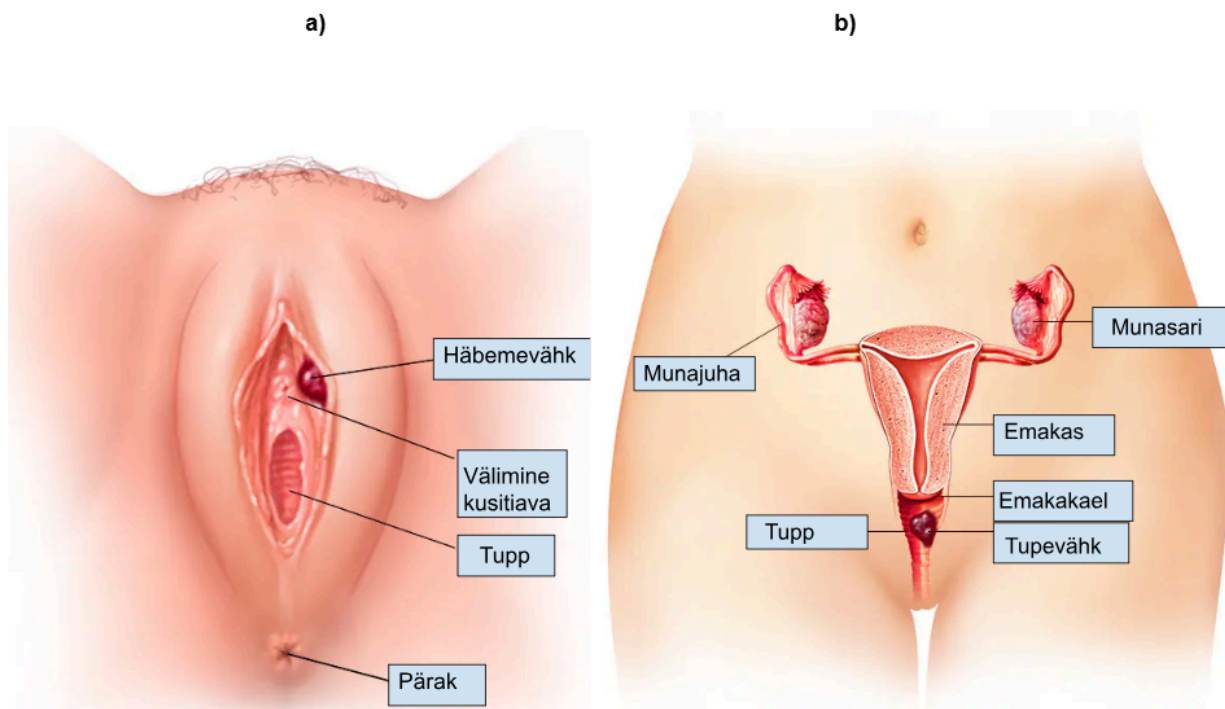
Häbemevähi ravi algab tavaliselt kirurgilise sekkumisega, mille käigus eemaldatakse kasvaja koos väikese osa ümbritseva terve koega. Raskematel juhtudel võib osutuda vajalikuks kogu häbeme eemaldamine. Mida varem haigus diagnoositakse, seda väiksem on vajadus ulatusliku operatsiooni järele (Mayo Clinic Staff, 2025a).

Tavaliselt algab tupevähi ravi samaaegse kiiritus- ja keemiaraviga. Väiksemate kasvajate korral, kus vähk pole tupest edasi levinud, võib esmaseks ravimeetodiks olla vaginektoomia, kus operatsiooni teel eemaldatakse kas osa tupest või terve tupp. Samuti võidakse teostada vaagna eksenteratsioon, kus opereeritakse paljud vaagnaelundid sh. põis, munasarjad, emakas, tupp ja pärasool. Seda võidakse kasutada, kui vähk taastub või ei allu muudele ravimeetoditele. Kõhuõõnde luuakse avad, mis võimaldavad uriinil ja jääkainetel kehast väljuda. Ravimeetoditele mitteallumise korral võib kasutada ka immunoteraapiat (Mayo Clinic Staff, 2025b).

Varasemalt raviti enamikku pärakuvähi juhtumeid operatsiooniga, mille käigus eemaldati kasvaja ning sageli loodi alternatiivne viis roojamassi väljumiseks kehast. Tänapäeval on

peamiseks ravimeetodiks keemia- ja kiiritusravi kombinatsioon või ka immunoteraapia, mis võimaldab paljudel juhtudel operatsiooni vältida. Väga väikesed kasvajaad võidakse ka opereerida, kuid see on üsna harv juhus ning rakendatakse vaid siis, kui operatsioon ei kahjusta anaalkanalit ümbritsevaid päraku sulgurlihaseid. Kui vähk ei allu keemia- ja kiiritusravile, võib arst soovitada abdominoperineaalset resektsiooni (ingl *abdominoperineal resection*), kus operatsiooni käigus eemaldatakse päkakukanal, päkasool ja osa käärsoolest. Seejärel luuakse kõhule avaus, mille kaudu jäätmed väljuvad ning kogunevad kolostoomikotti. Sõeluuringud ja biopsia võivad aidata tuvastada nii päkakuvähki kui ka vähieelseid rakumuutusi, mis võivad aja jookul areneda vähiks (Mayo Clinic Staff, 2023).

Inimese papilloomiviirusest tingitud häbeme-, tupe- ja päkakuvähi ennetamisel on oluline roll HPV-vastasel vaksineerimisel, mis aitab vähendada nakatumise ja sellest tuleneva vähitekke riski. Kuigi kondoomi kasutamine ei taga täielikku kaitset, võib see samuti vähendada viiruse ülekande tõenäosust. Päkakuvähi riski aitab vähendada ka seksuaalpartnerite arvu piiramine. Lisaks loetakse oluliseks suitsetamisest hoidumist, kuna suitsetamine on seotud kõrgema haigestumise riskiga. Tupevähki ei ole võimalik kindlalt ennetada, kuid regulaarne günekoloogiline kontroll ja HPV-vastane vaksineerimine võivad aidata kaasa selle varajasele avastamisele ja leviku tõkestamisele (Mayo Clinic Staff, 2023; 2025a; 2025b).



Joonis 4. a) Häbemevähk (Mayo Clinic Staff, 2025a, kohandatud), b) tupevähk (Mayo Clinic Staff, 2025b, kohandatud).

1.2.4. Peenisevähk

Peenisevähk (vt. Joonis 5) võib areneda kõikjal peenisel, kuid esineb kõige sagedamini ümberlõikamata meestel, lokaliseerudes eesnaha alla, peenise peale või otsa (Cancer Research UK, 2023b). Rahvusvahelise haiguste klassifikatsiooni 10. versiooni järgi on peenisevähi kood C60.

Peenisevähi sümptomid on üsna ebamäärasead, mistõttu võib neid segamini ajada mõne muu haigusseisundi või sugulisel teel leviva haigusega. Peenisevähile võivad viidata haavandid või mügarikud peenise piirkonnas, mis võivad näida tüügastena ning ei parane nelja nädala jooksul. Samuti verejooks, ebameeldiva lõhnaga eritis, lööve, peenise või eesnaha värvimuutus ning raskus eesnaha tagasi tõmbamisel. Hilisema staadiumi peenisevähi sümptomiteks peetakse ka tükki kubemes, väsimustunnet, valu kõhus ja luudes ning kaalulangust (Cancer Research UK, 2023c). Peenisevähi olemasolul võib muutuda urineerimine ja seksuaalvahekord seletamatult valulikuks. Patsiendid võivad tunda piinlikkust ning hirmu, mistõttu välditakse või lükatakse ravi saamist edasi (MOFFITT Cancer Center, i.a; b).

Peenisevähi tekkepõhjused ei ole alati üheselt määratletavad, kuna haigus võib esineda ka ilma teadaolevate riskifaktoriteta. Uuringud on siiski näidanud, et ümberlõikamata meestel on suurem tõenäosus nakatuda inimese papilloomiviirusesse, mille krooniliseks muutumine võib suurendada peenisevähi tekkeriski. Kuigi otsest seost ümberlõikamata oleku ja kõrgema nakatumisriskiga ei ole tõestatud, on seda seostatud smegma kogunemisega eesnaha alla, mis võib põhjustada kroonilist ärritust ja põletikku, soodustades pahaloomuliste muutuste teket. Lisaks võib suitsetamine suurendada peenisevähi riski, kuna tubakasuitsus leiduvad kantserogeensed kemikaalid võivad vereringe kaudu kahjustada peeniserakkude geneetilist materjali (American Cancer Society, 2018). Samuti on peenisevähi tekkerisk kõrgem nõrgenenud immuunsüsteemiga inimestel, näiteks HIV-positiivsetel patsientidel või immuunsupressiivset ravi saavatel isikutel (Cancer Research UK, 2023a). Peenisevähi kujunemisel võivad olulised olla ka geneetilised riskifaktorid, millest on pikemalt juttu peatükis 1.3.2.

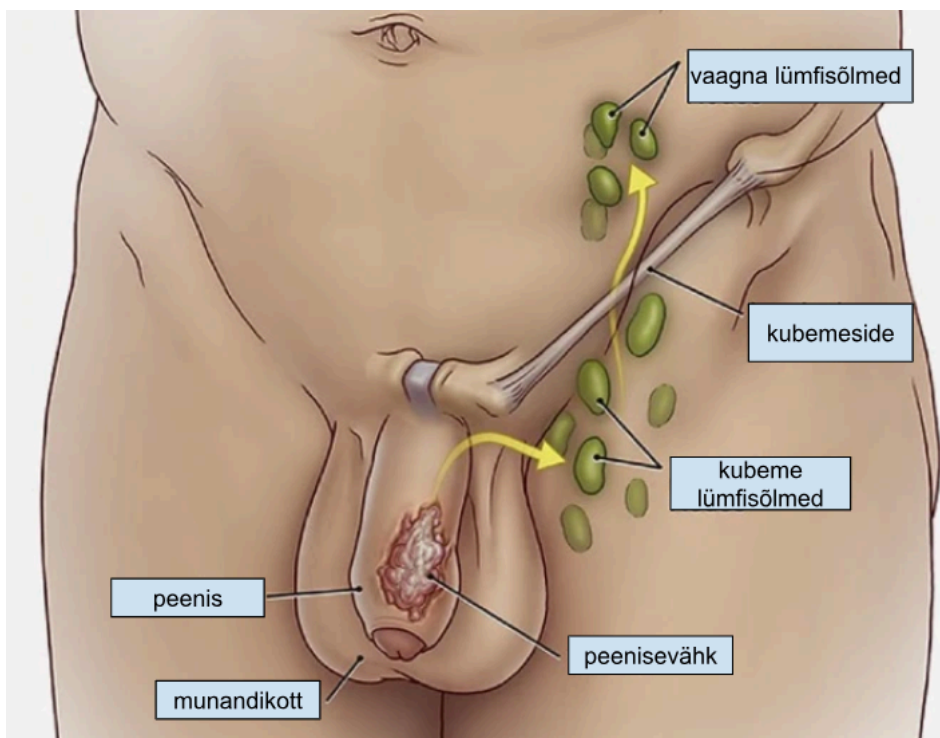
Umbes 95% peenisevähkidest on lamerakk-kartsinoomid (MOFFITT Cancer Center, i.a; b). Ligikaudu 60% peenisevähi juhtudest on seotud inimese papilloomiviiruse nakkusega, kusjuures kõige levinumad on HPV tüved 16 ja 18 (Cancer Research UK, 2023a).

Peenisevähk on haruldane pahaloomuline kasvaja, mida esineb peamiselt eakatel meestel. Haiguse diagnoosimisel on patsientide keskmine vanus 60 aastat ning esinemissagedus suureneb vanusega, saavutades haripunkti pärast 70. eluaastat (Bleeker *et al.*, 2008). Ameerika Ühendriikides on selle esinemissagedus valgete meeste seas ligikaudu 0,29 juhtu 100 000 kohta aastas (Frisch ja Goodman, 2000). Ühendkuningriikides diagnoositakse aastas ligikaudu 760 mehel peenisevähk (Cancer Research UK, 2023b). Taanis läbi viidud uuring näitas, et meestel, kes polnud kunagi kasutanud kondoomi, oli peenisevähi risk enam kui kahekordne võrreldes meestega, kes kasutasid kondoomi. See viitab võimalusele, et kondoomide kasutamine võib vähendada HPV-nakkuse ja seeläbi ka peenisevähi riski (Cancer Research UK, 2023a).

Peenisevähi tekkimise tõenäosus on väiksem meestel, kellel on hea hügieen, praktiseerivad turvalist seksuaalvahekorda ja kellel on eesnahk osaliselt või täielikult eemaldatud (Urology Care Foundation, 2024).

Peenisevähki diagnoositakse biopsiaga, kus peenisel olevat kude uuritakse mikroskoobi all ning seejärel hinnatakse rakkude kasvu. Vähi leviku kahtlustuse korral võetakse lisaks veel proov lümfisõlmelt ning sooritatakse kompuutertomograafia või magnetresonantstomograafia uuring (Urology Care Foundation, 2024).

Varajasel avastamisel on võimalik peenisevähki edukalt ja madala riskiga ravida. Kui kasvaja on väike ja nahapealne, saab ravimiseks kasutada kreemi, laserit ning välist kiiritusravi (ingl *beam radiation*). Keskmise, umbes hernetera suuruse kasvaja korral võib rakendada ebanormaalse koepinna eemaldamist. Suuremõõtmelise kasvaja puhul eemaldatakse rohkem kude ning eemaldataks või tühjendatakse ka lümfisõlmed. Võidakse rakendada ka kirurgia-, kiiritus- ja keemiaravi kombinatsiooni. Kaugele arenenud vähi korral võib vajalikuks osutuda osa või terve peenise eemaldamine (Urology Care Foundation, 2024).



Joonis 5. Peenisevähi illustratsioon (Stecca *et al.*, 2021, kohandatud).

1.2.5. Suu- ja neeluvähk

Suu- ja neeluvähk (vt. Joonis 6) kuuluvad pea- ja kaelapiirkonna pahaloomuliste kasvajate hulka ning on enamikul juhtudest lamerakk-kartsinoomi tüüpi (Ramqvist ja Dalianis, 2010). Halvaloomulised kasvajad on enimlevinud mandlitel (RHK-10 koodiga C09), kuid võivad tekkida ka keelepõhjale (RHK-10 koodiga C01), pehmele suulaele, kõri seintele ning neelu tagumisele osale (RHK-10 koodiga C10) (Cleveland Clinic, i.a; b).

Suu- ja neeluvähi varajased sümptomid võivad olla mittespetsiifilised ning sarnaneda vähem tõsiste seisundite tunnustega. Tüüpilisteks sümptomiteks loetakse püsivat kurguvalu, neelamisraskusi ja -valulikkust, suu avamise või keele liigutamise takistusi. Lisaks võivad võimalikule kasvajale viidata seletamatu kaalulangus, hääle muutused ning valkjas katt keelel või suu limaskestal. Haiguse progresseerudes võivad ilmned raskemad sümptomid, nagu pidev valu kõrvas, vereeritus köhimisel ning tuntav moodustis kõris, suus või kaelas (Cleveland Clinic, i.a; b).

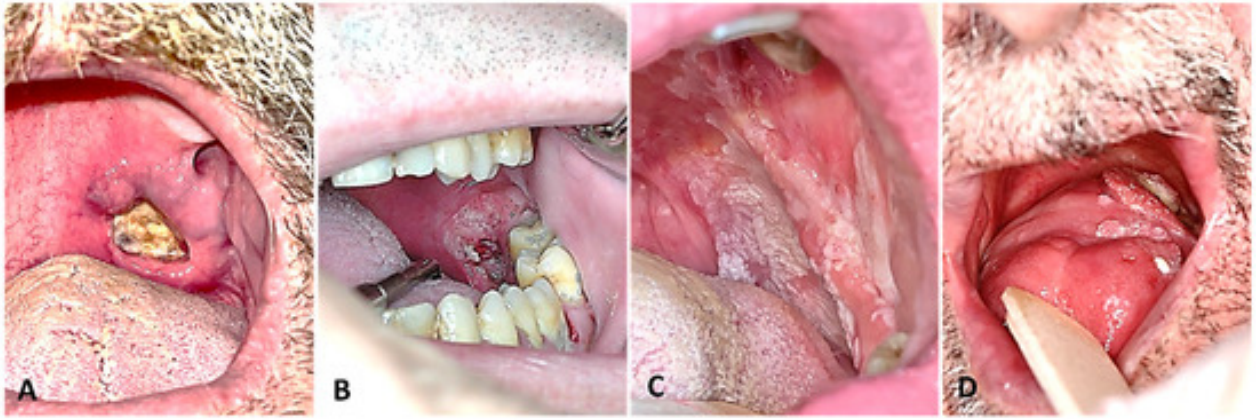
Vähi sagedaseks põhjustajaks on inimese papilloomiviiruse tüüp 16, mille toodetud valgud häirivad suu ja neelu limaskesta rakkude normaalset geeniregulatsiooni. Selle tulemusel kaotavad rakud kontrolli oma kasvu üle ning võivad moodustada kasvaja (Cleveland Clinic, i.a; b). Tubakatoodete, sealhulgas sigarettide, sigarite ja närimistubaka tarvitamine kahjustab

kõri limaskesta rakke, millele reageerides hakkavad rakud tavapärasest intensiivsemalt jagunema. Selle tõttu suureneb risk DNA replikatsioonivigade tekkeks ning need võivad kujuneda vähirakkudeks. Samuti kahjustab kõri limaskesta rakke alkoholi tarbimine (Cleveland Clinic, i.a; b). Suu- ja neeluvähile on leitud ka geneetilisi riskifaktoreid, mis on välja toodud peatükis 1.3.2.

Suu- ja neeluvähk on üsna haruldane ning seda diagnoositakse Ameerika Ühendriikides iga-aastaselt ligikaudu vaid 53 tuhat juhtu (Cleveland Clinic, i.a; b). Seire-, epidemioloogia- ja lõpptulemuste (SEER; ingl *Surveillance, Epidemiology, and End Results*) programmi biopanga andmetel vähenes aastatel 1988–2004 suitsetamise esinemissageduse vähenemise tõttu HPV-negatiivse suu- ja neeluvähi levimus 50%, samal ajal kui HPV-positiivsete juhtude esinemissagedus kasvas 225% (Chaturvedi *et al.*, 2011). Rootsi vähiandmestiku põhjal täheldati Stockholmi piirkonnas aastatel 1970–2002 mandlivähi esinemissageduse 2,8-kordset tõusu. Paralleelselt kasvas sama perioodi jooksul HPV-positiivsete juhtude osakaal 23%-lt 68%-ni, viidates HPV nakkuse olulisele rollile haiguse sagenemises (Hammarstedt *et al.*, 2006). Eestis diagnoositakse ligikaudu 300 pea- ja kaelapiirkonna vähi esmajuhtu, mis on eelneva 30 aastaga võrrelduna kahekordistunud (Kliinikum, i.a). 2022. aastal diagnoositi mandli- ja suuneelu pahaloolumuline kasvaja 60 juhtu, millest 37 meestel ja 23 naistel. Samal aastal registreeriti 48 keelepära pahaloolumulise kasvaja juhtu (Tervisestatistika ja terviseuuringute andmebaas, i.a).

Suu- ja neeluvähi kahtluse korral teostatakse esmalt suuõõne, neelu ning kaela palpatsioon, et tuvastada võimalikke moodustisi. Diagnoosi täpsustamiseks kasutatakse sageli endoskoopilist biopsiat, mille käigus uuritakse koeproovi võimalike pahaloolumuliste muutuste ning inimese papilloomiviiruse esinemise suhtes. Kui vähk on kinnitust leidnud, hinnatakse kasvaja suurust ja levikut kompuutertomograafia või magnetresonantstomograafia abil. Väiksemad lokaalsed kasvaja on sageli kirurgiliselt eemaldatavad, kuid ulatuslikuma kasvaja korral rakendatakse kombineeritud kiiritus- ja keemiaravi. Lisaks võib ravis kasutada laboratoorselt toodetud monoklonaalseid antikehi, mis blokeerivad vähirakkude kasvu ja levikut soodustavaid valke. Vähi taastekke või laialdase leviku korral kasutatakse ka immunoteraapiat (Cleveland Clinic, i.a; b).

Suu- ja neeluvähi esinemise riski suurendavad mitmed tegurid, sealhulgas kokkupuude inimese papilloomiviirusega, tubakatoodete kasutamine ning rohke alkoholi tarbimine. Riskifaktoriteks peetakse ka varasemalt diagnoositud pea- ja kaelapiirkonna kasvaja ning nende raviks saadud kiiritusravi (Cleveland Clinic, i.a; b).



Joonis 6. Suu- ja neeluvähi asukohad (Nuñez-Vera *et al.*, 2024). A) pehme suulagi. B) mandlite lohk. C) külgmine neelu sein. D) keelepõhi.

1.3. Inimese papilloomiviirusega seotud diagnooside geneetilised riskifaktorid

Mitmete eelpoolmainitud diagnooside riskifaktoritena on mainitud ka geneetilist eelsoodumust. Kuigi vastavaid uuringuid pole selles kontekstis palju tehtud, annan järgnevas peatükis ülevaate geneetiliste seoseuuringute läbiviimise põhimõtetest ning peamistest tulemustest eelpoolkirjeldatud diagnooside kontekstis.

1.3.1. Geneetiliste seoseuuringute ülevaade

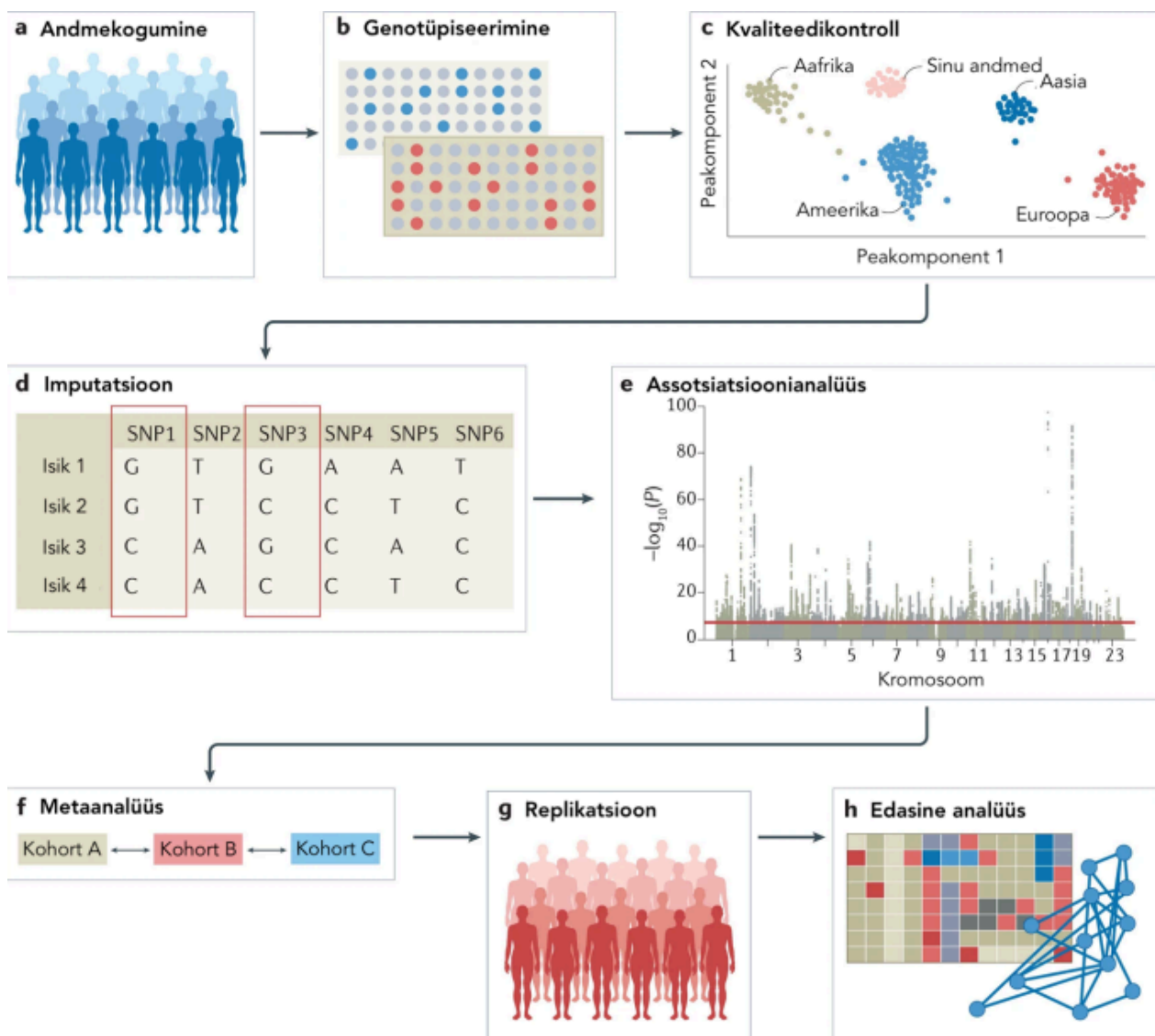
Geneetiliste seoseuuringute (GWAS, ingl *genome-wide association study*) peamine eesmärk on tuvastada geneetilisi variatsioone, mis on seotud erinevate tunnuste ja haigustega. Enim uuritud variandid on üksiknukleotiidsed polümorfismid (SNP-d, ingl *single nucleotide polymorphisms*) (Uffelmann *et al.*, 2021), mis esinevad inimese genoomis keskmiselt iga tuhande nukleotiidi kohta. Hinnanguliselt on ühe indiviidi genoomis neli kuni viis miljonit SNP-d, millest enamik paiknevad geenidevahelistes või mittekodeerivates piirkondades. Kodeerivates piirkondades olevad SNP-d on olulised, sest need võivad mõjutada valgu struktuure ja funktsioone ja seega ka bioloogilisi protsesse haigusriskile (CD Genomics; 2023). Funktsionaalsete mittekodeerivate variantide kõige olulisem roll on osaleda geeniekspressiooni regulatsioonis (Zhang ja Lupski, 2015), muutes ekspressiooni ajastust, asukohta või taset (Learn.Genetics; i.a).

Geneetiliste seoseuuringute töövoog koosneb mitmest etapist, mis on üldjoontes sarnased allpool kirjeldatule (vt. Joonis 7), kuid võib esineda erisusi tulenevalt uuringute eripärast. Esmalt kaasatakse uuringusse individid, kellelt võetakse vereproov ning andmed fenotüübi

kohta, sealhulgas vanus, sugu ja haigusseisund. Teiseks, indiviidide DNA genotüpiseeritakse kasutades mikrokiipe eelnevalt valitud markeritega. Kolmandaks, teostatakse kogutud andmetele kvaliteedikontroll (Uffelmann *et al.*, 2021), milles eemaldatakse haruldased ning puudulike andmetega variandid, mille genotüpiseerimine ei õnnestunud (Coleman *et al.*, 2015). Neljandaks, täidetakse genotüpiseerimata jäänud lüngad imputatsiooni abil (Uffelmann *et al.*, 2021), mis on statistiline meetod otseselt genotüpiseerimata variantide ennustamiseks (Li *et al.*, 2009). Tänapäeval on laialt levinud suurte biopankade geeni- ja terviseandmete kasutamine ja sellisel juhul on punktid 1–4 sageli juba eelnevalt teostatud. Viiendaks, viiakse läbi statistiline analüüs, et tuvastada huvipakkuva tunnusega seotud geneetilised variandid. Seejärel on võimalus teostada metaanalüüs väiksemate kohortide tulemuste liitmiseks, suurendamaks uuringu võimsust. Kuuendaks, kinnitatakse saadud tulemusi võimalusel geneetiliselt sarnases sõltumatus kohordis. Saadud uuringu väljundiks on tunnusega seotud geneetilised variandid ning nende efekti suurus. Tulemuste interpreteerimiseks on näiteks võimalik edasistes analüüsides vaadata, milliseid geene ning bioloogilisi protsesse tuvastatud geneetilised variandid mõjutavad (Uffelmann *et al.*, 2021). Näiteks on üheks levinumaks ja lihtsamini kasutatavaks lähenemiseks FUMA veebipõhine platvorm, mis kombineerib statistilisi leide ja bioloogilisi andmeid (Watanabe *et al.*, 2017).

Tüüpiline binaarse tunnuse GWAS põhineb juhtude-kontrollide uuringul, milles juhud ja kontrollid määratakse vastavalt teatud fenotüübi olemasolule või puudumisele (Pirinen *et al.*, 2012).

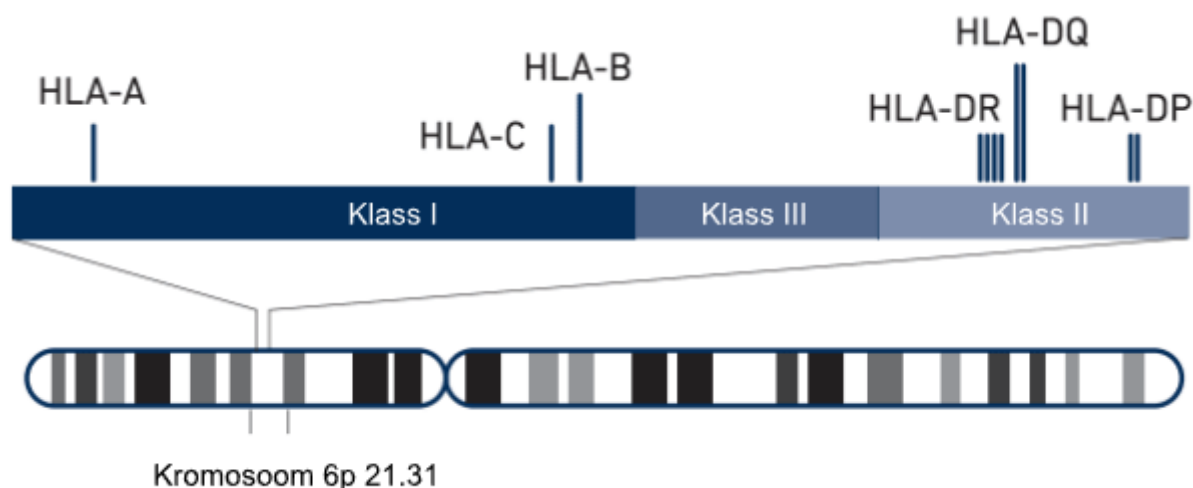
GWAS-i riskilookuste iseloomustamise esimeseks osaks on geneetiliste variantide omavahelise ahelduse tasakaalutuse (LD, ingl *linkage disequilibrium*) hindamine, võttes arvesse populatsioonispetsiifilist LD struktuuri (Watanabe *et al.*, 2017). LD on genoomis samas piirkonnas paiknevate SNP-de omadus, mis näitab, kui suure tõenäosusega ühe SNP alleel on pärandatud koos teise SNP alleeliga konkreetses populatsioonis või kui tugevalt need omavahel korreleeruvad (Bush ja Moore, 2012). Infot variantide omavahelise ahelduse kohta kasutatakse ka sõltumatute oluliste SNP-de tuvastamiseks (Watanabe *et al.*, 2017).



Joonis 7. Skemaatiline ülevaade ülegenoomse seoseuuringu töövoost (Uffelmann *et al.*, 2021, kohandatud).

1.3.2. Ülevaade inimese papilloomiviiruse põhjustatud diagnooside geneetilistest riskifaktoritest

Peamised geneetika uuringud inimese papilloomiviirusest tingitud haigustega seoses on teostatud emakakaelavähi kontekstis (Adebamowo *et al.*, 2020; Beckhaus *et al.*, 2025), kus eriti tugev seos on leitud 6. kromosoomil peamise koosobivuskompleksi (MHC)/ inimese leukotsüütide antigeeni (HLA) geeniklastris *HLA-DPB1/2* ja *HLA-DPA1* geenidega (Shi *et al.*, 2013). Inimese leukotsüütide antigeenide süsteem jaguneb kolmeks peamiseks alampiirkonnaks: I, II ja III klassi geenid (vt. Joonis 8). Kõik need piirkonnad mängivad olulist rolli immuunvastuse moduleerimisel ning immuunsüsteemi aktiivsuse reguleerimisel, sealhulgas ka immunosupressioonis (Klein ja Sato, 2000).



Joonis 8. HLA geeniklastri regiooni diagramm (Clark, 2016, kohandatud).

Valdav enamus senistest töödest on oma olemuselt kandidaatgeenide uuringud (ülevaade Amos *et al.*, 2010), kuid need ei võimalda tuvastada uusi gene ega geenikombinatsioone. Seevastu ülegenoomsed seoseuuringud ehk GWAS uuringud võimaldavad kogu genoomi ulatuses tuvastada seoseid geenivariantide ja haigusriskide vahel eelneva bioloogilise hüpoteesita (Cooke *et al.*, 2008).

Viirustüügastega on seostatud *HLA-DRB1* geeni ja *HLA-DRB1 PheSerTyr13* aminohapete polümorfismi *HLA-DRB1* peptiidi siduvates lõhedes. Viirustüügaste kujunemisega on seostatud ka varianti rs6692209, mis asub *LCE3E* geeni läheduses (Tian *et al.*, 2017). HLA geenikompleks paikneb 6. kromosoomil ning kodeerib rakupinnal ekspresseeruvaid valke, millel on oluline roll immuunregulatsioonis. HLA geenide variatsioonid on seostatud erinevate haigustega, sealhulgas infektsioonide, autoimmuun- ja degeneratiivsete haiguste ning vähiga (Dendrou *et al.*, 2018).

On leitud, et *HLA-DRB1* ja *DQB1* geenivariandid mõjutavad kõrge riski HPV tüüpide 16, 18, 31, 33, 45 ja 58 elimineerimist organismist. Näiteks *HLA-DRB1*12:01:01G* alleel soodustab HPV-16 ja HPV-45 eemaldamist, kuid on seotud HPV 18, 31 ja 58 tüüpide püsima jäämisega. See viitab sellele, et teatud HLA haplotüüpide mõju sõltub spetsiifilistest viiruse epitoopidest, mida need molekulid suudavad T-rakkudele esitleda (Del Río-Ospina *et al.*, 2020).

HPV-positiivse häbemevähi märkimisväärt suurenenud riski on seostatud geenide *CSF2* (rs25882 ja rs27438) ning *IL-12B* (rs2569254 ja rs3181225) variantidega (Hardikar *et al.*, 2015). Samuti on kõrge risk seostatud ka *IL12B* geeni rs3181225, rs3181224 ja rs3212227 variantidega (Hussain *et al.*, 2013). Olulist seost häbemevähi riskiga on näidanud

kasvaja nekroosifaktori (TNF, ingl *tumor necrosis factor*) piirkonnas asuvad SNP-d, nagu *LST1* geeni rs2256965 G-alleel ja *LTA* geeni 5' UTR-is paiknev SNP rs2239704 A-alleel (Bodelon *et al.*, 2013).

Peenisevähi geneetilisi riskifaktoreid on uuritud väga vähe ning seni avaldatud andmed on vastuolulised (Whibley *et al.*, 2009). Üheks enim uuritud geneetiliseks markeriks HPV-ga seotud vähkide puhul on kandidaatgeenuuringute meta-analüüsides esile kerkinud *TP53* geeni polümorfism *p.Arg72Pro* (rs1042522). Siiski pole selle variandi seos peenisevähi tekkega seniste andmete põhjal olnud statistiliselt usaldusväärselt kinnitatud (Hou *et al.*, 2015). HLA geenipiirkonna seos peenise intraepiteliaalse neoplaasiaga (PeIN) on leidnud kinnitust Ühendkuningriigis läbiviidud uurimistöös, kus tuvastati *HLA-C15* alleeli esinemisel suurenenud risk PeIN kujunemiseks, samas kui *HLA-DQA101* alleelil täheldati kaitsvat toimet. Pärast mitmese testimise korrektsiooni ei olnud seos HPV16-positiivse PeIN ja HLA genotüüpide vahel enam statistiliselt oluline (Shim *et al.*, 2021).

Pea- ja kaelapiirkonna kasvajate riski suurenemine on seotud kroonilise põletiku ning pikaajalise kokkupuutega kantserogeensete teguritega, mis soodustavad kumulatiivseid geneetilisi muutusi olulistes regulatoorsetes geenides, nagu *TP53*, *CDKN2A*, *NOTCH1* ja *CCND1* (Leemans *et al.*, 2018). Samuti seostatakse HPV nakkuse püsimisega *TGFβ* geeni, mis kodeerib põletikuvastast tsütokiini transformeerivat kasvufaktori beetavalku (*TGFβ* valku), mis toimib olulise immuunvastuse negatiivse regulaatorina (Wakabayashi *et al.*, 2019). Guan *et al.* uuring näitas positiivset seost *TGFβ1* geeni T869C variandi ja HPV16-positiivse pea- ja kaelapiirkonna vähi vahel (Guan *et al.*, 2010). Sama variant mõjutas ka ravivastust ja ellujäämist HPV-positiivse suuneelu pahaloomulise kasvaja patsientidel (Tao *et al.*, 2018). See viitab, et immuun geenide variandid võivad suurendada geneetilist eelsoodumust HPV-ga seotud pea- ja kaelapiirkonna kasvajate tekkeks (Espinoza *et al.*, 2021). Suu- ja neeluvähi GWAS-uuringus Euroopa, Põhja- ja Lõuna-Ameerika populatsioonides tuvastati kaheksa riskilookust. Kombineeritud suu- ja neeluvähki seostati piirkondadega 6p21.32 (rs3828805, *HLA-DQB1*), 10q26.13 (rs201982221, *LHPP*) ja 11p15.4 (rs1453414, *OR52N2-TRIM5*). Suuvähiga seostusid veel piirkonnad 2p23.3 (rs6547741, *GPNI*) ja 9q34.12 (rs928674, *LAMC3*), 9p21.3 (rs8181047, *CDKN2B-AS1*) ja 5p15.33 (rs10462706, *CLPTMIL*). Neeluvähi puhul piirdusid geneetilised seosed *HLA* piirkonnaga (Lesseur *et al.*, 2016) 2p23.1 (rs4318431, *GALNT14*), 6p21.33 (rs13211972, *MUC21*) ja 6p21.32 (rs34518860, *HLA-DQA1*) lookustes (Shete *et al.*, 2020).

Kuigi HPV-ga seotud haiguste geneetilist tausta on uuritud erinevates kontekstides, on ülegenoomseid seoseuuringuid seni läbi viidud suhteliselt vähe (ülevaade Amos *et al.*, 2010).

Olemasolevad suuremahulised uuringud viitavad siiski sellele, et geneetilistel teguritel on oluline roll HPV infektsiooni püsimise ja sellest tulenevate pahaloomuliste muutuste kujunemisel (Shi *et al.*, 2013). HPV on seotud erinevate haigustega, millest tuntuim ja enim uuritud on emakakaelavähk. Vähem on teada teiste HPV-ga seotud haiguste (viirustüükad, genitaalkondüloomid, kurgu- ja genitaalpiirkonna muud vähid) geneetilistest riskifaktoritest. Tulemused näitavad, et geneetilised seosed ilmnevad eelkõige immuunsüsteemi geenides, nagu HLA kompleks (Shi *et al.*, 2013), tsütokiinide kodeerivad geenid (Wakabayashi *et al.*, 2019) ja kasvajasupressorid (MOFFITT Cancer Center, i.a; a). Neist joonistub välja ühine muster – organismi võimekus HPV nakkust kontrolli all hoida ja sellele efektiivselt immuunvastust anda võib olla osaliselt määratud päriliku geneetilise eelsoodumusega. See loob eeldused tulevasteks uuringuteks, kus piisava valimi suuruse ja täpsete fenotüübiliste andmete abil oleks võimalik tuvastada uusi riskilookusi ning paremini mõista HPV-ga seotud haiguste geneetilisi põhjuseid. Inimese papilloomiviirusega seotud haiguste geneetiliste riskifaktorite uurimine on oluline haiguse eelsoodumuse ja bioloogia mõistmiseks.

Varasemate analüüside kirjeldatud geneetilised variandid ja kandidaatgeenid genitaalpiirkonna ning suu- ja neeluvähi riskideks on esitatud ülevahtlikult Tabelis 2.

Tabel 2. Inimese papilloomiviiruse seos varasemates uuringutes kirjeldatud genitaalpiirkonna ning suu- ja neeluvähi geneetiliste variantidega.

Diagnoos	Geneetiline variant	Kandidaatgeenid	Uuringutüüp	Allikas
Häbemevähk	rs25882 rs27438	<i>CSF2</i>	kandidaatgeeniuuring	(Hardikar <i>et al.</i> , 2015)
	rs2569254 rs3181225	<i>IL-12B</i>		
	rs2256965 rs2239704	<i>LST1</i> <i>LTA</i>	kandidaatgeeniuuring	(Bodelon <i>et al.</i> , 2013)
	rs3181225 rs3181224 rs3212227	<i>IL12B</i>	kandidaatgeeniuuring	(Hussain <i>et al.</i> , 2013)
Peenisevähk	rs1042522	<i>TP53</i> geeni polümorfism <i>p.Arg72Pro</i>	kandidaatgeeniuuring	(Stoehr <i>et al.</i> , 2018)

Suuvähk	rs6547741 rs928674 rs8181047 rs10462706	<i>GPNI</i> <i>LAMC3</i> <i>CDKN2B-AS1</i> <i>CLPTMIL</i>	GWAS-uuring	(Lesseur <i>et al.</i> , 2016)
Suu- ja neeluvähk kombineerituna	rs3828805 rs201982221 rs1453414	<i>HLA-DQB1</i> <i>LHPP</i> <i>OR52N2-TRIM5</i>	GWAS-uuring	(Lesseur <i>et al.</i> , 2016)
Neeluvähk	rs4318431 rs13211972 rs34518860	<i>GALNT14</i> <i>MUC21</i> <i>HLA-DQA1</i>	GWAS-uuring	(Shete <i>et al.</i> , 2020)

2. EKSPERIMENTAALOSA

2.1. Töö eesmärgid

Käesoleva töö eesmärgiks oli tuvastada ja iseloomustada inimese papilloomiviirusega seotud viirustüügaste, genitaalkondüloomide ning häbeme-, tupe-, päraku-, peenise-, suu- ja neeluvähi geneetilisi riskifaktoreid, võttes aluseks Eesti geenivaramu andmed.

2.2. Materjal ja metoodika

2.2.1. TÜ Eesti geenivaramu valim ja uuritavad tunnused

TÜ Eesti geenivaramu (TÜ EGV) geenidoonoriteks on vähemalt 18-aastased täisealised, kes on vabatahtlikult allkirjastanud nõusolekuvormi, täitnud laiaulatusliku tervise-seoselise küsimustiku ning loovutanud oma vereproovi DNA, valgeverelibled ja plasma eraldamiseks. Geenidoonorite andmed on pidevas täiustamises üleriigilise Tervisekassa andmebaasi kaudu (Leitsalu *et al.*, 2015; Milani *et al.*, 2025). TÜ EGV-ga on liitunud üle 200 000 Eesti elaniku, mis hõlmab ligikaudu 20% täiskasvanud populatsioonist (Milani *et al.*, 2025).

Uuringusse valiti isikud, kellele on diagnoositud viirustüükad (B07) ja genitaalkondüloomid (A63.0). Harva esinevate HPV-ga seotud pahaloormuliste kasvaja analüüsimiseks grupeeriti samasse piirkonda kuuluvad juhud genitaal-, päraku- ja peenisevähi (C21_C60_C52_C51) ning suu- ja neeluvähi (C10_C09_C01) diagnoosidega. Uuringu juhtudeks valiti isikud, kellel esines mõni eelpool loetletud diagnoosidest. Kontrollideks valiti isikud, kellel pole meile teadaolevalt vastavat haigust diagnoositud. Nii olid näiteks viirustüügaste analüüsis juhtudeks inimesed, kellel esines B07 diagnoos ja kontrollideks inimesed, kellel B07 diagnoosi pole. Vähi diagnooside analüüsis välistati kontrollide hulgast ka inimesed mõne teise vähipaikme diagnoosiga isikud. Käesolevas töös esitatud analüüsid on teostatud Eesti bioetika ja inimuuringute nõukogu poolt väljastatud eetikalubade 1.1-12/2732 ning 1.1-12/624 raames.

Töös kasutatud valimite suurused on esitatud Tabelis 3.

Tabel 3 . Uuritavate diagnooside juhtude ja kontrollide arv TÜ Eesti geenivaramu kohordis.

Diagnoos	Juhtude defineerimiseks kasutatud RHK-10 kood	Juhtude arv	Kontrollide arv
Teravad genitaalkondüloomid	A63.1	3281	96127
Viirustüükad ehk soolatüükad	B07	25263	132039
Suu- ja neeluvähk	C10_C09_C01	124	183930
Genitaal- ja pärakuvähk, peenisevähk	C21_C60_C52_C51	305	183930

2.2.2. Genotüüpiseerimine

Käesolevas bakalaureusetöös kasutati olemasolevaid genotüüpiseerimise andmeid. Geenidoonorite proovide genotüüpiseerimine toimus keskselt TÜ Eesti geenivaramu tuumiklaboris. Genotüüpiseerimist teostati kasutades Illumina GSAv1.0, GSAv2.0 ja GSAv2.0_EST kiipe.

Genotüüpiseerimise järgselt teostati kvaliteedikontroll, kus indiviidi- ja markeripõhises kontrollis arvestati järgmiseid parameetreid ja näitajaid:

- edukalt genotüüpiseeritud markerite protsentuaalne osakaal (ingl *call rate*, CR);
- genotüübilise ja fenotüübilise soo mittevastavus;
- Hardy-Weinbergi tasakaalustatus (ingl *Hardy-Weinberg equilibrium*, HWE);
- vähemesineva alleeli sagedus (ingl *minor allele frequency*, MAF);
- indiviidide omavaheline sugulus;
- heterosügootsus.

Pärast genotüüpiseerimist jäeti analüüsist välja individid, kelle edukalt genotüüpiseeritud markerite osakaal (CR) oli alla 95% ning need, kelle genotüüpide põhjal määratud sugu ei ühtinud fenotüübis märgitud sooga. Imputatsiooni eelselt eemaldati geneetilised variandid, mille CR oli alla 95%, Hardy-Weinbergi tasakaalu (HWE) P-väärtus oli väiksem kui 1×10^{-4}

(autosomaalsed variandid) ning mitte-autosomaalsed markerid ja variandid, mille harvema alleeli sagedus (MAF) jäi alla 1% (Pujol-Gualdo *et al.*, 2022). Eelfaasimine viidi läbi Eagle v2.3 tarkvaraga (Loh *et al.*, 2016; Pujol-Gualdo *et al.*, 2022). Genotüpiseerimata SNP-de imputatsioon teostati Beagle v.28Sep18.793 tarkvaraga (Browning ja Browning, 2007), kasutades Eesti-spetsiifilist referentspaneeli, mis koosneb 2297 geenidoonori kogugenoomidest (Mitt *et al.*, 2017).

2.2.3. Ülegenoomne seoseanalüüs TÜ Eesti geenivaramu andmetel

Seoseanalüüs TÜ EGV andmetel viidi läbi kasutades tarkvarapaketti REGENIE v2.2.4 (Mbatchou *et al.*, 2021). REGENIE on kiire ja mälukasutuselt efektiivne tarkvara, mis võimaldab analüüsida mitmeid tunnuseid paralleelselt. Samuti võimaldab REGENIE segamudel analüüsi kaasata ka suguluses olevaid isikuid (Mbatchou *et al.*, 2021). Analüüsis kasutati kovariaatidena sugu, sünniaastat ja 10 esimest geneetilist peakomponenti (PC). Madala imputatsiooni kvaliteediga (INFO skoor < 0,4) geneetilised variandid ning variandid, mille harvema alleeli koguarv < 5, jäeti analüüsist välja (Pujol-Gualdo *et al.*, 2022). Inimese papilloomiviirusega seotud diagnooside GWAS-analüüse teostas Mariann Võsa.

2.2.4. Seoseanalüüsi tõlgendamine

Enamik GWAS-i leide paikneb mittekodeerivates piirkondades (Maurano *et al.*, 2012), ja seega on sageli keeruline geneetilise seoseanalüüsi tulemuste põhjal teha järeldusi seotud geenide ja funktsionaalsete mehhanismide kohta. Sageli katab GWAS-i leid suuremat genoomipiirkonda (GWAS-i riskilookus), mis sisaldab mitmeid omavahel seotud SNP-sid ja rohkem kui üht geeni. Mõned neist geenidest võivad olla haiguse seisukohast olulised, teised mitte, kuid lähestikku paiknevate geneetiliste variantide korrelatsiooni tõttu on keeruline määrata, millised geenid on olulised vaid assotsiatsioonipõhiste P-väärtuste alusel. Oluliste geenide ja variantide tuvastamiseks tuleb arvestada piirkonna geneetilise aheldatuse mustreid ning SNP-de funktsionaalseid omadusi, sealhulgas variantide mõju geeniekspressioonile ning nende rolli kromatiini interaktsioonikohtades (Watanabe *et al.*, 2017).

FUMA on veebipõhine platvorm (Watanabe *et al.*, 2017, <https://fuma.ctglab.nl/>), mis aitab tõlgendada geneetilise seoseanalüüsi tulemusi, kombineerides statistilisi leide bioloogiliste andmetega. Selle eesmärk on teha selgeks, millised SNP-d võivad bioloogiliselt olulised olla

ja millised geenid või bioloogilised mehhanismid võivad nende taga olla (Watanabe *et al.*, 2017).

GWAS tulemuse annoteerimiseks kasutasin FUMA platvormi SNP2GENE funktsiooni. SNP2GENE sisendina kasutatakse GWAS-i koondstatistikuid. Analüüsi käigus valitakse välja eelnevalt määratud P-väärtuse lävendi alusel olulised SNP-d ning nendega kasutaja määratud r^2 alusel LD-s olevad SNP-d, kus $r^2 = 1$ viitab täielikule LD-le ehk genotüübid on täielikult üksteisega aheldunud (Watanabe *et al.*, 2017).

2.2.5. Funktsionaalne kaardistamine ja geneetiliste variantide prioritseerimine

Selgitamaks välja, millised geenid võivad olla kaasatud inimese papilloomiviirusega seotud diagnoosidega, visualiseerisin ülegenoomsete seoseanalüüside tulemused FUMA platvormi (v1.5.2) SNP2GENE funktsiooni abil. Selleks laadisin üles geneetilise seoseanalüüsi koondstatistikute failid, mis sisaldasid teavet iga analüüsitud variandi kohta, sealhulgas kromosoomi, genoomset asukohta, P-väärtust ja rsID ehk variandi identifikaatorit. Kandidaatgeenide kaardistamisel FUMA-s kaasasin ka MHC regioonid, mis muidu on vaikimisi kaardistamisel välja jäetud keerulise LD struktuuri tõttu. Esimeses etapis analüüsiti üksikute geneetiliste variantide seoseid uuritavate diagnoosidega. Seejärel esitab veebiplatvorm GWAS-i tulemuste ülevaate, kus on esile toodud tunnusega seotud geneetilised riskilookused ning nende sees paiknevad statistiliselt olulised ($P < 5 \times 10^{-8}$) juhtivad geneetilised variandid. FUMA visualiseeris automaatselt Manhattani diagrammi, mis kuvab GWAS-i analüüsi tulemusi. Juhtivate ja kandidaat SNP-de määramisel kasutasin programmi vaikeväärtusi.

FUMA kaardistab sõltumatud olulised SNP-d (ülegenoomselt olulised ja $r^2 < 0,6$ teiste ülegenoomselt olulise SNP-dega). Nende hulgast valitakse juhtsignaalid ($r^2 < 0,1$) ning iga juhtiva SNP jaoks kaardistatakse sama lookuse piires kandidaatvariandid, mis korreleeruvad juhtiva SNP-ga ($r^2 \geq 0,6$). Selle protsessi tulemusel saadakse riskilookused, mis koosnevad sõltumatult olulistest SNP-dest ja kandidaat-SNP-dest.

Kandidaatgeenide prioritseerimisel kasutasin järgnevat strateegiat:

1. FUMA platvormi kasutasin riskilookuste ja juhtsignaalide määramiseks. Kuna enamus/ kõik signaalid olid MHC/ HLA regioonis, siis ei toetunud FUMA eQTL ja kromatiininteraktsioonide abil kaardistamisele, sest HLA piirkonna LD struktuur on

väga keeruline ning see raskendab nende analüüside tõlgendamist.

2. Seoseanalüüsist saadud geneetiliste variantide lähimate geenide leidmiseks kasutasin Open Targets Genetics tööriista (Ghoussaini *et al.*, 2021; Mountjoy *et al.*, 2021). Veebiaadressi <https://genetics.opentargets.org/> kasutasin andmekogumiseks ajavahemikus september 2024 kuni märts 2025. Iga juhtsignaali variandi puhul sisestasin selle tähise otsinguväljale ning analüüsisin tulemuste loetelus lähima geeni (ingl *nearest gene*) rida, milles oli välja toodud lähima geeni nimetus ning selle kaugus vastavast variandist. Samuti kasutasin *National Center for Biotechnology Information*-i (NCBI) ühe nukleotiidi polümorfismi andmebaasi (ingl *The Single Nucleotide Polymorphism Database*, dbSNP) (Sherry *et al.*, 2001). Veebiaadressi <https://www.ncbi.nlm.nih.gov/snp/> kasutasin andmekogumiseks detsember 2025 juhtsignaalide puhul, millele Open Targets Genetics tööriist tulemusi ei andnud. Otsinguväljale sisestasin variandi tähise ning tulemustes kerisin reale *Genomic regions, transcripts, and products*, kus jooniselt otsisin lähima geeni.
3. MAGMA (ingl *Multi- marker Analysis of GenoMic Annotation* (Leeuw *et al.*, 2015)) v1.08 geenipõhine kaardistamismeetod kasutab mitme muutujaga lineaarse regressiooni mudelit geenide P-väärtuste arvutamiseks. Meetod hindab geenisiseste SNP-de assotsiatsioone uuritava tunnusega, võttes arvesse ka SNP-de vahelist ahelduse tasakaalutust (Leeuw *et al.*, 2015). Statistiliselt olulisteks peeti vaid neid geene, mille $P < 2,5 \times 10^{-6}$, et vältida mitmesest testimisest (testitud valku kodeerivate geenide arv on 19760) tingitud valepositiivseid tulemusi.

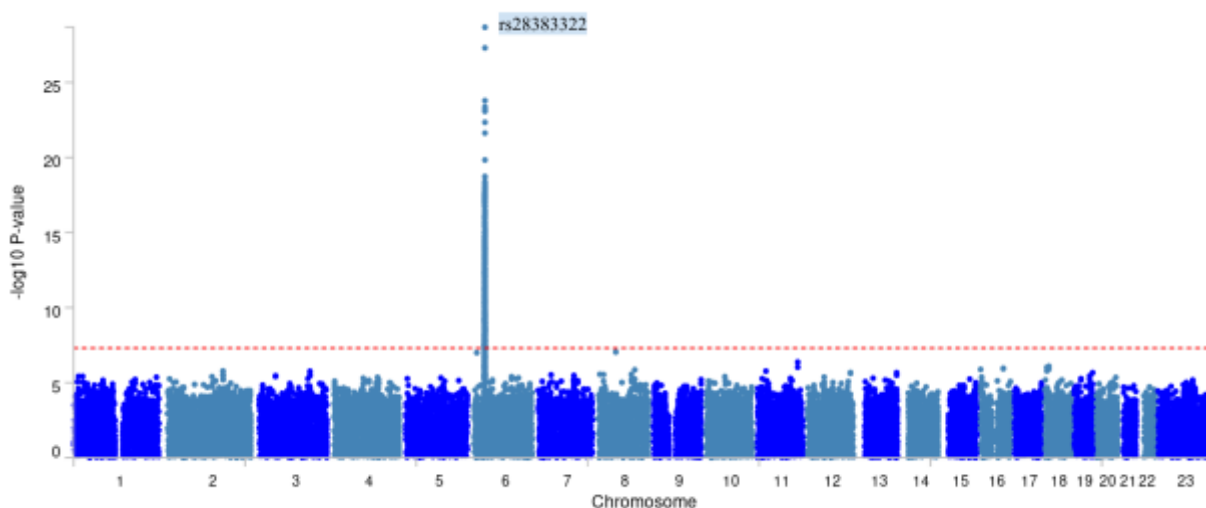
2.2.6. Varasemate assotsiatsioonide leidmine ja tulemuste replitseerumine

Juhtsignaaliga seotud varasemate assotsiatsioonide leidmiseks kasutasin FUMA GWAS kataloogi, kuhu on koondatud tulemused avaldatud ülegenoomsetest seoseuuringutest. Otsingus sisestasin juhtiva SNP tunnuse ning analüüsisin selle esinemist eelnevates uuringutes (Watanabe *et al.*, 2017). Lisaks uurisin tulemuste replitseerumist Soome FinnGen andmebaasi 12. väljaande (R12) (FinnGen, 2023, <http://r12.finnngen.fi/>) kaudu, et hinnata geneetiliste variantide seotust sarnaste diagnoosidega sõltumatus populatsioonis. Otsinguribale sisestasin juhtsignaalid ning võrdlesin tulemusi FUMA platvormi tulemustega.

2.3. Tulemused

2.3.1. Teravad genitaalkondüloomid

Teravate genitaalkondüloomide seoseanalüüsil kaardistati 128 statistiliselt olulist ($P < 5 \times 10^{-8}$) varianti 6. kromosoomil (vt. Joonis 9). Seoseanalüüsi tulemused on esitatud Tabelis 4. Madalaima P-väärtusega $2,11 \times 10^{-29}$ juhtiv SNP (rs28383322; 6:32592796:C:T) asub 3160 kb ülesvoolu *HLA-DQA1* geenist. Leitud varianti on seostatud mitmete teiste tunnustega, näiteks astma ja erinevad vere biomarkerid (Sun *et al.*, 2018; Zhu *et al.*, 2019; Zhu *et al.*, 2020). Soome FinnGen R12 andmebaasis on seostatud rs28383322 geneetilist varianti genitaalkondüloomidega ($P=1,8 \times 10^{-77}$) ning ka viirustüügaste ($P=2,0 \times 10^{-21}$) ja *in situ* emakakaelavähiga ($P=1,5 \times 10^{-11}$). Lisaks on FinnGen R12 andmebaasis leitud *HLA-DQA1* geeni muude variantide oluline seos viirustüügaste (rs9271669, $P=1,6 \times 10^{-48}$), genitaalkondüloomide (rs9272324, $P=1,3 \times 10^{-79}$) ja emakakaelavähiga (rs144081603, $P=1,6 \times 10^{-8}$). Samas andmebaasis on geenivariandid nominaalselt olulises ($P < 0,05$) seoses päraku- (rs9501184, $P=4,4 \times 10^{-7}$), peenise- (rs16870693, $P=1,9 \times 10^{-6}$), neelu- (rs756361095, $P=7,4 \times 10^{-6}$), hääbeme- (rs2856713, $P=1,2 \times 10^{-5}$) ja tupevähiga (rs768487433, $P=1,0 \times 10^{-3}$) (FinnGen, 2023).

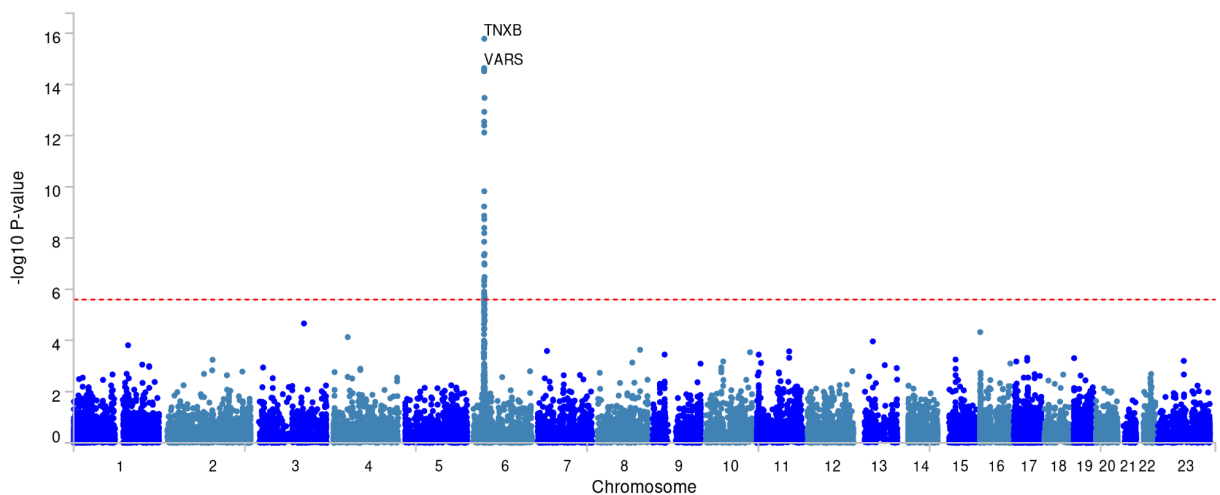


Joonis 9. Teravate genitaalkondüloomide (RHK-10 A63.0) ülegenoomse seoseuuringu Manhattan'i joonis. X-teljel on märgitud variantide genoomsed positsioonid ning Y-teljel P-väärtuste negatiivne kümnendlogaritm. Punane katkendjoon tähistab P-väärtust 5×10^{-8} . Madalaima P-väärtusega SNP ehk juhtiv SNP on märgitud 6. kromosoomil rs28383322 (6:32592796:C:T).

MAGMA geenipõhises analüüsis tuvastati 32 statistiliselt olulist geeni, mis ületas soovitatud olulisuse piiri ($P=2,5 \times 10^{-6}$), millest kaks olulisemat on *TNXB* ja *VARS* (vt. Joonis 10). *TNXB* geen asub 6.kromosoomi MHC klass III piirkonnas (NCBI, i.a; a) ning sellel on oluline roll sidekoe struktuuris ja funktsioonis, säilitamaks kudede elastsust (MedlinePlus Genetics, i.a; a). *VARS* geen asub 6. kromosoomi MHC klass III piirkonnas ning kodeerib aminohapete sidumiseks tRNA-dega vajalikku ensüümi (NCBI, i.a; b), mis osaleb tRNA aminoatsetüleerimises ja valkude biosünteesis, mängides olulist rolli rakulises ainevahetuses ja neuroarengulistes protsessides (Gene Cards, i.a; a).

FinnGen R12 andmebaasis on oluline seos *TNXB* geeniga viirustüügastel (rs115463529, $P=2,7 \times 10^{-13}$), *in situ* emakakaelavähil (rs206018, $P=3,4 \times 10^{-11}$) ja genitaaltüügastel (rs3132941, $P=7,6 \times 10^{-75}$). Soome FinnGen R12 andmebaasis on *TNXB* geen nominaalselt olulises ($P < 0,05$) seoses suu- (rs2070600, $P=4,5 \times 10^{-6}$), neelu- (rs568874520, $P=8,5 \times 10^{-6}$), peenise- (rs1052386239, $P=2,1 \times 10^{-4}$), häbeme- (rs13209014, $P=2,6 \times 10^{-5}$), tupe- (rs3130346, $P=4,0 \times 10^{-4}$) ja pärakuvähiga (rs12215107, $P=4,1 \times 10^{-4}$) (FinnGen, 2023), kuid suurema valimi korral võiksid need seosed muutuda ka ülegenoomselt oluliseks.

VARS geeni seos viirustüügaste (rs113525864, $P=6,9 \times 10^{-13}$) ja genitaalkonduloomidega (rs501942, $P=4,3 \times 10^{-70}$) on leitud ka FinnGen R12 andmebaasis. FinnGen R12 andmebaasis on *VARS* geen nominaalselt olulises ($P < 0,05$) seoses emakakaela- (rs743399, $P=3,7 \times 10^{-6}$), neelu- (rs373597058, $P=8,6 \times 10^{-6}$), häbeme- (rs112947839, rs749030782, rs3037244, $P=1,5 \times 10^{-5}$), tupe- (rs1297672328, $P=8,9 \times 10^{-4}$), pära- (rs1488538040, $P=8,8 \times 10^{-5}$) ja peenisevähiga (rs149807555,rs70990287, $P=1,6 \times 10^{-4}$) (FinnGen, 2023).

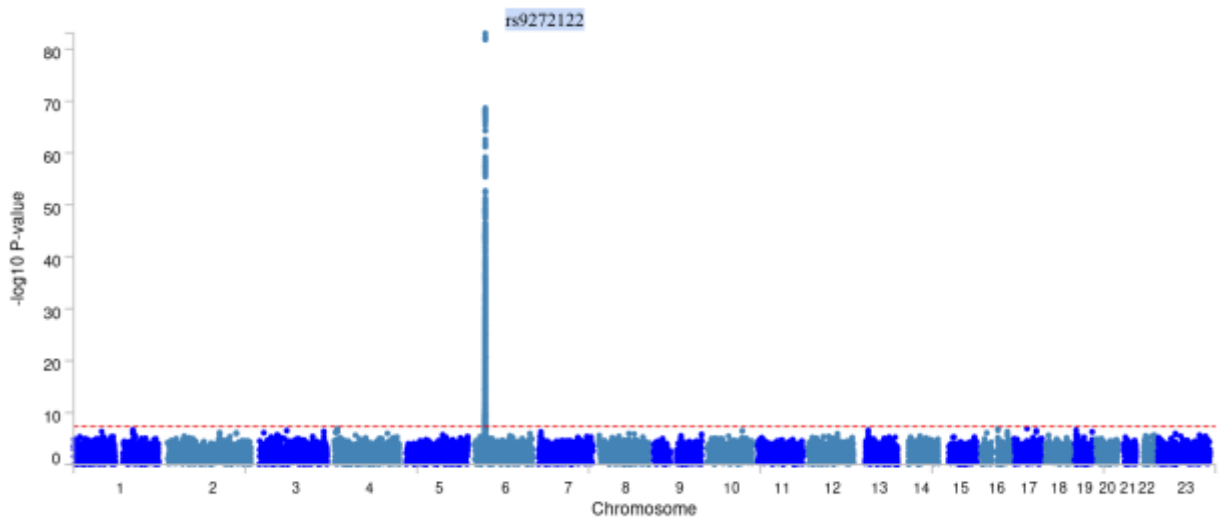


Joonis 10. Manhattan'i joonis MAGMA geenipõhise seoseanalüüsi kohta. X-teljel on märgitud geenide genoomsed positsioonid ning Y-teljel P-väärtuste negatiivne kümnendlogaritm. Ülegenoomse statistilise olulisuse piiriks on $P=2,5 \times 10^{-6}$ (punane katkendjoon).

2.3.2. Viirustüükad ehk soolatüükad

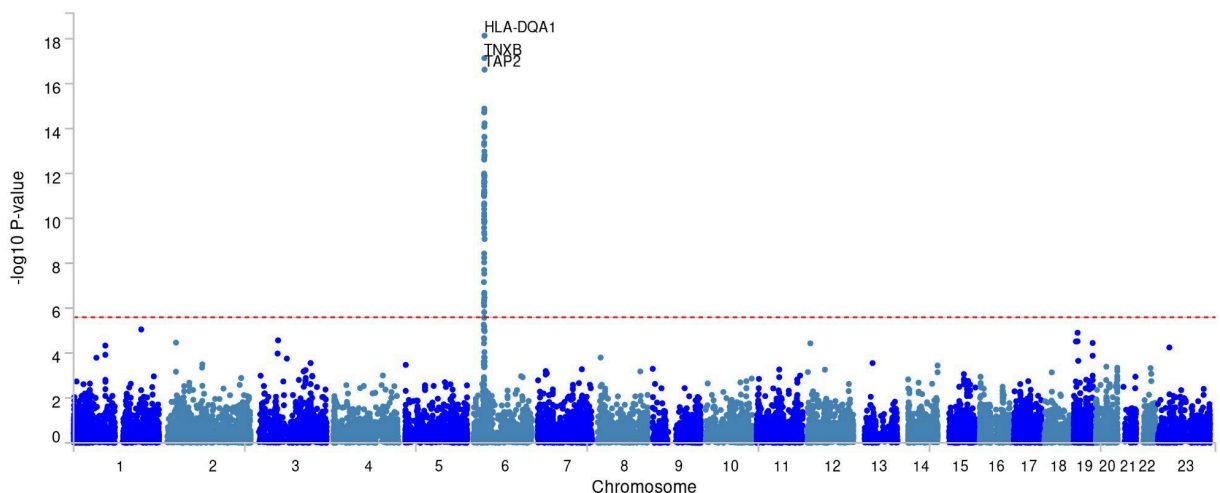
Soolatüügaste seoseanalüüsil kaardistati 353 statistiliselt olulist ($P < 5 \times 10^{-8}$) varianti 6. kromosoomil (vt. Joonis 11). Seoseanalüüsi tulemused on esitatud Tabelis 4. Madalaima P-väärtusega ($P=8.9 \times 10^{-84}$) juhtiv SNP (rs9272122; 6:32600562:C:T) asub *HLA-DQA1* geeni 5'UTR regioonis.

Geneetiline variant rs9272122 on tugevas LD-s ($r^2=0,89$) rs9272050 variandiga, mida on seostatud soolatüügaste ja emakakaelavähiga (Tian *et al.*, 2017; Rashkin *et al.*, 2020). Juhtiv SNP on tugevas LD-s ($r^2=0.70$) ka rs9272226 variandiga, mida on seostatud täiskasvanute astmaga (An *et al.*, 2021). Soome FinnGen R12 andmebaasis on seostatud rs9272122 geneetilist varianti viirustüügaste ($P=1,6 \times 10^{-48}$), astma ($P=3,3 \times 10^{-26}$), emakakaelavähi ($P=1,1 \times 10^{-20}$) ja atoopilise dermatiidiga ($P=3,1 \times 10^{-17}$) (FinnGen, 2023).



Joonis 11. Soolatüügaste (RHK-10 B07) ülegenoomse seoseuuringu Manhattan'i joonis. X-teljel on märgitud variantide genoomsed positsioonid ning Y-teljel P-väärtuste negatiivne kümnendlogaritm. Punane katkendjoon tähistab P-väärtust 5×10^{-8} . Madalaima P-väärtusega SNP ehk juhtiv SNP on märgitud 6. kromosoomil rs9272122 (chr 6:32600562:C:T).

MAGMA geenipõhises analüüsis tuvastati 63 statistiliselt olulist geeni, mis ületas soovitud olulisuse piiri ($P=2,5 \times 10^{-6}$, Bonferroni korrigeerimine 19760 valgu kodeeriva geeni assotsiatsiooni testil), millest kolm kõige olulisemat on *HLA-DQA1*, *TNXB* ja *TAP2* (vt. Joonis 12). *HLA-DQA1* geen asub 6. kromosoomil ning kodeerib HLA klass II molekuli alfa-ahelat (NCBI, i.a; c), mis on vajalik rakuväliste peptiidide esitlemiseks T-rakkudele ja T-raku vahendatud immuunvastuse käivitamiseks (Gene Cards, i.a; b). *TNXB* ja *HLA-DQA1* geenide seos esines ka genitaalkonduloomidega (vt. peatükk 1.3.2). *TAP2* geen kodeerib ABC-transporterit, mis moodustab TAP-kompleksi ja vahendab peptiidide ülekannet tsütosoolist endoplasmaatilisse retiikulumi MHC I antigeenide jaoks (NCBI, i.a; d). *TAP2* geenivariantide seos viirustüügaste (rs241441, $P=1,5 \times 10^{-24}$) ja genitaalkonduloomidega (rs9280183, rs371952381, $P=2,8 \times 10^{-65}$) on leitud ka FinnGen R12 andmebaasis. FinnGen R12 andmebaasis on *TAP2* geen nominaalselt olulises ($P < 0,05$) seoses peenise- (rs16870693, $P=1,9 \times 10^{-6}$), päraku- (rs181565586, $P=5,4 \times 10^{-6}$), emakakaela- (rs544834814, $P=1,0 \times 10^{-5}$), neelu- (rs115433903, $P=1,7 \times 10^{-5}$), hääme- (rs12207915, $P=2,3 \times 10^{-5}$) ja tupevähiga (rs3132133, $P=5,1 \times 10^{-4}$) (FinnGen, 2023).



Joonis 12. Manhattan'i joonis MAGMA geenipõhise seoseanalüüsi kohta. X-teljel on märgitud geenide genoomsed positsioonid ning Y-teljel P-väärtuste negatiivne kümnendlogaritm. Ülegenoomse statistilise olulisuse piiriks on $P=2,5 \times 10^{-6}$ (punane katkendjoon).

Tabel 4. Inimese papilloomiviirusega seotud diagnooside seosealüüside tulemused. EAF – efekti-alleeli sagedus (ingl *effect allele frequency*); OR – šansside suhe (ingl *odds ratio*); 95% CI – šansside suhte 95%-usaldusintervall (ingl *confidence interval*).

Diagnoos, RHK-10	Marker	Kromosoom: positsioon	EAF (Efekti-alleel/ teine alleel)	OR (95% CI)	P-väärtus	Lähim geen
Teravad genitaal-kondüloomid, A63.0	rs28383322	6:32592796	0,20 (C/T)	0,68 (0,64–0,73)	$2,1 \times 10^{-29}$	<i>HLA-DQA1</i>
Viirustüükad ehk soolatüükad, B07	rs9272122	6:32600562	0.63 (C/T)	0,81 (0,79–0,83)	$8,9 \times 10^{-84}$	<i>HLA-DQA1</i>

2.3.3. Teised analüüsitud vähid

Lisaks analüüsisin suu- ja neeluvähi ning teiste genitaalpiirkonna vähkide geneetilisi risifaktoreid. Seoseanalüüsil ühtki statistiliselt olulist ($P < 5 \times 10^{-8}$) varianti ei leitud (vt. Joonis 13), mis võib olla põhjustatud liiga väikese valimi kasutamisest. Varasemates suu-, neelu-,

genitaal- ja pärakuvähi uuringutes seostatud geneetilised variandid ja kandidaatgeenid on esitatud Tabelis 2. Seoseanalüüsi koondstatistikute failist saadud varasemate geneetiliste variantide P-väärtused on välja toodud Lisas 1.

Varasemalt häbemevähiga seostatud variandid rs2256965 ja rs2239704 (Bodelon *et al.*, 2013) käesoleva töö raames teostatud analüüsides ülegenoomselt statistiliselt olulised ei olnud. Variandid olid nominaalselt olulised ($P < 0,05$), seega replitseeruvad teostatud analüüsi tulemustega (P-väärtused vastavalt 0,02 ja 0,006). Need variandid on seotud geenidega *LST1* ja *LTA*. *LST1* geeni poolt kodeeritud membraanivalk on võimeline pärssima lümfotsüütide proliferatsiooni (NCBI; i.a; e). *LTA* geeni valk vahendab põletikulisi, immunostimuleerivaid ja viirusevastaseid reaktsioone ning mängib rolli apoptoosis (NCBI; i.a; f). Soome FinnGen R12 andmebaasis on rs2256965 geneetiline variant olulises seoses ning variant rs2239704 nominaalselt olulises seoses genitaalkondüloomidega ($P=1,5 \times 10^{-9}$; $P=6,5 \times 10^{-3}$). *LST1* geen on FinnGen R12 andmebaasis olulises seoses genitaalkondüloomidega (rs3132450, $P=6,0 \times 10^{-69}$) ning nominaalselt olulises seoses pea- ja kaela- (rs760293, $P=3,0 \times 10^{-7}$), päraku- (rs2260051, $P=5,2 \times 10^{-5}$), häbeme- (rs9267532, $P=5,3 \times 10^{-5}$), suu- ja neelu- (rs115128935, $P=9,0 \times 10^{-5}$; rs373597935, $P=8,9 \times 10^{-5}$), emakakaela- (rs2269475, $P=1,0 \times 10^{-4}$), peenise- (rs147926185, $P=2,7 \times 10^{-4}$) ning tupevähiga (rs9380263, $P=7,2 \times 10^{-4}$). *LTA* geen on samuti Soome andmebaasis olulises seoses genitaalkondüloomidega (rs3132450, $P=6,0 \times 10^{-69}$) ning nominaalselt olulises seoses päraku- (rs2260051, $P=5,2 \times 10^{-5}$), häbeme- (rs9267532, $P=5,3 \times 10^{-5}$), suu- ja neelu- (rs115128935, $P=9,0 \times 10^{-5}$; rs373597935, $P=8,9 \times 10^{-5}$), emakakaela- (rs2269475, $P=1,0 \times 10^{-4}$), peenise- (rs147926185, $P=2,7 \times 10^{-4}$) ja tupevähiga (rs9380263, $P=7,2 \times 10^{-4}$) (FinnGen, 2023).

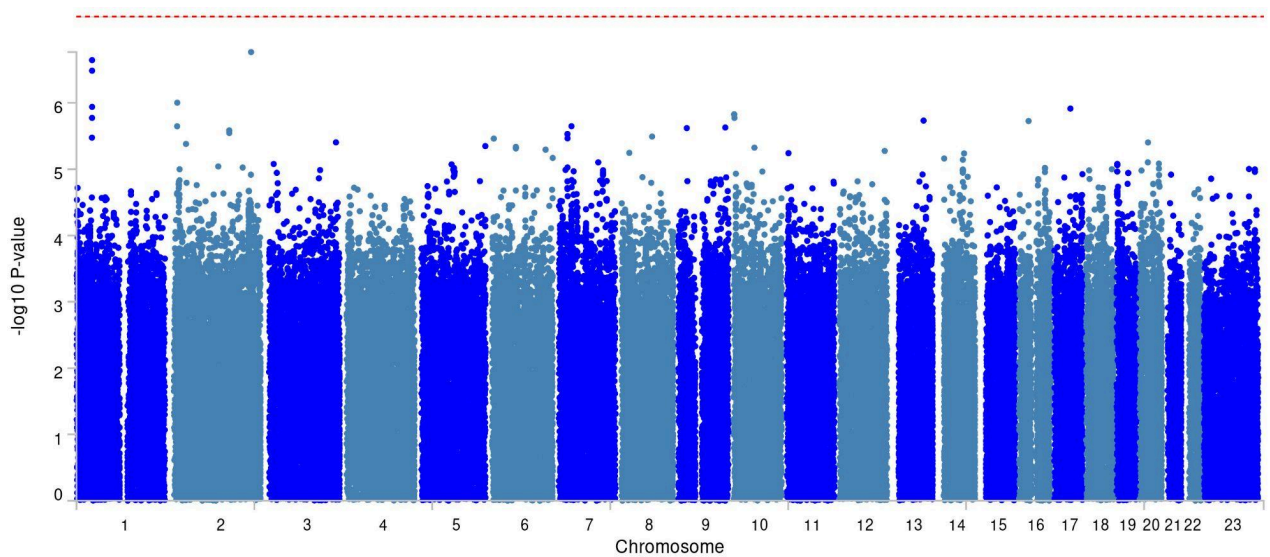
Varasemalt peenisevähiga seostatud variant rs1042522 (Stoeckert *et al.*, 2018) käesoleva töö raames teostatud analüüsides statistiliselt oluline polnud ($P=0,79$). Kandidaatgeeniks leiti *TP53* geeni polümorfism *p.Arg72Pro*. *p.Arg72Pro* on ühe aminohappe asendus *TP53* geenis, kus arginiin (Arg) asendub proliiniga (Pro) 72. positsioonil (Damin *et al.*, 2006). See polümorfism mõjutab p53 valgu võimet indutseerida rakusurma ehk apoptoosi, kus Arg-variant soodustab tugevamat apoptoosi kui Pro-variant, mistõttu võib suurendada vähiriski (Petitjean *et al.*, 2007). Soome FinnGen R12 andmebaasis on rs1042522 geneetiline variant nominaalselt olulises ($P < 0,05$) seoses häbemevähiga ($P=1,4 \times 10^{-2}$) (FinnGen, 2023).

Varasemalt suuvähiga seostatud variant rs8181047 (Lesseur *et al.*, 2016) käesoleva töö raames teostatud analüüsides statistiliselt oluline polnud ($P=0,22$). Kandidaatgeeniks leiti funktsionaalset RNA-d tootev *CDKN2B-AS1*, mis vaigistab teisi gene samas geeniklastris. Seda geeni on seostatud südame-veresoonkonnahaiguste, mitmete vähkide ja Alzheimeri

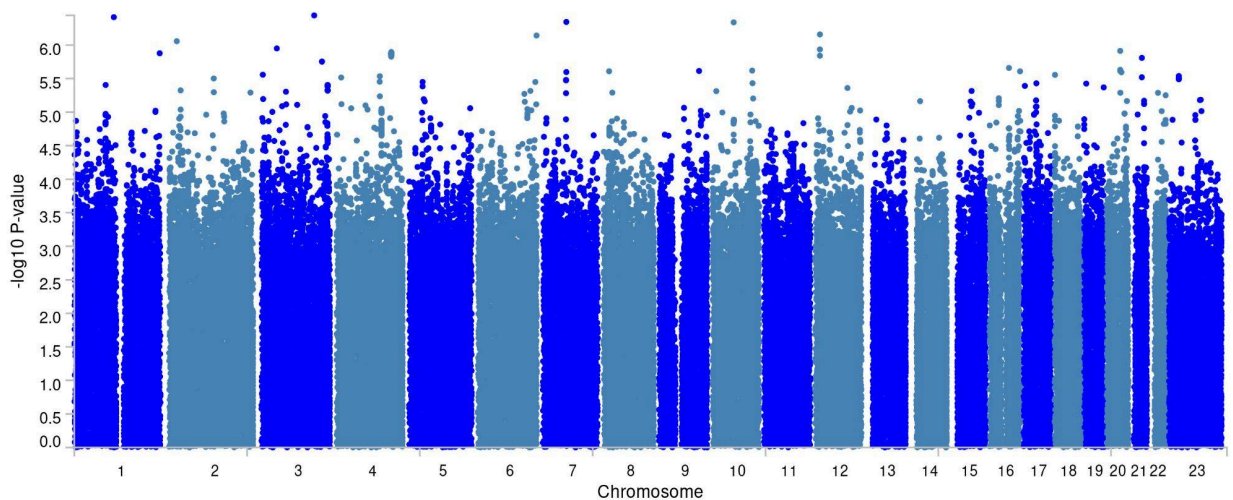
tõvega (NCBI, i.a; g). Eelnevalt neeluvähiga seostatud variant rs4318431 (Shete *et al.*, 2020) ei osutunud samuti seoseanalüüsil oluliseks ($P=0,05$). Kandidaatgeeniks leiti *GALNT14*, mis kodeerib Golgi valku ning selle geeni muutused võivad rolli mängida vähi progresseerumisel ja keemiaravile reageerimisel (NCBI, i.a; h). Soome FinnGen R12 andmebaasis pole rs4318431 geneetiline variant oluline vastavate diagnoosiga, kuid *CDKN2B-AS1* geen on nominaalselt oluline ($P < 0,05$) genitaalkondüloomide (rs72655459, $P=6,9 \times 10^{-4}$), suu- ja neelu- (rs543261612, $P=1,3 \times 10^{-4}$; rs1399803208, $P=2,6 \times 10^{-4}$), peenise- (rs1362205941, $P=2,1 \times 10^{-4}$), häbeme- (rs578002564, $P=5,7 \times 10^{-4}$) ning tupevähiga (rs887622648, $P=6,4 \times 10^{-4}$) (FinnGen, 2023).

Varasemalt suu- ja neeluvähiga (kombineerituna) seostatud variant rs3828805 (Lesseur *et al.*, 2016) käesoleva töö raames teostatud analüüsides statistiliselt oluline ei olnud (P -väärtus 0,1). Kandidaatgeeniks leiti *HLA-DQB1*, mis kodeerib HLA klass II beeta-ahelat ja mängib olulist rolli antigeenide esitamisel immuunrakkudele, aidates kaasa immuunvastuse kujunemisele (NCBI, i.a; i). Soome FinnGen R12 andmebaasis on rs3828805 geneetiline variant olulises seoses genitaalkondüloomide ($P=5,5 \times 10^{-36}$), pea- ja kaelapiirkonna halvaloomuliste kasvajatega ($P=8,2 \times 10^{-10}$) ning nominaalselt olulises ($P < 0,05$) seoses mandli- ($P=1,6 \times 10^{-4}$) ja emakakaelavähiga ($P=3,8 \times 10^{-2}$) (FinnGen, 2023).

A



B



Joonis 13. (A) Suu- ja neeluvähi markerite seoseuuringu Manhattan'i joonis. (B) Genitaal- ja pärakuvähi ning peenisevähi markerite seoseuuringu Manhattan'i joonis. X-teljel on märgitud variantide genoomsed positsioonid ning Y-teljel P-väärtuste negatiivne kümnendlogaritm. Punane katkendjoon tähistab P-väärtust 5×10^{-8} . Statistiliselt olulised seosed (üle P-väärtuse ulatuvad märged) puuduvad.

2.4. Arutelu

Antud bakalaureusetöö eesmärgiks oli kirjeldada inimese papilloomiviirusega (HPV) seotud diagnooside, nagu viirustüügaste, genitaalkondüloomide, häbeme-, tupe-, päraku-, peenise-, suu- ja neeluvähi geneetilisi riskifaktoreid Tartu Ülikooli Eesti geenivaramu andmetel. Selleks

kaardistati GWAS analüüsi tulemusi, kirjeldati diagnoosidega seotud geneetilisi riskilookuseid ning pakuti välja kandidaatgeene.

Uuringu käigus leiti, et nii genitaalkondüloomide kui ka soolatüügastega on seotud üks lookus kuuendal kromosoomil. Statistiliselt oluline seos viirustüügaste ja genitaalkondüloomidega esines *HLA-DQA1* geeni piirkonnas paiknevate SNP-dega, vastavalt rs9272122 ja rs28383322.

Teravate genitaalkondüloomidel pole varasemalt geneetilisi riskifaktoreid publitseeritult kirjeldatud, mistõttu võib käesolevat uurimistööd pidada esimeseks katseks selle diagnoosi geneetilise tausta kaardistamisel. Töö käigus tuvastati võimalik geneetilise seose olemasolu variandiga rs28383322, mis replitseerub ka Soome FinnGen R12 andmebaasis. Kuna publitseeritud andmed puuduvad, siis võib viidata see senitundmata eelsoodumusele genitaaltüügaste kujunemisel. Tuvastatud rs28383322 variant paikneb *HLA-DQA1* geenist ülesvoolu, mis kuulub peamise koesobivuskompleksi (MHC) klass II reguleerimisalasse. On võimalik, et see variant mõjutab geeniekspressiooni, reguleerides *HLA-DQA1* geeni transkriptsiooni ja seeläbi mõjutades peremeesorganismi immuunvastust madala riskiga HPV tüüpide suhtes.

Viirustüügastega seos geneetilise variandiga rs9272122 replitseerib varasemates uuringutes kirjeldatud tulemusi (Tian *et al.*, 2017), mis kinnitab rakendatud analüüsilähenemise korrektsust ja usaldusväärust.

Statistiliselt oluliseks geeniks nii genitaalkondüloomide kui ka viirustüügaste puhul osutus MAGMA geenipõhise analüüsi põhal *TNXB* geen, mis kodeerib tenaskiini-XB ekstratsellulaarse maatriksi glükoproteiini basaalmembraani välimises retikulaarses kihis (Pénisson-Besnier *et al.*, 2013). See valk mängib kesket rolli keha lihaseid, liigeseid, siseorganeid ja nahka toetavate sisekudede organiseerimises ning struktuurse terviklikkuse säilitamises (MedlinePlus Genetics, i.a; a). *TNXB* geeni defitsiiti on seostatud sidekoehaiguse Ehlers-Danlose sündroomiga (NCBI, i.a; a). Soome FinnGen andmebaasis seostus *TNXB* geen viirustüügaste, emakakaelavähi ning genitaalkondüloomidega. Genitaalpiirkonna ning suu- ja neeluvähiga võib nominaalne seos selle geeniga muutuda oluliseks valimi suurendamisel.

Viirustüügaste puhul olid MAGMA geenipõhise analüüsi alusel olulised kandidaatgeenid ka *HLA-DQA1* ja *TAP2* ning genitaalkondüloomide puhul *VARS* geen.

HLA-DQA1 geen osaleb keskse komponendina immuunvastuses, esitledes väliskeskkonnast rakuvälistest valkudes pärinevaid peptiide (NCBI, i.a; c). *HLA-DQA1* geen on seotud

autoimmuunhaigustega nagu tsöliaakia ja tsöliaakia tüüp 1 (Gene Cards, i.a; b), mida tihti nimetatakse ka gluteenitalumatuseks (Ida-Tallinna Keskhaigla, i.a). Samuti on geeni seostatud inimese papilloomiviirusest tingitud emakakaelavähi tekkeriskiga (Chen *et al.*, 2013) ning soolatüügaste tekkega (Sánchez-Barrientos *et al.*, 2021). Publitseeritud andmed *HLA-DQA1* geeni ja genitaalkonduloomide vahelisest seosest puuduvad, kuid siiski võivad selle geeni variatsioonid mõjutada organismi võimet tuvastada ja reageerida patogeenidele, sealhulgas nakkusi põhjustavatele viirustele (MedlinePlus Genetics, i.a; b).

TAP2 geen kodeerib membraaniga seotud ABC-transportervalku, mis kuulub *MDR/TAP* alamperekonda ning mängib olulist rolli antigeenide esitluses. Geeni alternatiivne splaissing võib mõjutada peptiidide valikut ja MHC I molekulide rakupinnal ekspressiooni. *TAP2* geeni mutatsioonid võivad olla seotud insuliinsõltuva suhkurtõve, tsöliaakia ja anküloseeriva spondüliidiga (NCBI, i.a; d) ehk lülisamba liigshaigusega (Inimene, i.a). Gostout *et al.* (2003) uuringus leiti, et emakakaelavähiga patsientidel esineb statistiliselt oluliselt sagedamini *TAP2* geeni A/B heterosügootne alleelivariant võrreldes tervete kontrollisikutega (Gostout *et al.*, 2003). Homosügootne *LMP7/ TAP2* haplotüüp C võib suurendada söögitoruvähi tekkeriski ning mõjutada kasvaja arengut HPV infektsiooniga isikutel (Cao *et al.*, 2005), kuid suurema valimi korral võiksid need seosed muutuda ka ülegenoomselt oluliseks.

VAR5 geenis kodeeritud valk kuulub klass I aminoatsüül-tRNA süntetaaside hulka, mis katalüüsivad tRNA aminoatsüleerimist vastava aminohappega (NCBI, i.a; b). *VAR5* geen on seotud mitme haruldase haigusega, sealhulgas neuroarenguhäirega, mida iseloomustavad ajukahjustus, krambid ja kortikaalne atroofia, ning oksüdatiivse fosforüleerimise kombineeritud defitsiidi tüübiga 20 (ingl *Combined Oxidative Phosphorylation Deficiency 20*). *VAR5* geenil on oluline roll valkude sünteesis ja neuroarengulistes protsessides, kuid teaduslikud tõendid selle seotusest HPV infektsioonide või genitaalkonduloomidega puuduvad.

TNXB ja *VAR5* geenide rolli inimese papilloomiviiruse infektsiooni puhul pole teaduslikult veel kirjeldatud, kuid mõlemad geenid paiknevad MHC klass III piirkonnas, mis on tuntud oma keeruka regulatoorse struktuuri ja paljude immuunvastust mõjutavate geenide poolest. Võimalik, et tuvastatud seosed peegeldavad piirkonna laiemat regulatiivset mõju immuunvastusele.

Mitmed mainitud geenidega seotud variandid on statistiliselt oluliselt seotud viirustüügaste, genitaalkonduloomide ja emakakaelavähi fenotüüpidega FinnGen R12 andmebaasis.

Mõningate fenotüüpide korral, nagu häbeme-, tupe-, päraku-, peenise-, suu- ja neeluvähk, on seosed nominaalselt olulised, kuid ilmselt on asi valimi suuruses ning on võimalik, et laiendatud valimiga võiksid need seosed muutuda ka ülegenoomselt oluliseks.

Leitud geneetilised seosed paiknevad *HLA* piirkonnas, millel on oluline roll immuunvastuse kujunemisel (Dendrou et al., 2018). See on kooskõlas varasemate teadustöödega, kus *HLA* geenide polümorfismid on seostatud nii HPV nakkuse püsimise kui ka viirusest tingitud haiguste, sealhulgas häbemevähi, tekkeriskiga (Wakabayashi et al., 2019). Tian *et al.* (2017) uuringus kirjeldati rs9272122 seost HPV-ga seotud infektsiooniga, sealhulgas viirustüügastega (Tian *et al.*, 2017). Häbemevähiga seostatud variandid rs2256965 ja rs2239704 (Bodelon *et al.*, 2013) replitseerusid seoseanalüüsil nominaalselt. Teiste HPV-ga seotud vähkide osas statistiliselt olulisi lookusi ei tuvastatud, mis võib olla seotud valimi ebapiisava suurusega.

Läbiviidud uuringu kitsaskohaks võib olla asjaolu, et kasutatud valimite suurus on väike, eriti harvamate diagnooside jaoks, mis vähendas statistilist võimsust võimalike seoste tuvastamiseks. Lisaks võib MHC piirkonna keeruline ahelduse struktuur raskendada üksikute variantide bioloogilise tähenduse tõlgendamist. Tulemustel võib olla populatsioonispetsiifiline iseloom, kuna valim piirdus Tartu Ülikooli Eesti geenivaramu andmetega, mistõttu ei pruugi need olla täielikult üldistavad teistele populatsioonidele. Samas replitseeruvad tulemused Soome andmetel, mis näitab, et need pole ilmselt ainult Eesti-spetsiifilised. Analüüsi võimsuse suurendamiseks võiks kaasata teisi kohorte teistest populatsioonidest, et kirjeldada geneetilisi riskifaktoreid teistes populatsioonides ning seega parandada haiguse tekkemehhanismide arusaama. Kuna Soome FinnGen andmebaasi 12. väljaande (R12) seoseuuringute tulemused kattuvad TÜ Eesti geenivaramu andmetel saaduga soolatüügaste ja genitaaltüügaste puhul, siis võiks tulevikus neile teostada meta-analüüsi. Tulevastes uuringutes võiks keskenduda ühelt poolt konkreetsete HPV tüvedega seotud geneetilistele seostele ning teiselt poolt laiendada analüüsi ka nakkuse kulgu iseloomustavatele tunnustele, nagu infektsiooni kestus ja taasnakatamine. Funktsionaalsed uuringud leitud variantide bioloogilise mõju hindamiseks võimaldaksid täpsustada variantide mõjule immuunvastuses ja haiguste kujunemisel. Rakukultuuri- ja loomkatsete abil funktsionaalne valideerimine võimaldaks kinnitada geneetiliste variantide seost konkreetse bioloogilise efektiga ning seejärel seostada bioloogiline efekt uuritava fenotüübi või haigusega.

KOKKUVÕTE

Käesoleva bakalaureusetöö eesmärgiks oli tuvastada ja iseloomustada inimese papilloomiviirusega (HPV) seotud viirustüügaste, genitaalkondüloomide ning häbeme-, tupe-, päraku-, peenise-, suu- ja neeluvähi geneetilisi riskifaktoreid. Selleks kasutati Tartu Ülikooli Eesti Geenivaramu andmeid, viidi läbi ülegenoomne seoseuuring ning annoteeriti saadud tulemusi FUMA veebiplatvormil. Lisaks analüüsiti saadud tulemuste replitseerumist Soome FinnGen R12 andmebaasis, et hinnata geneetiliste variantide seotust sarnaste diagnoosidega sõltumatus populatsioonis.

Töö tulemusel leiti üks lookus kuuendal kromosoomil, mis on seotud genitaalkondüloomide ja soolatüügastega. Mõlema fenotüübi puhul oli lähim geen *HLA-DQA1*, mis on oluline immuunvastuse kujundamisel ning varasemalt seostatud inimese papilloomiviirusnakkusest tingitud vähkide tekkeriskiga. Need tulemused viitavad sellele, et peremeesorganismi geneetilised faktorid, eriti immuunvastusega seotud geenid, võivad mõjutada vastuvõtlikkust erinevatele inimese papilloomiviirusega seotud haigutele. Kuigi erinevate HPV tüvede põhjustatud diagnoosid ja nende patogenees on erinevad, on vähemalt osa immuunvastusest nende vastu universaalne. Statistiliselt oluline kandidaatgeen nii soolatüügaste kui ka genitaalkondüloomide puhul oli *TNXB* geen. Soolatüügaste puhul osutus oluliseks kandidaatgeeniks ka *TAP2* geen ning genitaalkondüloomidel *VARS* geen. Saadud tulemused replitseeruvad FinnGen R12 andmebaasis, mis kinnitab nende geenide olulisust haigusseisundite kujunemisel.

Genitaal- ning pea- ja kaelapiirkonna vähkide harva esinemise tõttu grupeeriti sama paikme juhud, moodustades kaks gruppi häbeme-, tupe-, päraku- ja peenisevähi ning suu- ja neeluvähi diagnoosidega. Siiski jäi valim liiga väikseks, mistõttu statistiliselt olulisi variante seoseanalüüsil ei tuvastatud. Analüüsi täienduseks võrreldi varasemalt kirjeldatud geneetilisi variante nii käesoleva töö raames läbi viidud analüüsidega kui Soome FinnGen R12 andmebaasiga ning ilmnesid nominaalselt olulised ($P < 0,05$) seosed immuunsüsteemiga seotud geenides nagu *LST1* ja *LTA*, viidates nende potentsiaalsele bioloogilisele tähtsusele.

The genetic risk factors for human papillomavirus-related diagnoses

Steffi Salu

SUMMARY

Human papillomavirus (HPV) is one of the most common viral infections worldwide, with the majority of individuals exposed during their lifetime (Chesson *et al.*, 2014). While most infections are transient and manifest as benign lesions such as viral warts or genital condylomas, persistent infection can lead to malignancies in the anogenital, head and neck regions (Kidd *et al.*, 2017; American Cancer Society, i.a; American Cancer Society, i.a; Ramqvist ja Dalianis, 2010). The progression of HPV-related disease depends on several factors, including viral type, transmission patterns, immune system status, co-existing health conditions, and lifestyle, including smoking and drinking alcohol (DermNet, i.a; BHC, i.a; Mayo Clinic Staff, 2023; Cancer Research UK, 2023a; Cleveland Clinic, i.a; b). Genetic susceptibility also plays an important role in the development of these HPV-related diseases, but the exact biological mechanisms remain poorly understood compared to cervical cancer, the most widely recognized consequence of HPV infection (Adebamowo *et al.*, 2020; Beckhaus *et al.*, 2025).

The aim of this bachelor's thesis was to identify and characterize the genetic susceptibility factors associated with HPV-related diagnoses, including viral warts, genital condylomas, and cancers of the vulva, vagina, anus, penis, mouth, and oropharynx. For this purpose, data from the Estonian Biobank were used to conduct a genome-wide association study (GWAS), and the results were annotated using the FUMA web platform. In addition, the findings were evaluated for replication in the independent Finnish FinnGen R12 database to assess their relevance in an external population.

As a result of the conducted analyses, one locus on chromosome 6 was found to be associated with both genital condylomas and viral warts. For both phenotypes, the nearest gene was *HLA-DQA1*, which plays a key role in immune response and antigen presentation. Additional statistically significant candidate genes included *TNXB* for both phenotypes, *TAP2* for viral warts, and *VARS* for genital condylomas. These associations were supported by the FinnGen R12 dataset, indicating their potential relevance to the development of these conditions.

Due to the low prevalence of HPV-related cancers in the studied cohort, cases of vulvar, vaginal, anal, and penile cancer cases were grouped into one broader phenotype, while oral and oropharyngeal cancers were grouped into another. No genome-wide significant variants

were identified, likely due to limited statistical power. However, based on comparisons with the FinnGen R12 database and findings from previous studies, nominally significant associations ($P < 0,05$) were observed in genes involved in immune regulation, such as *LST1* and *LTA*. These results suggest that host genetic factors, particularly those related to immune function, may influence susceptibility to various HPV-related conditions.

TÄNUSÕNAD

Eelkõige tahan tänada enda juhendajat Triin Laiska, kes oli igal hetkel mulle toeks ning motiveeris edasi pürgima. Suured tänud ka kaasjuhendaja Mariann Võsale, kes teostas tööks vajalikud GWAS-analüüsid ning kelle nõuanded olid abistavad. Samuti tänan Reedik Mäge, kelle suunamisel (töö) temaatilisele kursusele aitas bakalaureusetöö tulemuste sisukamale mõistmisele.

Täna ka oma sõpru, lähedasi ja kursusekaaslasi, kes olid mulle töö kirjutamisel toeks ja aitasid kaasa positiivsuse hoidmisel.

KIRJANDUSE LOETELU

- Adebamowo, S. N., Adeyemo, A. A., Rotimi, C. N., Olaniyan, O., Offiong, R., Adebamowo, C. A. (2020). Genome-wide association study of prevalent and persistent cervical high-risk human papillomavirus (HPV) infection. *BMC Medical Genetics*, 21(1). DOI: 10.1186/s12881-020-01156-1
- Amos, W., Driscoll, E., Hoffman, J. I. (2010). Candidate genes versus genome-wide associations: which are better for detecting genetic susceptibility to infectious disease? *Proceedings of the Royal Society B: Biological Sciences*, 278(1709), 1183–1188. DOI: 10.1098/rspb.2010.1920
- An, J., Ra, A. D., Kang, H. Y., ... Kim, T. (2021). Genome-Wide Association Study of Korean Asthmatics: A comparison with UK asthmatics. *Allergy Asthma and Immunology Research*, 13(4), 609. DOI: 10.4168/aair.2021.13.4.609
- Beckhaus, T., Kachuri, L., Nakase, T., ... Dörk, T. (2025). Genome-Wide Association Analyses of HPV16 and HPV18 Seropositivity Identify Susceptibility Loci for Cervical Cancer. *Journal of Medical Virology*, 97(2). DOI: 10.1002/jmv.70195
- Bleeker, M. C. G., Heideman, D. a. M., Snijders, P. J. F., Horenblas, S., Dillner, J., Meijer, C. J. L. M. (2008). Penile cancer: epidemiology, pathogenesis and prevention. *World Journal of Urology*, 27(2). DOI: 10.1007/s00345-008-0302-z
- Bodelon, C., Madeleine, M. M., Johnson, L. G., Du, Q., Galloway, D. A., Malkki, M., Petersdorf, E. W., Schwartz, S. M. (2013). Genetic variation in the TLR and NF- κ B pathways and cervical and vulvar cancer risk: A population-based case-control study. *International Journal of Cancer*, 134(2), 437–444. DOI: 10.1002/ijc.28364
- Browning, S. R., 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.
- Bush, W. S., Moore, J. H. (2012). Chapter 11: Genome-Wide Association Studies. *PLoS Computational Biology*, 8(12), e1002822. DOI: 10.1371/journal.pcbi.1002822
- Cao, B., Tian, X., Li, Y., Jiang, P., Ning, T., Xing, H., Zhao, Y., Zhang, C., Shi, X., Chen, D., Shen, Y., Ke, Y. (2005). LMP7/TAP2 gene polymorphisms and HPV infection in esophageal

carcinoma patients from a high incidence area in China. *Carcinogenesis*, 26(7), 1280–1284. DOI: 10.1093/carcin/bgi071

Centers for Disease Control and Prevention (CDC) (2024). HPV Vaccination. <https://www.cdc.gov/hpv/vaccines/index.html>

Chaturvedi, A. K., Engels, E. A., Pfeiffer, R. M., ... Gillison, M. L. (2011). Human Papillomavirus and Rising Oropharyngeal Cancer Incidence in the United States. *Journal of Clinical Oncology*, 29(32), 4294–4301. DOI: 10.1200/jco.2011.36.4596

Chen, D., Juko-Pecirep, I., Hammer, J., Ivansson, E., Enroth, S., Gustavsson, I., Feuk, L., Magnusson, P. K. E., McKay, J. D., Wilander, E., Gyllensten, U. (2013). Genome-wide Association Study of Susceptibility Loci for Cervical Cancer. *JNCI: Journal of the National Cancer Institute*, 105(9), 624–633. DOI: 10.1093/jnci/djt051

Chesson, H. W., Dunne, E. F., Hariri, S., Markowitz, L. E. (2014). The estimated lifetime probability of acquiring human papillomavirus in the United States. *Sexually Transmitted Diseases*, 41(11), 660–664. DOI: 10.1097/olq.0000000000000193

Clark, S. (2016). More Complete HLA Analysis. <https://www.10xgenomics.com/blog/more-complete-hla-analysis>

Clifford, G., Gallus, S., Herrero, R., ... Franceschi, S. (2005). Worldwide distribution of human papillomavirus types in cytologically normal women in the International Agency for Research on Cancer HPV prevalence surveys: a pooled analysis. *The Lancet*, 366(9490), 991–998. DOI: 10.1016/s0140-6736(05)67069-9

Coleman, J. R. I., Euesden, J., Patel, H., Folarin, A.A., Newhouse, S., Breen, G. (2015). Quality control, imputation and analysis of genome-wide genotyping data from the Illumina HumanCoreExome microarray. *Briefings in Functional Genomics*. 15(4):298–304. DOI: 10.1093/bfgp/elv037

Confido Meditsiinikeskus (2022). SPETSIALIST SELGITAB: mis vahe on konnasilmal ja soolatuukal? <https://www.confido.ee/mis-vahe-on-konnasilmal-ja-soolatuukal/>

Cooke, G. S., Campbell, S. J., Bennett, S., ... Hill, A. V. S. (2008). Mapping of a Novel Susceptibility Locus Suggests a Role for MC3R and CTSZ in Human Tuberculosis. *American Journal of Respiratory and Critical Care Medicine*, 178(2), 203–207. <https://doi.org/10.1164/rccm.200710-1554oc>

- Cubie, H. A. (2013). Diseases associated with human papillomavirus infection. *Virology*, 445(1–2), 21–34. DOI: 10.1016/j.virol.2013.06.007
- Curado, MP., Shin, HR., Storm, H., Ferlay, J., Heanue, M., Boyle, P. (2007). Cancer incidence in five continents, Vol. IX. International Agency for Research on Cancer (IARC)
- Daling, J. R., Madeleine, M. M., Johnson, L. G., Schwartz, S. M., Shera, K. A., Wurscher, M. A., Carter, J. J., Porter, P. L., Galloway, D. A., McDougall, J. K. (2004). Human papillomavirus, smoking, and sexual practices in the etiology of anal cancer. *Cancer*, 101(2), 270–280. DOI: 10.1002/cncr.20365
- Damin, A. P., Frazzon, A. P., Damin, D. C., Roehe, A., Hermes, V., Zettler, C., Alexandre, C. O. (2006). Evidence for an association of TP53 codon 72 polymorphism with breast cancer risk. *Cancer Detection and Prevention*, 30(6), 523–529. DOI: 10.1016/j.cdp.2006.09.007
- Del Río-Ospina, L., Camargo, M., Soto-De León, S. C., Sánchez, R., Moreno-Pérez, D. A., Patarroyo, M. E., Patarroyo, M. A. (2020). Identifying the *HLA DRB1-DQB1* molecules and predicting epitopes associated with high-risk HPV infection clearance and redetection. *Scientific Reports*, 10. DOI: 10.1038/s41598-020-64268-x
- Dendrou, C. A., Petersen, J., Rossjohn, J., Fugger, L. (2018). HLA variation and disease. *Nature Reviews Immunology*, 18(5), 325–339. DOI: 10.1038/nri.2017.143
- Dudley, W. (1995). The psychological impact of warts on patients' lives. *Professional Nurse (London, England)*. 11(2): 99-100. PMID: 7480061
- Espinoza, H., Ha, K. T., Pham, T. T., Espinoza, J. L. (2021). Genetic Predisposition to Persistent Human Papillomavirus-Infection and Virus-Induced Cancers. *Microorganisms*, 9(10), 2092. DOI: 10.3390/microorganisms9102092
- Frisch, M., Goodman, M., T. (2000). Human papillomavirus-associated carcinomas in Hawaii and the mainland U.S. *Cancer*, 88 (2000), 1464-1469. DOI: 10.1002/(SICI)1097-0142(20000315)88:6<1464::AID-CNCR26>3.0.CO;2-O
- Ghoussaini, M., Mountjoy, E., Carmona, M., ... Dunham I. (2021) Open Targets Genetics: systematic identification of trait-associated genes using large-scale genetics and functional genomics. *Nucleic Acids Res.* 49(D1): D1311-D1320. DOI: 10.1093/nar/gkaa840

Gostout, B. S., Poland, G. A., Calhoun, E. S., Sohni, Y. R., Giuntoli, R. L., II, McGovern, R. M., Sloan, J. A., Cha, S. S., Persing, D. H. (2003). TAP1, TAP2, and HLA-DR2 alleles are predictors of cervical cancer risk. *Gynecologic Oncology*, 88(3), 326–332. DOI: 10.1016/s0090-8258(02)00074-4

Hammarstedt, L., Lindquist, D., Dahlstrand, H., ... Munck-Wikland, E. (2006). Human papillomavirus as a risk factor for the increase in incidence of tonsillar cancer. *International Journal of Cancer*, 119(11), 2620–2623. DOI: 10.1002/ijc.22177

Hardikar, S., Johnson, L. G., Malkki, M., Petersdorf, E. W., Galloway, D. A., Schwartz, S. M., Madeleine, M. M. (2015). A population-based case–control study of genetic variation in cytokine genes associated with risk of cervical and vulvar cancers. *Gynecologic Oncology*, 139(1), 90–96. DOI: 10.1016/j.ygyno.2015.07.110

HealthCentral (2019). Condyloma (Genital Warts).
<https://www.healthcentral.com/condition/condyloma>

Hou, J., Gu, Y., Hou, W., Wu, S., Lou, Y., Yang, W., Zhu, L., Hu, Y., Sun, M., Xue, H. (2015). P53 codon 72 polymorphism, human papillomavirus infection, and their interaction to oral carcinoma susceptibility. *BMC Genomic Data*, 16(1). DOI: 10.1186/s12863-015-0235-7

Hussain, S. K., Madeleine, M. M., Johnson, L. G., Du, Q., Galloway, D. A., Daling, J. R., Malkki, M., Petersdorf, E. W., Schwartz, S. M. (2013). Nucleotide variation in IL-10 and IL-12 and their receptors and cervical and vulvar cancer risk: a hybrid case-parent triad and case-control study. *International Journal of Cancer*, 133(1), 201–213. DOI: 10.1002/ijc.28000

International Agency for Research on Cancer. (2007). Human papillomaviruses. NCBI Bookshelf. <https://www.ncbi.nlm.nih.gov/books/NBK321760/>

Kalczyński, J., M. (2024). Condyloma Acuminatum (Genital Warts). Medscape. <https://emedicine.medscape.com/article/781735-overview?form=fpf>

Kidd, L.C., Chaing, S., Chipollini, J., Giuliano, A.R., Spiess, P.E., Sharma, P. (2017). Relationship between human papillomavirus and penile cancer—implications for prevention and treatment. *Translational Andrology and Urology*. 6(5): 791-802. DOI: 10.21037/tau.2017.06.27

- Kilkenny, M., Merlin, K., Young, R., Marks, R. (1998). The prevalence of common skin conditions in Australian school students: 1. Common, plane and plantar viral warts. *British Journal of Dermatology*. 138(5):840–845. DOI: 10.1046/j.1365-2133.1998.02222.x
- Klein, J., Sato, A. (2000). The HLA System. *New England Journal of Medicine*, 343(10), 702–709. DOI: 10.1056/nejm200009073431006
- Leemans, C. R., Snijders, P. J. F., Brakenhoff, R. H. (2018). The molecular landscape of head and neck cancer. *Nature Reviews Cancer*, 18(5), 269–282. DOI: 10.1038/nrc.2018.11
- Leeuw, C. A. de, Mooij, J. M., Heskes, T., Posthuma, D. (2015). MAGMA: Generalized Gene-Set Analysis of GWAS Data. *PLoS Comput Biol*, 11(4): e1004219.
- 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. DOI: 10.1093/ije/dyt268
- Leslie, S.W., Sajjad, H., Kumar, S. (2023). Genital Warts. *StatPearls*. <https://www.ncbi.nlm.nih.gov/books/NBK441884/>
- Lesseur, C., Diergaarde, B., Olshan, A. F., ... Robinson, M. (2016). Genome-wide association analyses identify new susceptibility loci for oral cavity and pharyngeal cancer. *Nature Genetics*, 48(12), 1544–1550. DOI: 10.1038/ng.3685
- Li, Y., Willer, C., Sanna, S., Abecasis, G. (2009). Genotype imputation. *Annual Review of Genomics and Human Genetics*, 10(1), 387–406. DOI: 10.1146/annurev.genom.9.081307.164242
- Loh, P.-R., Danecek, P., Palamara, P. F., ... L Price, A. (2016). Reference-based phasing using the Haplotype Reference Consortium panel. *Nat Genet*, 48(11): 1443–1448.
- Loo, S. K., Tang, W. Y. (2014). Warts (non-genital). *BMJ Clinical Evidence*, 2014, 1710. <https://pmc.ncbi.nlm.nih.gov/articles/PMC4054795/>
- Maurano, M. T., Humbert, R., Rynes, E., ... Stamatoyannopoulos, J. A. (2012). Systematic localization of Common Disease-Associated variation in regulatory DNA. *Science*, 337(6099), 1190–1195. DOI: 10.1126/science.1222794
- Mayo Clinic Staff (2023). Anal cancer. *Mayo Clinic*. <https://www.mayoclinic.org/diseases-conditions/anal-cancer/symptoms-causes/syc-20354140>

Mayo Clinic Staff (2025a). Vulvar cancer. Mayo Clinic. <https://www.mayoclinic.org/diseases-conditions/vulvar-cancer/symptoms-causes/syc-2036805>
1

Mayo Clinic Staff (2025b). Vaginal cancer. Mayo Clinic. <https://www.mayoclinic.org/diseases-conditions/vaginal-cancer/symptoms-causes/syc-203524>
47

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. DOI: 10.1038/s41588-021-00870-7

Milani, L., Alver, M., Laur, S., ... Metspalu, A. (2025). The Estonian Biobank's journey from biobanking to personalized medicine. *Nature Communications*, 16(1). DOI: 10.1038/s41467-025-58465-3

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.

Mountjoy, E., Schmidt, E., M., Carmona, M., ... Ghoussaini M. (2021) An open approach to systematically prioritize causal variants and genes at all published human GWAS trait-associated loci. *Nature Genetics*, 53(11): 1527-1533. DOI: 10.1038/s41588-021-00945-5

Mühr, L. S. A., Guerendiain, D., Cuschieri, K., Sundström, K. (2021). Human Papillomavirus Detection by Whole-Genome Next-Generation Sequencing: Importance of validation and quality assurance procedures. *Viruses*. 13(7): 1323. DOI: 10.3390/v13071323

Núñez-Vera, V., Garcia-Perla-Garcia, A., Gonzalez-Cardero, E., Esteban, F., Infante-Cossio, P. (2024). Impact of Treatment on Quality of Life in Oropharyngeal Cancer Survivors: A 3-Year Prospective Study. *Cancers*, 16(15), 2724–2724. DOI: 10.3390/cancers16152724

Okbay, A., Wu, Y., Wang, N., Jayashankar, H., Bennett, M., Nehzati, S. M., Sidorenko, J., Kweon, H., Goldman, G., Gjorgjieva, T., Jiang, Y., Hicks, B., Tian, C., Hinds, D. A., Ahlsgog, R., Magnusson, P. K. E., Oskarsson, S., Hayward, C., Campbell, A., ... Young, A. I. (2022). Polygenic prediction of educational attainment within and between families from genome-wide association analyses in 3 million individuals. *Nature Genetics*, 54(4), 437–449. DOI: 10.1038/s41588-022-01016-z

- Penisson-Besnier, I., Allamand, V., Beurrier, P., Martin, L., Schalkwijk, J., van Vlijmen-Willems, I., Gartioux, C., Malfait, F., Syx, D., Macchi, L., Marcorelles, P., Arbeille, B., Croue, A., De Paepe, A., Dubas, F. (2013). Compound heterozygous mutations of the TNXB gene cause primary myopathy. *Neuromuscular Disorders*, 23(8), 664–669. DOI: 10.1016/j.nmd.2013.04.009
- Petitjean, A., Achatz, M. I. W., Borresen-Dale, A. L., Hainaut, P., Olivier, M. (2007). TP53 mutations in human cancers: functional selection and impact on cancer prognosis and outcomes. *Oncogene*, 26(15), 2157–2165. DOI: 10.1038/sj.onc.1210302
- Petry, K. U., Luyten, A., Justus, A., Iftner, A., Strehlke, S., Schulze-Rath, R., Iftner, T. (2012). Prevalence of low-risk HPV types and genital warts in women born 1988/89 or 1983/84 -results of WOLVES, a population-based epidemiological study in Wolfsburg, Germany. *BMC Infectious Diseases*, 12(1). DOI: 10.1186/1471-2334-12-367
- Pirinen, M., Donnelly, P., Spencer, C. C. A. (2012). Including known covariates can reduce power to detect genetic effects in case-control studies. *Nature Genetics*, 44(8), 848–851. DOI: 10.1038/ng.2346
- Pujol-Gualdo, N., Läll, K., Lepamets, M., Rossi, H.-R., Arffman, R. K., Piltonen, T. T., Mägi, R., Laisk, T. (2022). Advancing our understanding of genetic risk factors and potential personalized strategies for pelvic organ prolapse. *Nat Commun*, 13: 3584
- Ramqvist, T., Dalianis, T. (2010). Oropharyngeal Cancer Epidemic and Human Papillomavirus. *Emerging Infectious Diseases*, 16(11), 1671–1677. DOI: 10.3201/eid1611.100452
- Rashkin, S. R., Graff, R. E., Kachuri, L., ... Witte, J. S. (2020). Pan-cancer study detects genetic risk variants and shared genetic basis in two large cohorts. *Nature Communications*, 11(1), 4423. DOI: 10.1038/s41467-020-18246-6
- Redzic, N., Pereira, A. R., Menon, S., Bogers, J., Coppens, A., Kehoe, K., Vanden Broeck, D. (2023). Characterization of type-specific HPV prevalence in a population of persistent cutaneous warts in Flanders, Belgium. *Scientific Reports*, 13(1). DOI: 10.1038/s41598-023-44154-y

- Rübben, A., Kalka, K., Spelten, B., Grußendorf-Conen, E. (1997). Clinical features and age distribution of patients with HPV 2/27/57-induced common warts. *Archives of Dermatological Research*, 289(6), 337–340. DOI: 10.1007/s004030050201
- Sánchez-Barrientos, G., Vega-Memije, E., García-Corona, C., Cuevas-González, J. C., Zavaleta-Villa, B., Ibarra-Arce, A., Olivo-Díaz, A. (2021). Human Leukocyte Antigens -DQA1 and -DQB1 Alleles in Patients With Common Warts. *Cureus*. DOI: 10.7759/cureus.18933
- Sherry, S.T., Ward, M.H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. *Nucleic Acids Research*, 29(1), 308–311. DOI: 10.1093/nar/29.1.308
- Shete, S., Liu, H., Wang, J., Yu, R., Sturgis, E. M., Li, G., Dahlstrom, K. R., Liu, Z., Amos, C. I., Wei, Q. (2020). A Genome-Wide Association Study Identifies Two Novel Susceptible Regions for Squamous Cell Carcinoma of the Head and Neck. *Cancer Research*, 80(12), 2451–2460. DOI: 10.1158/0008-5472.CAN-19-2360
- Shi, Y., Li, L., Hu, Z., ... Huang, K. (2013). A genome-wide association study identifies two new cervical cancer susceptibility loci at 4q12 and 17q12. *Nature Genetics*, 45(8), 918–922. DOI: 10.1038/ng.2687
- Shim, T. N., Harwood, C. A., Marsh, S. G., ... Bunker, C. B. (2021). The prevalence of human leucocyte antigen and human papillomavirus DNA in penile intraepithelial neoplasia in England 2011-2012. *International Journal of STD & AIDS*, 32(5), 388–395. DOI: 10.1177/0956462420970727
- Spiess, P. E., Dhillon, J., Baumgarten, A. S., Johnstone, P. A., Giuliano, A. R. (2016). Pathophysiological basis of human papillomavirus in penile cancer: Key to prevention and delivery of more effective therapies. *CA a Cancer Journal for Clinicians*. 66(6): 481–495. DOI: 10.3322/caac.21354
- Stecca, C. E., Alt, M., Jiang, D. M., Chung, P., Crook, J. M., Kulkarni, G. S., Sridhar, S. S. (2021). Recent Advances in the Management of penile Cancer: A Contemporary Review of the literature. *Oncology and Therapy*, 9(1), 21–39. DOI: 10.1007/s40487-020-00135-z

- Stoehr, R., Weisser, R., Wendler, O., ... Hartmann, A. (2018). P53 Codon 72 Polymorphism and risk for squamous cell carcinoma of the penis: a Caucasian Case-Control Study. *Journal of Cancer*, 9(22), 4234–4241. DOI: 10.7150/jca.26050
- Sun, B. B., Maranville, J. C., Peters, J. E., ... Morrell, N. W. (2018). Genomic atlas of the human plasma proteome. *Nature*, 558(7708), 73–79. DOI: 10.1038/s41586-018-0175-2
- Zhang, F., Lupski, J. R. (2015). Non-coding genetic variants in human disease: Figure 1. *Human Molecular Genetics*, 24(R1), R102–R110. DOI: 10.1093/hmg/ddv259
- Zhu, Z., Guo, Y., Shi, H., ... Liang, L. (2020). Shared genetic and experimental links between obesity-related traits and asthma subtypes in UK Biobank. *Journal of Allergy and Clinical Immunology*, 145(2), 537–549. DOI: 10.1016/j.jaci.2019.09.035
- Zhu, Z., Zhu, X., Liu, C. L., Shi, H., Shen, S., Yang, Y., Hasegawa, K., Camargo, C. A., Liang, L. (2019). Shared genetics of asthma and mental health disorders: a large-scale genome-wide cross-trait analysis. *European Respiratory Journal*, 54(6), 1901507. DOI: 10.1183/13993003.01507-2019
- Tao, Y., Sturgis, E. M., Huang, Z., Wang, Y., Wei, P., Wang, J. R., Wei, Q., Li, G. (2018). TGFβ1 Genetic Variants Predict Clinical Outcomes of HPV-Positive Oropharyngeal Cancer Patients after Definitive Radiotherapy. *Clinical Cancer Research : An Official Journal of the American Association for Cancer Research*, 24(9), 2225–2233. DOI: 10.1158/1078-0432.CCR-17-1904
- Terviseamet (2024). Eestis haigestub igal aastal HPV põhjustatud vähki enam kui 300 inimest. <https://rb.gy/ypdxch>
- Tian, C., Hromatka, B. S., Kiefer, A. K., Eriksson, N., Noble, S. M., Tung, J. Y., Hinds, D. A. (2017). Genome-wide association and HLA region fine-mapping studies identify susceptibility loci for multiple common infections. *Nature Communications*, 8(1). DOI: 10.1038/s41467-017-00257-5
- Uffelmann, E., Huang, Q. Q., Munung, N. S., de Vries, J., Okada, Y., Martin, A. R., Martin, H. C., Lappalainen, T., Posthuma, D. (2021). Genome-wide association studies. *Nat Rev Methods Primers*, 1(1): 1–21.
- Uusküla, A., Reile, R., Rezeberga, D., Karnite, A., Logminiene, Z., Padaiga, Ž., Nygård, M. (2014). The prevalence of genital warts in the Baltic countries: findings from national

cross-sectional surveys in Estonia, Latvia and Lithuania. *Sexually Transmitted Infections*, 91(1), 55–60. DOI: 0.1136/sextrans-2014-051540

van Haalen, F.M., Bruggink, S.C., Gussekloo, J., Assendelft, W.J., Eekhof, J.A. (2009). Warts in primary schoolchildren: prevalence and relation with environmental factors. *British Journal of Dermatology*, 161(1):148–152. DOI: 10.1111/j.1365-2133.2009.09160.x

Wakabayashi, R., Nakahama, Y., Nguyen, V., Espinoza, J. L. (2019). The Host-Microbe Interplay in Human Papillomavirus-Induced Carcinogenesis. *Microorganisms*, 7(7), 199. DOI: 10.3390/microorganisms7070199

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

Williams, H., Pottier, A., Strachan, D. (1993). The descriptive epidemiology of warts in British schoolchildren. *British Journal of Dermatology*, 128(5), 504–511. DOI: 10.1111/j.1365-2133.1993.tb00226.x

Woodhall, S., Ramsey, T., Cai, C., Crouch, S., Jit, M., Birks, Y., Edmunds, W. J., Newton, R., Lacey, C. J. N. (2008). Estimation of the impact of genital warts on health-related quality of life. *Sexually Transmitted Infections*, 84(3), 161–166. DOI: 10.1136/sti.2007.029512

Yang, D., Zhang, J., Cui, X., Ma, J., Wang, C., Piao, H. (2022). Risk factors associated with human papillomavirus infection, cervical cancer, and precancerous lesions in Large-Scale Population screening. *Frontiers in Microbiology*, 13. DOI: 10.3389/fmicb.2022.914516

Yuan, H., Ma, H., Lu, F., Yuan, Z., Wang, R., Jiang, H., Hu, Z., Shen, H., & Chen, N. (2012). Genetic variants at 4q23 and 12q24 are associated with head and neck cancer risk in China. *Molecular Carcinogenesis*, 52(S1), 2–9. DOI: 10.1002/mc.21929

KASUTATUD VEEBIAADRESSID

Accelerating Medicines Partnership (AMP) (i.a). Understanding genetic association data. Kasutatud 27.04.2025, <https://hugeamp.org/help.html?page=1036>

American Cancer Society (i.a). Risk Factors for Anal Cancer. Kasutatud 30.10.2024, <https://www.cancer.org/cancer/types/anal-cancer/causes-risks-prevention/risk-factors.html>

American Cancer Society (2018). Risk Factors for Penile Cancer. Kasutatud 13.03.2025, <https://www.cancer.org/cancer/types/penile-cancer/causes-risks-prevention/risk-factors.html>

Better Health Channel (BHC) (i.a). Warts. Kasutatud 16.02.2025, <https://www.betterhealth.vic.gov.au/health/conditionsandtreatments/warts#symptoms-of-warts>

Cancer Research UK (2023a). Risks and causes of penile cancer. Kasutatud 13.03.2025, <https://www.cancerresearchuk.org/about-cancer/penile-cancer/risks-causes>

Cancer Research UK (2023c). Symptoms of penile cancer. Kasutatud 13.03.2025, <https://www.cancerresearchuk.org/about-cancer/penile-cancer/symptoms>

Cancer Research UK (2023b). What is penile cancer? Kasutatud 13.03.2025, <https://www.cancerresearchuk.org/about-cancer/penile-cancer/what-is-penile-cancer>

CD Genomics (2023). Fact Sheet: Single Nucleotide Polymorphisms (SNPs). Kasutatud 16.04.2025, <https://www.cd-genomics.com/blog/fact-sheet-single-nucleotide-polymorphisms-snps/>

Cleveland Clinic (i.a; a). Genital Warts. Kasutatud 12.03.2025, <https://my.clevelandclinic.org/health/diseases/4209-genital-warts>

Cleveland Clinic (i.a; b). Oropharyngeal Cancer. Kasutatud 31.03.2025, <https://my.clevelandclinic.org/health/diseases/12180-oropharyngeal-cancer>

DermNet. (i.a). Viral wart. Kasutatud 30.10.2024, <https://dermnetnz.org/topics/viral-wart>

European Commission (i.a). Union Register of Medicinal Products for human use. Kasutatud 01.04.2025, <https://ec.europa.eu/health/documents/community-register/html/h080.htm>

FinnGen and mebaas (2023). Kasutatud 17.04.2025, <http://r12.finnngen.fi/>

Gene Cards (i.a; b). HLA-DQA1 Gene - Major Histocompatibility Complex, Class II, DQ Alpha 1. Kasutatud 12.04.2025, <https://www.genecards.org/cgi-bin/carddisp.pl?gene=HLA-DQA1>

Gene Cards (i.a; a). VARS1 Gene - Valyl-TRNA Synthetase 1. Kasutatud 12.04.2025, <https://www.genecards.org/cgi-bin/carddisp.pl?gene=VARS1>

Ida-Tallinna Keskhaigla (i.a). Tsöliaakia. Kasutatud 12.04.2025, <https://www.itk.ee/patsiendile/patsiendi-infomaterjalid/haigused/tsoliaakia>

- Inimene (i.a). Anküloseeriv spondüliit. Kasutatud 12.04.2025, <https://www.inimene.ee/haigused-ja-seisundid/list/haigused-ja-seisundid/ankuloseeriv-sponduliit>
- Kliinikum (i.a). Pea- ja kaelapiirkonna kasvaja. Kasutatud 16.05.2025, <https://www.kliinikum.ee/ho/info-haiguste-kohta/2-uncategorised/98-pea-ja-kaelapiirkonna-kasvaja>
- Learn.Genetic (i.a). Making SNPs Make Sense. Kasutatud 16.05.2025, <https://learn.genetics.utah.edu/content/precision/snips/>
- MedlinePlus Genetics (i.a; b). HLA-DQA1 gene. Kasutatud 12.04.2025, <https://medlineplus.gov/genetics/gene/hla-dqa1/>
- MedlinePlus Genetics (i.a; a). TNXB gene. Kasutatud 12.04.2025, <https://medlineplus.gov/genetics/gene/tnxb/>
- MOFFITT Cancer Center (i.a; a). Can HPV Cause Vulvar Cancer? Kasutatud 30.10.2024, <https://www.moffitt.org/cancers/vulvar-cancer/faqs/can-hpv-cause-vulvar-cancer/>
- MOFFITT Cancer Center (i.a; b). Penile Cancer Signs and Symptoms. Kasutatud 13.03.2025, <https://www.moffitt.org/cancers/penile-cancer/symptoms/>
- Mount Sinai Tisch Cancer Center (i.a). HPV and Cervical Cancer Myths vs. Facts. Kasutatud 31.03.2025, <https://www.mountsinai.org/care/cancer/services/gynecologic/conditions/cervical/myths-facts>
- NCBI (i.a; g). CDKN2B-AS1 CDKN2B and CDKN2A antisense cis and trans regulatory RNA 1 [Homo sapiens (human)]. Kasutatud 14.05.2025, <https://www.ncbi.nlm.nih.gov/gene/100048912>
- NCBI (i.a; h). GALNT14 polypeptide N-acetylgalactosaminyltransferase 14 [Homo sapiens (human)]. Kasutatud 14.05.2025, <https://www.ncbi.nlm.nih.gov/gene/79623>
- NCBI (i.a; c). HLA-DQA1 major histocompatibility complex, class II, DQ alpha 1 [Homo sapiens (human)]. Kasutatud 12.04.2025, <https://www.ncbi.nlm.nih.gov/gene/3117>
- NCBI (i.a; i). HLA-DQB1 major histocompatibility complex, class II, DQ beta 1 [Homo sapiens (human)]. Kasutatud 27.04.2025, <https://www.ncbi.nlm.nih.gov/gene/3119>

NCBI (i.a; e). LST1 leukocyte specific transcript 1 [Homo sapiens (human)]. Kasutatud 14.05.2025, <https://www.ncbi.nlm.nih.gov/gene/7940>

NCBI (i.a; f). LTA lymphotoxin alpha [Homo sapiens (human)]. Kasutatud 14.05.2025, <https://www.ncbi.nlm.nih.gov/gene/4049>

NCBI (i.a; d). TAP2 transporter 2, ATP binding cassette subfamily B member [Homo sapiens (human)]. Kasutatud 12.04.2025, <https://www.ncbi.nlm.nih.gov/gene/6891>

NCBI (i.a; a). TNXB tenascin XB [Homo sapiens (human)]. Kasutatud 12.04.2024, <https://www.ncbi.nlm.nih.gov/gene/7148>

NCBI (i.a; b). VARS1 valyl-tRNA synthetase 1 [Homo sapiens (human)]. Kasutatud 14.04.2025, <https://www.ncbi.nlm.nih.gov/gene/7407>

National Cancer Institute (2023). HPV and Cancer. Kasutatud 30.10.2024, <https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-and-cancer>

Open Targets Genetics (2022). Kasutatud september 2024 – mai 2025, <https://genetics.opentargets.org/>

Rahvusvahelise haiguste klassifikatsiooni 10. väljaanne (RHK-10) (2019). Kasutatud jaanuar 2024 – mai 2025, <https://icd.who.int/browse10/2019/en>

Rebrandly, URL shortener. Kasutatud 16.05.2025, <https://www.rebrandly.com/>

Tervise Arengu Instituut (TAI) (i.a). RHK ehk rahvusvaheline haiguste klassifikatsioon. Kasutatud 21.01.2025, <https://www.tai.ee/et/instituudist/meditsiini terminoloogia-kompetentsikeskus/who-klassifikaat-orid-rfk-ja-rhk/rhk-ehk>

Tervisestatistika ja terviseuuringute andmebaas (i.a). PK10: Pahaloomuliste kasvajate esmasjuhud paikme, soo ja vanuserühma järgi. Kasutatud 16.05.2025, https://statistika.tai.ee/pxweb/et/Andmebaas/Andmebaas_02Haigestumus_04PahaloomulisedKasvajad/PK10.px/

Urology Care Foundation (2024). What is Penile Cancer? Kasutatud 13.03.2025, <https://www.urologyhealth.org/urology-a-z/p/penile-cancer>

vaktsineeri.ee (i.a). HPV - inimese papilloomiviirus. Kasutatud 05.05.2025,
<https://vaktsineeri.ee/et/hpv>

WordlGODay (i.a). Vaginal Cancer. Kasutatud 16.05.2025, <https://rb.gy/qs8f8k>

LISAD

Lisa 1.

Tabel 5. Inimese papilloomiviirusega seotud genitaalpiirkonna- ning suu- ja neeluvähi seos varasemates uuringutes kirjeldatud variantidega.

Diagnoos	Geneetiline variant	Seoseanalüüsi P-väärtus
Häbemevähk	rs25882 rs27438	0.201922 0.278152
	rs2569254 rs3181225	0.411101 0.415331
	rs2256965 rs2239704	0.0152918 0.00649397
	rs3181225 rs3181224 rs3212227	0.415331 0.771307 0.164613
Peenisevähk	rs1042522	0.787572
Suuvähk	rs6547741 rs928674 rs8181047 rs10462706	0.845284 0.420689 0.220429 0.56169
Suu- ja neeluvähk kombineerituna	rs3828805 rs201982221 rs1453414	0.148628 - -
Neeluvähk	rs4318431 rs13211972 rs34518860	0.0589807 0.900517 0.251183

LIHTLITSENTS

Mina, Steffi Salu,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose

Inimese papilloomiviirusega seotud diagnooside geneetilised riskifaktorid ,

mille juhendajad on Triin Laisk ja Mariann Võsa,

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 Commons'i 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.

Steffi Salu

23.05.2025