TARTU ÜLIKOOL

LOODUS- JA TÄPPISTEADUSTE VALDKOND

MOLEKULAAR- JA RAKUBIOLOOGIA INSTITUUT

BIOINFORMAATIKA ÕPPETOOL

MAAELU TEADMUSKESKUS

TAIMETERVISE JA MIKROBIOLOOGIA VALDKOND

Kõrge läbilaskevõimega sekveneerimise rakendamine viiruste tuvastamiseks Eesti seemnekartulis ja kartuli Y-viiruse tüvede identifitseerimine bioinformaatiliste meetoditega

Magistritöö

30 EAP

Anette Stražev

Juhendajad MSc Piret van der Sman

MSc Tarmo Puurand

TARTU 2025

INFOLEHT

Kõrge läbilaskevõimega sekveneerimise rakendamine viiruste tuvastamiseks Eesti seemnekartulis ja kartuli Y-viiruse tüvede identifitseerimine bioinformaatiliste meetoditega

Lühikokkuvõte: Peamine saagikust ja saagi kvaliteeti mõjutavad tegurid on taimehaigused. Märkmismväärseid saagikadusid ja mugulate kvaliteedi langust põhjustab kartuli Y-viirus (PVY) mistõttu on selle ennetamine ja varajane tuvastamine äärmiselt oluline. Magistritöö eesmärk oli iseloomustada Eestis esinevaid PVY tüvesid, hinnata nende geneetilist mitmekesisust ja viia läbi fülogeneetiline analüüs ning tuvastada kaasnevaid RNA-viirusi RNA-sekveneerimise ja bioinformaatika abil. Tuvastati rekombinantsed tüved $PVY^{NTN(a)}$ ja PVY^{N-Wi}. Fülogeneetiline analüüs näitas, et Eesti PVY isolaadid kuuluvad kahte erinevasse klastritesse koos teiste Euroopa, Lähis-Ida ja Aasia päritolu järjestustega. Lisaks PVY'le tuvastati mitmeid teisi viiruse järjestusi, kuid nende esinemine vajab täiendavat tõlgendust.

CERCS teaduseriala: B230 Mikrobioloogia, bakterioloogia, viroloogia, mükoloogia
Märksõnad: kartuli Y-viirus, genoomijärjestus, rekombinantsed tüved, fülogeneetiline puu, taksonoomiline klassifitseerimine

Application of High-Throughput Sequencing for Virus Detection in Estonian Seed Potato and Identification of Potato Virus Y Strains Using Bioinformatic Methods

Abstact: Plant diseases are among the main factors affecting crop yield and quality. Potato virus Y (PVY) causes significant yield losses and reduces tuber quality, making its prevention and early detection critically important. The aim of this thesis was to characterize PVY strains present in Estonia, assess their genetic diversity and phylogenetic origin, and detect co-occurring RNA viruses using RNA sequencing and bioinformatics. Recombinant strains $PVY^{NTN(a)}$ and PVY^{N-Wi} were identified. Phylogenetic analysis showed that Estonian PVY isolates belong to two different clusters together with sequences from Europe, the Middle East and Asia. In addition to PVY, several other viral sequences were detected, although their biological relevance requires further interpretation.

CERCS classification: B240 Microbiology, bacteriology, virology, mycology
Keywords: Potato virus Y, genome sequencing, recombinant strains, phylogenetic tree, taxonomic classification

SISUKORD

KASUTATUD LÜHENDID 5

SISSEJUHATUS 7

1. KIRJANDUSE ÜLEVAADE 9

1.1. Kartul 9

1.1.1. Kartuli tähtsus maailmas 9

1.1.2. Kartuli tootmine maailmas ja Eestis 9

1.2. Taimeviirused 10

1.3. Kartulit nakatavad viirushaigused 12

1.4. Kartuliviiruste tuvastamise meetodid 16

1.5. Sekveneerimine 18

1.5.1. Sekveneerimismeetodid 18

1.5.2. RNA sekveneerimise eripärad 20

1.6. Sekveneerimisandmete analüüs 20

1.6.1. *de novo* järjestamine 20

1.6.2. BLAST-analüüs 22

1.6.3. Järjestusandmete joondamine ja kaardistamine 22

1.6.4. Tüvede määramine metagenoomsetest andmetest 23

1.6.5. Taksonoomiline klassifitseerimine 23

1.6.6. Fülogeneetiliste puude koostamine ja analüüs 23

2. EKSPERIMENTAALNE OSA 25

2.1. Töö eesmärgid 25

2.2. Materjal ja metoodika 25

2.2.1. Kartuli proovid 27

2.1.1. Kartuliviiruste tuvastamise meetodid 28

2.1.2. TaqMan Reaalaja pöördtranskriptsiooni polümeraasi ahelreaktsioon 28

2.2.2. Kartuli Y-viiruse positiivsete proovide sekveneerimine 30

2.2.3. *de novo* järjestuse kokkupanek 30

2.2.4. BLAST-analüüs ja referentsgenoomi valik 31
2.2.5. Järjestusandmete kaardistamine referentsgenoomile 32
2.2.6. Kartuli Y-viiruse tüvede määramine 32
2.2.7. Kartuli Y-viiruse fülogeneetiline analüüs 33
2.2.8. Taksonoomiline klassifitseerimine RNA-viiruste tuvastamiseks 34
2.3. Tulemused 34
2.3.1. Eestis esinevad kartuli Y-viiruse tüved 34
2.3.2. Kartuli Y-viiruse tüvede fülogeneetiline analüüs 40
2.3.3. Teised kartulit nakatavad taimeviirused Eestis 42
2.4. Arutelu 44
KOKKUVÕTE 49
SUMMARY 51
TÄNUSÕNAD 53
KASUTATUD KIRJANDUS 54
LISAD 63
Lisa 1. PVY tuvastamine ELISA meetodil 64
Lisa 2. RNA ekstraktsioon Hughes ja Galau meetodil 66
Lisa 3. GenomeTester4 *gmer_counter* andmebaasi faili struktuur 67
Lisa 4. Fülogeneetiline puu koos harupikkustega 68
Lisa 5. Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks 69

KASUTATUD LÜHENDID

BLAST – *Basic Local Alignment Search Tool*

Ct – tsükli tundlikkuslävi

CTAB – tsetüültrimetüülammooniumbromiid

ddNTP – didesoksünukleotiid

DNA – desoksüribonukleiinhape

dsDNA – kaheahelaline desoksüribonukleiinhape

dsRNA – kaheahelaline ribonukleiinhape

ELISA – immunoensüümmeetod

FERA – The Food and Environment Research Agency

H&G ekstraktsioon – Hughes & Galau ekstraktsioon

METK – Maaelu Teadmuskeskus

mRNA – informatsiooni-RNA

HTS – kõrge läbilaskevõimega sekveneerimine

PLRV – kartuli keerdlehisuse viirus

PSTVd – kartuli värtnakujulisuse viroid

PVA – kartuli A-viirus

PVM – kartuli M-viirus

PVS – kartuli S-viirus

PVX – kartuli X-viirus

PVY – kartuli Y-viirus

RNA – ribonukleiinhape

RNA-seq – ribonukleiinhape sekveneerimine

rRNA – ribosomaalne RNA

RT – pöördtranskriptsioon

RT-qPCR – pöördtranskriptsioon polümeraasi ahelreaktsioon

SBS – s*equencing-by-synthesis*

SDS – naatrium dodetsüülsulfaat

ssDNA – üksikahelaline desoksüribonukleiinhape

ssRNA (–) – negatiivse polaarsusega üksikahelaline ribonukleiinhape

ssRNA (+) – positiivse polaarsusega üksikahelaline ribonukleiinhape

TGS – kolmanda põlvkonna sekveneerimine

TTMV – Taimetervise ja mikrobioloogia valdkond

SISSEJUHATUS

Kartul on maavitsaliste (*Solanaceae*) sugukonda ning maavitsa (*Solanum*) perekonda kuuluv toidukultuur, kuhu lisaks kuuluvad ka teised tähtsad kultuurid nagu tomat, paprika, baklažaan, tubakas ja vähetähtsam kultuur nagu melon (Loebenstein jt 2001: 20). Kartul on tähtsuselt kolmas toiduaine maailmas, sest sel on võime toota rohkem energiat ja valku pindalaühiku kohta kui ühelgi teisel üksikul toidukultuuril (Birch jt., 2012; Bajaj, 1987).

Taimehaiguseid peetakse saagikuse ja saagi kvaliteedi vähenemise peamisteks faktoriteks. Viirustest tingitud taimehaiguste kohta teatakse kõige vähem ning nende tõrjumine on seetõttu ka kõige raskem (Salazar 1996: 11). Toidulaual on kartuli tähtsus langenud, kuid seda nakatavate taimehaiguste hulk suureneb, mille tõttu on nende ennetamine oluline. Ennetamiseks on kehtestatud Euroopa Liidu seadusest tuleneb supereliitseemnekartuli taimekahjustajate piirmäär. Supereliitseemnekartulis on viiruste piirmäär 0,5%. Lisaks ei tohi seemnekartulis esineda ohtlikke taimekahjustajaid nagu on kartuli värtnakujulisuse viroidi (Euroopa Liidu Teataja 2020, Taimede paljundamise ja sordikaitse seadus 2006, § 3 lg 1). Taimehaiguste seast üks enim mõjutatavaum patogeene on kartuli Y-viirus, mis võib hävitada kartuli saagi täielikult ja mõjutab oluliselt mugulate kvaliteeti (Tsedaley, 2015; Singh jt., 2013). Taimeviiruste, sealhulgas kartuli Y-viiruse (PVY) tuvastamiseks kasutatakse seroloogilisi ja molekulaarseid meetodeid nagu näiteks on ELISA (immunoensüümmeetod) ja reaalaja pöördtranskriptaasi polümeraasi ahelreaktsioon (reaalaja RT-qPCR) (MacKenzie jt., 2015). Lisaks on järjest enam kasutusel kõrge läbilaskevõimega sekveneerimine andes teavet organismi geneetilise ülesehituse kohta (Fitzpatrick jt., 2021; Shendure jt., 2017) ning võimaldades tuvastada seni tundmatuid haigustekitajaid.

Käesolev magistritöö jätkab varasemates uuringutes alustatud teemakäsitlust, keskendudes PVY viiruse variatsioonide detailsemale analüüsile. Eelnevalt on autor käsitlenud sarnast teemat oma bakalaureusetöös, kus keskenduti kartuliviiruse tuvastamisele kasutades erinevaid ekstraktsioonimeetodeid. Kõrge kvaliteediga ekstraktsioon on iga järgmise molekulaabrioloogilise analüüsi õnnestumise aluseks. Magistritöö eesmärk on suurendada seniste kasutusolevate analüüsimeetodite valikut viiruste tuvastamiseks seemnekartulist ning rakendada põhjalikumaid meetodeid andmete analüüsiks. Antud magistritöös uuriti Eestis esinevaid kartuli Y-viiruse tüvede järjestusi ning analüüsiti neid erinevate bioinformaatiliste meetoditega. Samuti uuriti, milliseid kartulit nakatavaid RNA viiruseid veel Eestis esineb. Töös kasutatud kartuliproovide toorandmete jaoks viidi läbi sekveneerimine Ilumina MiSeq platvormil. Saadud andmed töödeldi kasutades järgmiseid bioinformaatilisi tööriistu: Velvet,

BLAST, BWA, SAMtools, SAMtools consensus, MEGA11, RAxML, GenomeTester4, Kraken2. Uurimises kasutatud taimne materjal pärineb Maaelu Teadmuskeskuses Taimetervise ja mikrobioloogia valdkonna laboratooriumis analüüsitud PVY positiivsetest kartuliproovidest aastatel 2011, 2012 ja 2019. Analüüsitud seemnekartuli sordid olid „Lucilla“, „Arielle“, „Red Lady“, „Tiina“, „Albertine“, „Teele“, „Milva“, „Manitou“ ja „Anuschka“.

1. KIRJANDUSE ÜLEVAADE

1.1. Kartul

1.1.1. Kartuli tähtsus maailmas

Kartulil (*Solanum tuberosum*) on võrreldes teiste maailmas olevate kultuurtaimedega suurim metsikute esivanemate hulk. Kartuli ajalugu ulatub vähemalt 14 000 aasta taha, kui Lõuna-Ameerikas elanud jahimehed ja korilased toitusid looduslikest taimedest, mille seas oli ka 236 erinevat kartuliliiki. Kõige olulisemaks neist kujunes *Solanum tuberosum*, mille kodustasid Andide piirkonna põllumehed umbes 10 000 aastat enne Kristust. Looduslike tingimuste ja geograafiliste eripärade tõttu sai kartulist seal kiiresti peamine toiduaine. Katse-eksitus meetodiga avastati, et kartulit on võimalik paljundada nii seemnete kui ka mugulate istutamise kaudu (Smith, 2011).

Esmakordselt jõudis kartul Euroopasse 1567. aastal (Hispaaniasse), Gran Canariale, kust see hiljem 1574. aastal jõudis Põhja-Prantsusmaale, täpsemalt Rouen'i (Hawkes ja Francisco-Ortega, 1993). Eestis hakati kartulit kasvatama 18. sajandi lõpul, esmalt mõisates, kust see järk-järgult talupõldudele levis (EE 1989 s. v. kartul).

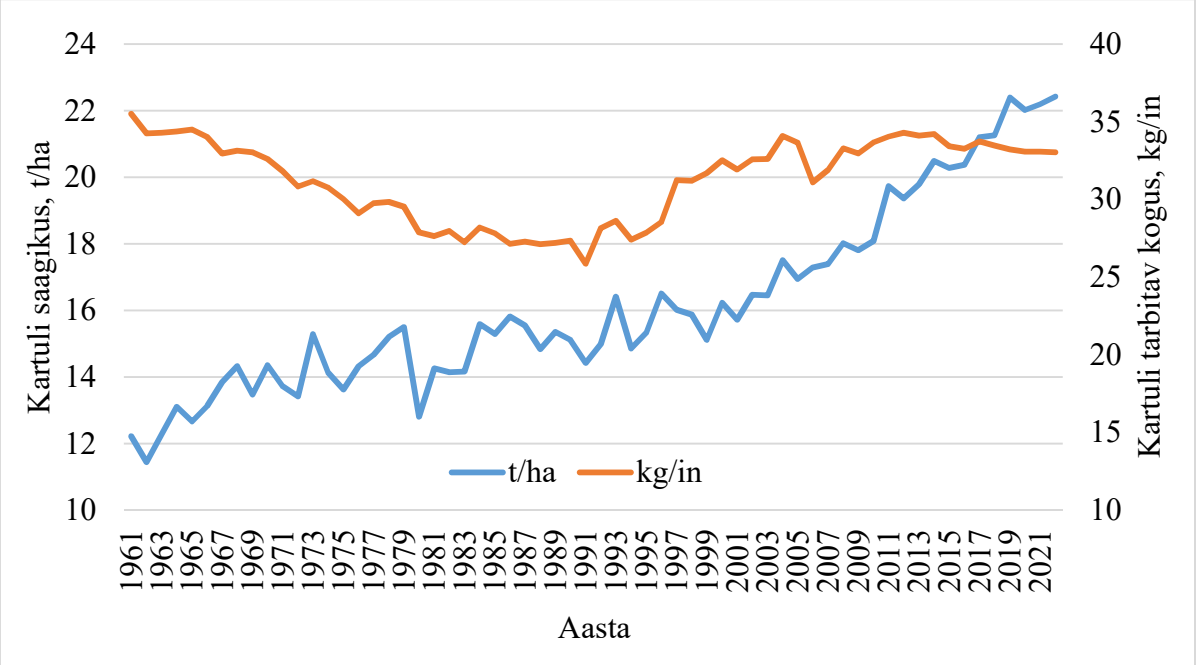
Tänapäeval kasvatatakse kartulit maailmas mitmesugustel eesmärkidel, sõltuvalt piirkonna, riigi arengu ja ajalooliste tegurite mõjust. Kartulit kasutatakse peamiselt toiduks, loomasöödaks, tööstuslikuks töötlemiseks ja seemnemugulate tootmiseks (Ortiz ja Mares, 2017). Kartul on väärtuslik toitaineallikas, sisaldades kiudaineid, kaaliumit, magneesiumi, C- ja B6-vitamiine, samuti antioksüdante (Beals, 2019).

1.1.2. Kartuli tootmine maailmas ja Eestis

Kartul on üks olulisemaid põllukultuure maailmas, olles eriti sobilik kasvatamiseks regioonides, kus maaressursid on piiratud, kuid elanikkond tihe. Kartulil on võime anda oluliselt suuremat saaki pindala- ja ajaühiku kohta kui näiteks levinumad teraviljad nagu nisu, riis või mais (Raigond jt., 2020).

Ajavahemikus 1961-2021 on maailma kartulisaagikus näidanud selget kasvutrendi, suurenedes ligikaudu 11 tonnilt hektari kohta enam kui 22 tonnini (Joonis 1). Selline kahekordne tõus viitab edusammudele sordiaretuses, agrotehnikas ja põllumajandustehnoloogias. Samal ajal on kartuli tarbimine inimese kohta olnud märksa kõikuvam. Kui 1960. aastate alguses tarbiti maailmas keskmiselt üle 35 kg kartulit inimese kohta aastas, siis järgnevate aastakümnete jooksul langes

see näitaja märkimisväärselt, saavutades madalseisu 1990. aastate alguseks. Alates 1990ndate keskpaigast on tarbimine siiski taas mõõdukalt tõusnud, stabiliseerudes vahemikus 33–36 kg/in aastas (FAOSTAT 2025).



Joonis 1. Maailma kartuli saagikus (t/ha) ja tarbitav kogus inimesed kohta (kg/in) (FAOSTAT 2025).

2023. aastal ulatus maailma kartulisaak ligikaudu 383 miljoni tonnini, kartulit kasvatati kokku 16 miljonil hektaril. Suurimateks tootjariikideks olid Hiina (93 miljonit tonni), India (60 miljonit tonni) ning Ukraina (21 miljonit tonni). Enim toodeti kartulit Aasias, kus kogutoodangu osakaal moodustas 53,4%. Euroopa andis 26,1% kogu maailmatoodangust, samas kui Ameerika, Aafrika ja Okeaania tootsid vastavalt 12,3%, 7,8% ja 0,5% ulatuses (FAOSTAT, 2025).

Maailmaga võrreldes oli Eestis kasvatatava kartuli saagikus 2021. aastal mõnevõrra väiksem. 2021. aastal oli kartuli keskmine saagikus Eestis 21 t/ha ning kasvupind oli 3370 hektarit. Aastal 2022 tõusis kartuli saagikus Eestis 23 tonnini hektari kohta ning samuti suurenes ka kasvupind 3381. hektarini. 2023. aastal suurenes saagikus ja kasvupind veelgi (24 t/ha, 3457 ha). 2024. aastal saagikus vähenes 24-lt tonnilt hektaril 23-le tonnile hektarile. Kasvupinnaks oli 3176 hektarit (Statistikaamet PM0281).

1.2. Taimeviirused

Viirused on mikroskoopilised nakkustekitajad, mis suudavad nakatada väga erinevaid elusorganisme, sealhulgas baktereid, taimi ja loomi. Viirused on ehituselt lihtsa struktuuriga, need koosnevad ühest või mitmest nukleiinhappe molekulist, mis on ümbritsetud valkudest või

lipoproteiinidest koosneva kapsiidiga (Hull, 2004). Viiruseid liigitatakse nende geneetilise materjali tüübi põhjal: eristatakse DNA (desoksüribonukleiinhape) ja RNA (ribonukleiinhape) viiruseid, kuid viirus sisaldab alati ainult ühte neist nukleiinhappetüüpidest (Loebenstein jt., 2001; Smith, 1979). Klassifikatsioonisüsteem eristab järgmisi rühmi: kaheahelaline DNA (dsDNA), üksikahelaline DNA (ssDNA), pöördtranskriptsiooniga dsDNA (dsDNA (RT)), kaheahelaline RNA (dsRNA), negatiivse polaarsusega ssRNA (ssRNA (−)), positiivse polaarsusega ssRNA (ssRNA (+)) ning pöördtranskriptsiooniga ssRNA (ssRNA (RT)). Kaheahelalised ja üksikahelalised DNA-viirused sisaldavad vastavalt kahte või ühte desoksüribonukleiinhappe ahelat, RNA-viirused aga ribonukleiinhappe molekule. Pöördtranskriptsiooniga viirused (RT) on võimelised oma RNA põhise genoomi transkripteerima DNA-ks, mis võimaldab neil integreeruda peremeesorganismi genoomi (Fauquet jt., 2005).

Viiruste paljunemine on võimalik vaid siis, kui nad satuvad sobivatesse peremeesrakkudesse (Hull, 2004). Taimedes saavad viirused liikuda peremeesrakkudesse üksnes passiivselt, näiteks väikeste vigastuste kaudu või viirust kandvate vektorite, nagu putukate abil (Takacs jt., 2014). Replikatsioon algab kui viirus nakatab peremeestaime ning siseneb peremeesrakku läbides rakuseina ja -membraani (Loebenstein jt., 2001).

Viirused mõjutavad negatiivselt taime tervist ja produktiivsust, vähendades taime saagikust ja elujõudu. Viirushaiguste avastamine on keeruline, kuna erinevad viirused ja nende tüved võivad tekitada väga erinevaid sümptomeid (Hooker, 1981). Levinud viirusnakkuse tunnusteks on näiteks lehtede laiksus, mosaiiksus, rõngaslaiksus, kudede kärbumine, nekroos, triibulised mustrid, närbumine, keerdlehisus või leherootsu kahvatumine (Scholthof jt., 2011). Viirushaigused võivad esineda ka latentsel kujul, see tähendab, et nakatunud taimel ei pruugi silmaga nähtavaid sümptomeid esineda. Iga viirusinfektsioon algab latentse faasiga, kus viirus üldjuhul ei ole nakkav, kuid teatud juhtudel võib leviku oht ilmneda ka enne sümptomite välja kujunemist (Chan ja Jeger, 1994).

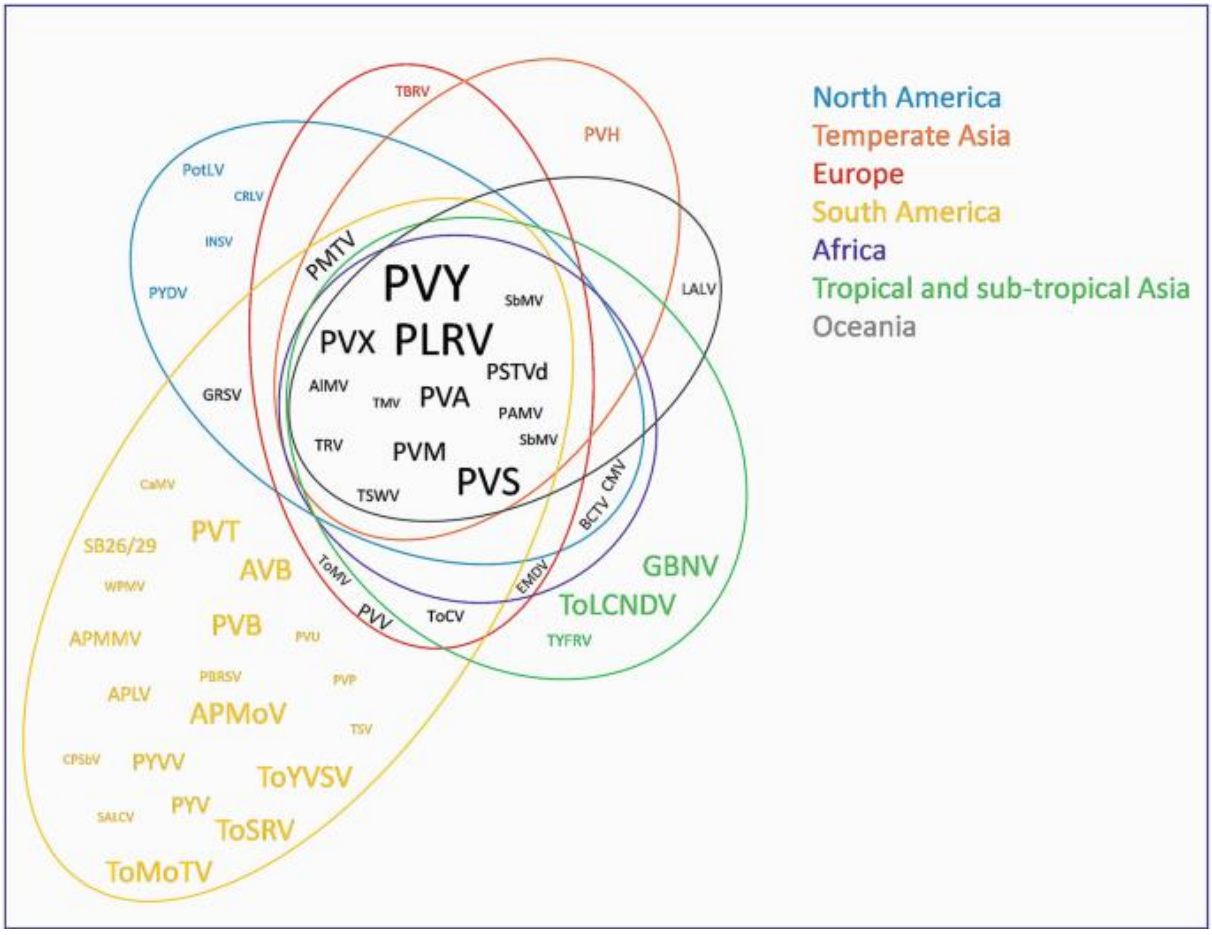
Enamasti toimub taimeviiruste edasikandumine selgrootute vektorite abil (Scholthof, 2005), looduslikes tingimustes võivad viirusekandjateks olla mitmed organismirühmad, sealhulgas seened, nematoodid, lehetäid, mardikad, ämblikulaadsed ja karilased (Singh, 2020). Viiruse ellujäämise ja leviku seisukohalt on oluline, et see suudaks nakatunud taimedelt liikuda edasi tervele taimele. Selline levik võib toimuda kas mehaaniliselt või vektororganismide abil. Ligikaudu 70% teadaolevatest enam kui 700 taimeviirusest levib edasi lülijalgsete, nematoodide või seente vahendusel. Kõige mõjukamad vektorid on lehetäid, kes on seotud üle

380 erineva taimeviiruse edasikandmisega (Emden ja Harrington, 2017). Lisaks bioloogilistele vektoritele kanduvad viirused edasi ka vegetatiivselt, näiteks mugulate, sibulate, võsundite, risoomide ja pungade kaudu. See on eriti oluline vegetatiivselt paljundatavates kultuurides nagu kartul, banaan, suhkruroog, bataat ja peet (Sastry 2013). Mehaaniline levik võib toimuda ka inimtegevuse kaudu, näiteks nakatunud tööriistadega või isegi mäletsejaliste kaudu, kes viiruseosakesi edasi kannavad (Pagan, 2022).

1.3. Kartulit nakatavad viirushaigused

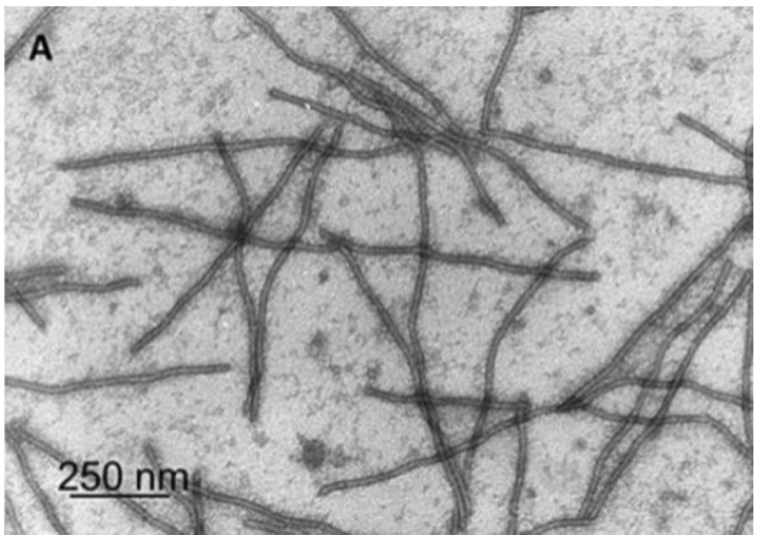
Kartul on viirushaiguste suhtes vastuvõtlik ning looduslikult suudab nakatada kartulit rohkem kui 40 erinevat viirust (Onditi jt., 2021). Nende viiruste levimus, kahjulikkus ja geograafiline jaotus varieeruvad sõltuvalt viiruse levikut soodustavatest teguritest, nagu kliima, vektororganismide olemasolu ning kartulitaimede käsitlemisviisid (Lacomme, 2017). Ohtlikumad kartuliviirused on näiteks kartuli X-viirus (PVX), kartuli Y-viirus (PVY) ja kartuli keerdlehisuse viirus (PLRV) (Onditi jt 2021). Taimehaigused on kartulikasvatuses eriti ohtlikud – näiteks kartuli Y-viirus võib põhjustada saagikadu, mis varieerub 10-100% vahel sõltuvalt kartulisordist ja nakkuse raskusastmest. (Tsedaley, 2015). Eestis testitakse Super Eliit seemnekartulis kuue kartuliviiruse olemasolu: kartuli A-viirus (PVA), kartuli Y-viirus (PVY), kartuli X-viirus (PVX), kartuli S-viirus (PVS), kartuli M-viirus (PVM) ja kartuli keerdlehisuse viirus (PLRV), madalamates seemneklassides ainult PVY (Taimekaitseseadus 2004, § 12 lg 2). Lisaks määratakse Eestis ka kartuli värtnakujulisuse viroidi (PSTVd) olemasolu (Taimede paljundamise ja sordikaitse seadus 2006, § 3 lg 1).

Joonisel 2 on kujutatud VENN'i diagramm, mis annab ülevaate kartuliviiruste georgraafilisest levikust maailmas. Diagrammi keskmes paiknevad viirused, mille levik kattub kõigi piirkondade vahel. Näiteks kartuli Y-viirus (PVY), kartuli keerdlehisuse viirus (PLRV), kartuli S-viirus (PVS), kartuli M-viirus (PVM) ja kartuli X-viirus (PVX) on esindatud diagrammi keskosas, mis viitab sellele, et need on globaalselt levinud. Äärmistes piirkondades paiknevad viirused esinevad valdavalt lokaalselt ning nende geograafiline levik on piiratud. Näiteks GBNV (*Groundnut bud necrosis virus*) ja ToLCNDV (tomati keerdlehisuse New Delhi viirus) esinevad peamiselt troopilises ja subtroopilises Aasias, samas kui ToRSV (*Tomato ringspot virus*) ja ApMoV (*Andean potato mottle virus*) on esindatud peamiselt Aafrikas (Kreuze jt., 2020).



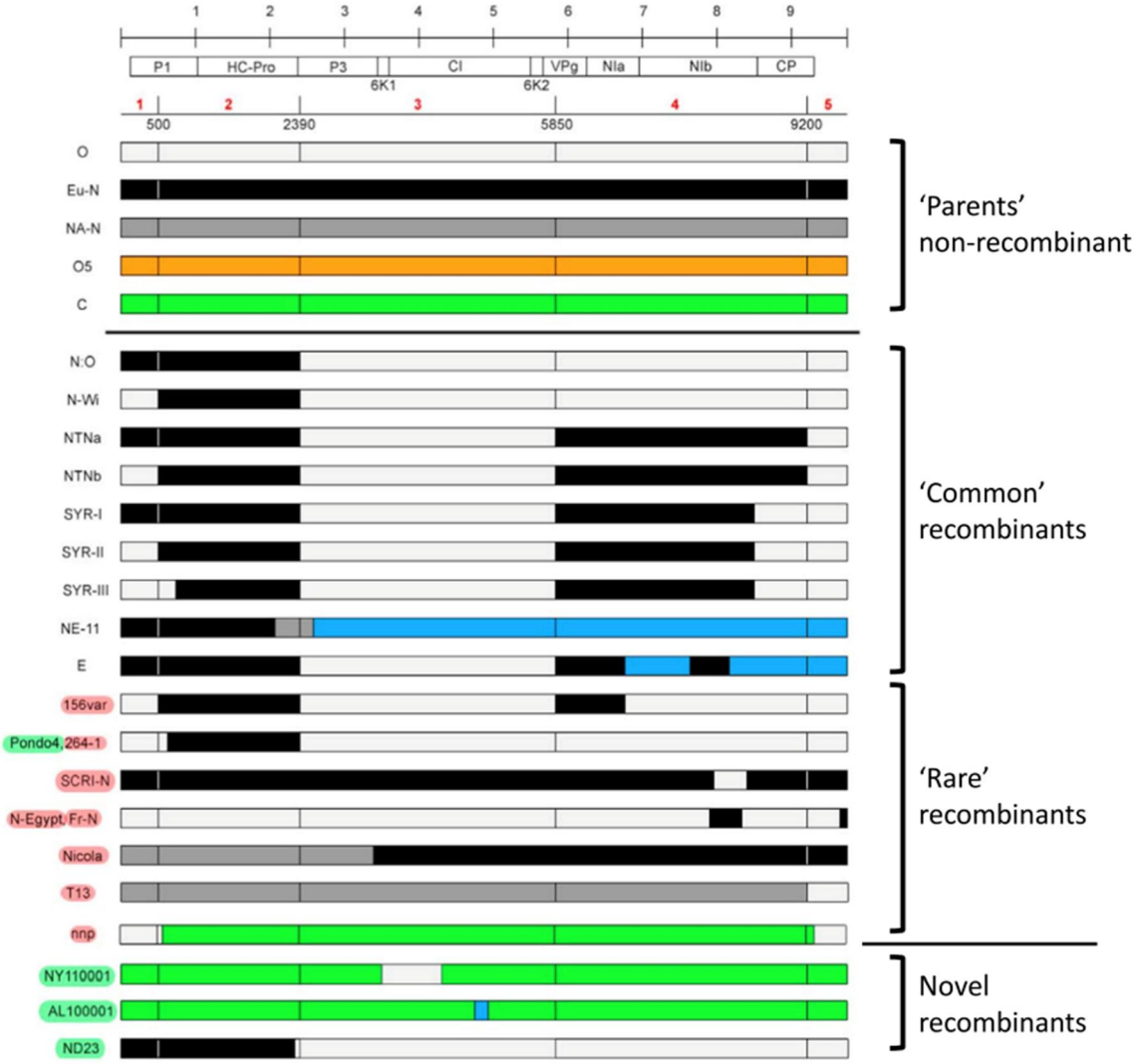
Joonis 2. VENN'i diagramm kujutatuna kartuliviiruste geograafilisest levikust maailmas (Kreuze jt., 2020).

Kartuli Y-viirus kuulub *Potyviridae* sugukonda ja on üks tõsisemaid viiruspatogeene *Solanaceae* sugukonna kultuurtaimedel, sealhulgas kartulil, tubakal ja paprikal. Viirus põhjustab tugevaid haigusnähte, vähendab saagikust ning levib kergesti. Esmakordselt kirjeldati kartuli Y-viirust 20. sajandi alguses, kui see avastati 1935. aastal tubakataimelt. Hiljem leiti see ka kartulitaimedelt Peruus ja Boliivias 1940. aastate alguses. 1950. aastatel põhjustasid need viirust kandvad kartulid Euroopas mitmeid epideemiaid. Lisaks kartulile võib PVY nakatada ka tomatit ja paprikat (Moury, 2010; Tsedaley, 2015). PVY viiruspartiklid on pikad, painduvad ning helikaalse ehitusega . Genoom koosneb umbes 9700 nukleotiidist ja on positiivse polaarsusega üksikahelaline RNA (ssRNA (+)). Joonisel 3 on toodud PVY virionide mikrofoto, mis on jäädvustatud elektronmikroskoobiga (Lacomme, 2017).



Joonis 3. Elektronmikroskoobi kartuli Y-viiruse osakeste (virionide) mikrofoto (Lacomme 2017).

Kartuli Y-viirus on kõrge mutatsiooni- ja rekombinatsioonisagedusega viirus, mille tulemusel on aja jooksul kirjeldatud arvukalt erinevaid tüvesid ja variante (Karasev & Gray, 2013; Lacomme jt, 2017; Green jt, 2017, 2018; Glasa jt, 2021; Samarskaya jt, 2024). Ajalooliselt on PVY tüvegrupid klassifitseeritud peamiselt peremeestaimede vastusreaktsioonide ja resistentsusgeenidega toimuvate interaktsioonide alusel (Singh jt, 2008; Lacomme jt, 2017). PVY algupärased klassikalised tüved hõlmavad PVY-O (tavaliselt mitte nekrootiline, enim levinud enne 1980ndaid) ja PVY-N (nekrootiline tüvi). PVY-O on seotud leebemate sümptomitega nagu mosaiiksus, samas kui PVY-N põhjustab nekroosi (Karasev ja Gray, 2013; Singh jt., 2008). Joonis 4 esitab skemaatilise ülevaate kartuli Y-viiruse rekombinantsetest struktuuridest. Ülal paiknev joonlaud kujutab PVY genoomi pikkusega ligikaudu 9,7 kilobaasi (kb), mille all on näidatud üksikud geenid ristkülikutena koos vastavate valgunimetustega (P1, HC-Pro, P3, CI, 6K1, 6K2, VPg, NIa, NIb ja CP). Joonisel on isolaadid jaotatud nelja kategooriasse. Esmalt on kujutatud „vanemlikud“ ehk mitte-rekombinantsed järjestused (Eu-N, NA-N ja O), mis on esitatud ülemise musta joone kohal ja värvitud erinevalt, tähistamaks eri päritolu järjestusi (Green jt., 2017). Üheksa rekombinanti (N:O, N-Wi, NTNa, NTNb, NE11, E, SYR-I, SYR-II ja SYR-III) on laialdaselt levinud ja esinevad mitmes geograafilises piirkonnas. Lisaks on olemas seitse haruldast rekombinanti, mida on tuvastatud vaid üksikutes proovides, näiteks $^{PVYN\text{-}Wi}$-156var, PVY$^{N\text{-}Wi}$-261-4, PVY-SCRI-N, PVY-FrN, PVY-Nicola, PVY-T13 ja PVY-nnp (Lorenzen jt., 2006; Chikh Ali jt., 2007; Schubert jt., 2007; Hu jt., 2009; Galvino-Costa jt., 2012a; Galvino-Costa jt., 2012b; Karasev ja Gray, 2013; Lorenzen jt., 2008; Chikh Ali jt., 2010); Ogawa jt., 2008; Ogawa jt., 2012).



Joonis 4. Varem tuntud ja avaldatud kartuli Y-viiruse 16 rekombinantstruktuuri (Green jt., 2017).

Joonisel on punaste numbritega tähistatud viis peamist rekombinatsioonipiirkonda, mis põhinevad sagedastel murdepunktidel ning võimaldavad jälgida rekombinatsioonide korduvust erinevates isolaatides. Selline genoomiline plastilisus ja rekombinatsioonide rohkus viitab PVY kõrgele evolutsioonilisele potentsiaalile, kohanemisvõimele ning võimalikele väljakutsetele haigustõrjes (Green jt., 2017).

Kartuli Y-viirus põhjustab mitmesuguseid haigustunnuseid, mille väljendumine sõltub viiruslikust tüvest, peremeestaimest ning kasvukeskkonna tingimustest. Samuti mõjutab sümptomite avaldumist see, kas nakatumine toimub primaarse või sekundaarse nakkuse teel. Primaarne nakkus toimub vektorite, nagu on lehetäid, kaudu. Sel juhul võib täheldada taimede ülemistes osades kerget kuni tugevat mosaiiksust, mida seostatakse lehtede moondumisega ehk kortsumisega. Alumistel lehtedel võivad ilmneda kolletumise ja nekroosi tunnused. Sümptomiteks võivad olla ka varisevad või pudenevad keskmised lehed. Sekundaarne nakkus toimub juhul, kui istutusmaterjal ehk seemnemugul on viirusega juba eelnevalt nakatunud.

Sellised taimed jäävad kasvu poolest kääbuseks, on habraste kortsus lehtedega ning mugulatel võib esineda nekroosi ja rõngaslaiksust (Scholthof jt., 2011).

PVY levib nii bioloogiliste kui mehaaniliste mehhanismide kaudu. Looduslikult kannavad viirust edasi enam kui 50 lehetäi liiki, kellest mitmed ei ole isegi püsivad kartulitaime kahjurid, vaid juhuslikud külastajad. Kartuli Y-viirus levib lehetäide vahendusel mitteperistentsel viisil, mis tähendab, et viirus suudab väga kiiresti liikuda ühelt taimelt teisele isegi siis, kui vektor taimel viibib vaid lühikest aega. Mehaaniline levik võib toimuda näiteks siis, kui üks taim on nakatunud ja kõrval olev taim saab vigastada, olgu see siis tuule, loomade või põllutöömasinate tõttu. Samuti võib viirus levida seemnekartuli mehaanilisel lõikamisel, mistõttu on oluline töövahendite põhjalik puhastamine ja desinfitseerimine enne uue seemnepartiiga töötamist (Tsedaley, 2015).

Kartuli Y-viiruse leviku tõkestamisel on üheks olulisemaks strateegiaks ennetusmeetmete rakendamine. Üks tõhusamaid viise nakkuse vältimiseks on kasutada viirusvaba seemnematerjali ja võimalusel taimekasvatuse ajastamine perioodile, mil lehetäide esinemine on minimaalne. Lisaks võib nakkusriski vähendada vältides viiruse suhtes vastuvõtlike taimeliikide kasvatamist üksteise läheduses. Tõhusaks meetmeks on seemnekartuli pealsete hävitamine enne kasvuperioodi lõppu, kuna see takistab viiruse edasikandumist järgmistesse taimedesse. Viiruse leviku piiramiseks kasutatakse ka mineraalõlidega pritsimist, mis vähendab lehetäide kaudu toimuva ülekande tõenäosust. Samuti on oluline roll resistentsete kartulisortide aretusel ja kasvatamisel (Brunt jt., 1996).

Lisaks agronoomilistele meetmetele on kehtestatud seaduslikud nõuded viiruse kontrolli all hoidmiseks. Näiteks nõuab seadus, et kui ettevõte kasvatab samaaegselt nii seemne- kui tarbekartulit, tuleb ka tarbekartuli tootmisel kasutada sertifitseeritud seemnekartulit. See aitab tagada, et kasutatav istutusmaterjal oleks võimalikult puhas viirusnakkustest (Taimede paljundamise ja sordikaitse seadus 2006, § 5 lg 5).

1.4. Kartuliviiruste tuvastamise meetodid

Kartuliviiruste tuvastamiseks kasutatakse traditsiooniliselt saagikoristusejärgsete mugulate testimist. Ajalooliselt kasutati taimede välja kasvatamist kuni kindla arengufaasini, et hinnata nende visuaalseid haigustunnuseid ja tuvastada viirust ELISA või RT-qPCR meetodil.

Immunoensüümmeetod ehk ELISA on laialdaselt kasutatav meetod taimset päritolu viiruspatogeenide tuvastamiseks (Naidu ja Hughes, 2001). Meetod põhineb spetsiifiliste

antikehade kasutamisel, mis seovad viirusosakesi mikrotiterplaadi kaevukeste pinnale. Esmalt kaetakse kaevukesed viirusespetsiifilise antikehaga ning inkubeeritakse. Seejärel lisatakse taimne proov, milles viirusosakesed seonduvad immobiliseeritud antikehadega. Pärast pesemist lisatakse ensüümiga märgistatud sekundaarne antikeha, mis seondub kompleksi. Viimases etapis lisatakse substraat, mis ensüümi toimel põhjustab värvimuutuse. Reaktsiooni intensiivsus on korrelatsioonis viiruse kogusega proovis ning seda hinnatakse visuaalselt või spektrofotomeetriliselt lainepikkusel 405 nm (Bioreba AG, 2023).

Nukleiinhapete eraldamiseks kartulitaimedest rakendatakse sageli CTAB (tsetüültrimetüülammooniumbromiid)-põhist meetodit ja kartuli mugulatest Hughesi ja Galau poolt välja töötatud liitiumkloriidimeetodit (Singh jt., 2013; Hema ja Konakalla, 2021). H&G meetod sobib taimeliikide ekstraheerimiseks, mis on raskemast materjalist, näiteks puuvilla puhul ning sama kehtib ka kõrge tärklise sisaldusega kartulimugulate kohta. Ameerika teadlased D. Wayne Hughes ja Glenn Galau töötasid 1988. aastal välja tõhusa meetodi RNA eraldamiseks puuvillast, mis väldib fenooli ja teiste denatureerivate ainete kasutamist ning põhineb madalal temperatuuril. Nende optimeeritud protokollis kasutatakse liitiumkloriidi, SDS-i (naatriumdodetsüülsulfaat), kaaliumatsetaati ja isopropanooli. Liitiumkloriid aitab parandada SDS-i lahustuvust, samas kui SDS lõhub valgumolekulide struktuure. Kaaliumatsetaat aitab seejärel valke koos SDS-iga sadestada, võimaldades need RNA-st eemaldada. Isopropanool aitab RNA sadestamisel. Selle meetodi edukaks rakendamiseks peab homogeniseeriv puhver olema kõrge ioonkontsentratsiooniga, aluseline (kõrge pH) ning sisaldama kõiki kolme detergenti (Hughes & Galau, 1988; He et al., 2021; Lab Reagents…).

Nendest meetoditest saadud ekstrakti põhjal on viiruste tuvastamiseks üks täpsemaid ja usaldusväärsemaid tehnikaid reaalaja pöördtranskriptsioon polümeraasi ahelreaktsioon (RT-qPCR), mis võimaldab viiruse tuvastamist nii lehest kui ka otse mugulast (Rubio jt., 2020). Reaalaja pöördtranskriptsiooni polümeraasi ahelreaktsioon (RT-qPCR) on viimastel aastatel kujunenud keskseks tööriistaks molekulaarbioloogia rakendustes ja teadusuuringutes (Bustin, 2007). TaqMan-tehnoloogial põhinev reaalaja RT-qPCR on väga tundlik meetod RNA olemasolu ja koguse määramiseks. Meetod võimaldab tuvastada ka madala ekspressioonitasemega transkripte, näiteks taimekudedest või -mahladest. Lisaks RNA olemasolu tuvastamisele võimaldab see tehnoloogia hinnata ka selle suhtelist ekspressiooni taset (Khanna ja Paoloni, 2007). Meetod tugineb pöördtranskriptsioonile, mille käigus sünteesitakse RNA alusel komplementaarne DNA (cDNA), mille amplifikatsiooni jälgitakse reaalajas spetsiaalse fluorestsentspõhise keemia ja seadmete abil. Tähtis roll on sihtmärk-RNA kvaliteedil. Madala kvaliteediga RNA või PCR-reaktsiooni inhibiitorite esinemine võib

oluliselt mõjutada tulemuste usaldusväärsust. Seetõttu on oluline tagada RNA efektiivne puhastamine ekstraheerimisprotsessi käigus (Bustin, 2005).

1.5. Sekveneerimine

DNA sekveneerimine on meetod, mille abil määratakse DNA molekuli nukleotiidide järjestus, andes teavet organismi geneetilise ülesehituse kohta. Algselt 1970. aastate lõpus välja töötatud Sangeri meetod (Sanger jt., 1977) oli aastaid standardseks lähenemiseks, kuid viimastel kümnenditel on see asendunud suure läbilaskevõimega tehnoloogiatega, mis võimaldavad töödelda miljoneid DNA fragmente paralleelselt. Sekveneerimine on osutunud asendamatuks tööriistaks nii teadusuuringutes kui ka kliinilises diagnostikas, sealhulgas haiguste ja kasvajate analüüsis. Uuemad tehnoloogiad võimaldavad järjestada geene ning tuvastada haigustekitajate olemasolu proovis kiiremini ja täpsemalt (Shendure jt., 2017).

1.5.1. Sekveneerimismeetodid

Sekveneerimismeetodid jagunevad arenguetappide kaupa kolmeks põlvkonnaks, esimese põlvkonna meetodid, mis hõlmavad peamiselt Sangeri sekveneerimist, teise põlvkonna meetodid ehk kõrge läbilaskevõimega sekveneerimine (HTS) ning kolmanda põlvkonna reaalajas üksikmolekuli sekveneerimine (Shendure jt., 2017).

Esimese põlvkonna meetodite hulka kuulub klassikaline Sangeri sekveneerimine, mis töötati välja 1977. aastal (Shendure jt., 2017). Meetod põhineb didesoksünukleotiidide (ddNTP) kasutamisel, mis põhjustavad DNA sünteesi lõppemise kindlas punktis. Sünteesitud fragmendid eraldatakse elektroforeesi teel (Sanger jt., 1977). Sangeri meetod on limiteeritud selle läbilaskevõime tõttu ning lisaks on see tehnoloogia aeglane (Fawcett ja Eterovic, 2017).

Teise põlvkonna sekveneerimine on kõrge läbilaskevõimega tehnoloogia (HTS), mis tuli laialdasse kasutusse 2000. aastate keskpaigas. Antud tehnoloogia võimaldab sekveneerida korraga rohkem fragmente ning selle kasvuga suurenesid ka lugemite pikkused. See tehnoloogia on oluline näiteks kogu genoomi, eksoomi ja RNA sekveneerimisel (RNA-Seq). Peamised kommertskasutuses olevad platvormid on Illumina, Ion Torrent ja (varem) Roche 454, mis kasutavad "*sequencing-by-synthesis*" (SBS) lähenemist. Selle tehnoloogia puhul kinnitatakse DNA fragmendid kindlale pinnale ja tehakse neist piisavalt palju koopiad, et saada tugev lugemissignaal. Kuna iga lugem põhineb ühe fragmendi järjestamisel, on SBS andmed digitaalsed ning see on eeliseks neid andmeid kvantitatiivselt hinnata. Enamik SBS-sekveneerimisseadmed toodavad lugemispikkusi vahemikus 100–400 aluspaari. Kuigi teise

põlvkonna sekveneerimise andmete tootmiskulud on viimaste aastatega oluliselt langenud ja genoomi-andmete maht hüppeliselt kasvanud siis tehnoloogial esineb endiselt piiranguid (Mardis, 2017). HTS turul on juhtival positsioonil Illumina platvormid (Bentley jt., 2008). Illumina sekvenaatorite valikusse kuuluvad mitmed seadmed, mis on optimeeritud erineva läbilaskevõime ja analüüside kiiruse jaoks. MiSeq platvorm on suunatud väiksematele laboritele ja kiiretele analüüsidele, võimaldades väikeste genoomide või suunatud sekveneerimise läbiviimist vähem kui 24 tunni jooksul. HiSeq 2500 on sobilik suurte andmemahtude genereerimiseks, võimaldades kuni 1 Tb andmeid 6 päevaga ning toetades ka kiiremat režiimi. HiSeq X Ten esindab suurt läbimurret populatsioonipõhise täisgenoomi sekveneerimise vallas. See platvorm võimaldab toota kuni 1,8 Tb andmeid kolme päevaga ning sekveneerida kuni 18 000 inimese genoomi aastas, saavutades 30× kaetavuse. Lisaks täiustatud optikale ja arvutusvõimekusele kasutab see platvorm uut *patterned flow cell* tehnoloogiat, mille abil standardiseeritakse klastrite paiknemine voolurakus. See võimaldab kõrgema tihedusega klastrite genereerimist ning oluliselt suuremat andmemahtu ühe jooksu kohta (Reuter jt., 2015).

Kolmanda põlvkonna sekveneerimine ehk reaalajas üksikmolekulide järjestus (TGS) nagu Pacific Biosciences SMRT-sekveneerimine ja Oxford Nanopore Technologies, võimaldavad DNA või RNA molekulide otsejärjestust ilma eelneva võimenduseta. See võimaldab reaalajas sekveneerimist, kus iga DNA molekuli loetakse läbi tervikuna, saavutades nii väga pika lugemispikkuse. PacBio SMRT kasutab suletud, üheahelalist tsüklilist DNA-malli, mida nimetatakse SMRTbell'iks. SMRTbell moodustatakse, kui sihtmärgiks oleva kaheahelalise DNA molekuli mõlemasse otsa liidetakse juuksenõeladapterid (*hairpin adaptors*). Oxford Nanopore sekveneerimine põhineb DNA või RNA molekuli läbitungimisel läbi valgusreguleeritud nanopoori, mis põhjustab muutuse ioonvoolus, mida saab mõõta ja iseloomustada (van Dijk jt., 2018). Kuna vajadus täpsete ja pikkade loetelude järele kasvab, siis PacBio uus sekveneerimismeetod HiFi'ga on võimalik saada väga täpseid pika lugemispikkusega järjestusandmeid (Hon jt., 2020). PacBio HiFI tehnoloogia võimaldab saavutada kõrge täpsusega järjestusandmeid, tuletades lõpliku loetelu mitmest järjestusläbimisest üheainsa DNA mallimolekuli kohta. Antud protsessi käigus kombineeritakse sama molekuli kohta saadud alamloetelud, mis eraldi võivad sisaldada juhuslikke vigu, kuid korduva järjestamise kaudu moodustub usaldusväärne konsensusjärjestus. Uuringud näitavad, et juba kümne järjestusläbimise korral saavutavad loetelud mediaanselt Q30 täpsusklassi, mis vastab ligikaudu 99,9% järjestustäpsusele. Täpsus paraneb läbitud tsüklite arvuga, pakkudes väärtuslikku kompromissi loetelu pikkuse ja kvaliteedi vahel (Wenger jt., 2019).

1.5.2. RNA sekveneerimise eripärad

RNA sekveneerimine (RNA-seq) on tehnoloogia, mis võimaldab uurida kogu rakus või koes väljendatud RNA molekulide järjestusi ja hulka (Wang jt., 2009). RNA-seq on tänapäeval laialdaselt kasutusel geenieksпressiooni analüüsiks, haiguste tuvastamiseks ning erinevate organismide uurimiseks (Wang jt., 2009; Li jt., 2011; Zhao jt., 2014). Enne sekveneerimist tuleb kogu RNA ekstraktist eemaldada ribosomaalne RNA (rRNA), mis moodustab kuni 90% kogu RNA massist, et keskenduda informatiivsematele molekulidele, nagu mRNA või viiruslik RNA. Lisaks tuleb RNA muuta cDNA-ks pöördtranskriptaasi abil (Zhao jt., 2014). RNA sekveneerimiseks on mitu strateegiat, sõltuvalt uurimise eesmärgist. Kogu transkriptoomi sekveneerimine võimaldab tuvastada erinevaid RNA molekule, uusi transkripte ja määrata täpselt transkriptsiooni algus- ja lõpp-punkte (Wang jt., 2009). mRNA-seq keskendub polü-A RNA-le ja sobib hästi eukarüootsete geenide ekspressiooni analüüsiks (Mortazavi jt., 2008). Ahelaspetsiifiline RNA-seq võimaldab avastada uusi transkripte, viia läbi genoomi annotatsiooni ning geenieksпressiooni (Levin jt., 2010).

RNA-seq on rakendatav ka viiruste uurimisel, selle lähenemisega on võimalik otse sekveneerida RNA-viiruseid, viroide ning aktiivselt replitseeruvate DNA-viiruste RNA-vorme (Adams jt., 2009). See teeb RNA-seq-ist asendamatu tööriista viirusdiagnostikas ja patogeenide jälgimises põllumajanduses (Maclot jt., 2020).

1.6. Sekveneerimisandmete analüüs

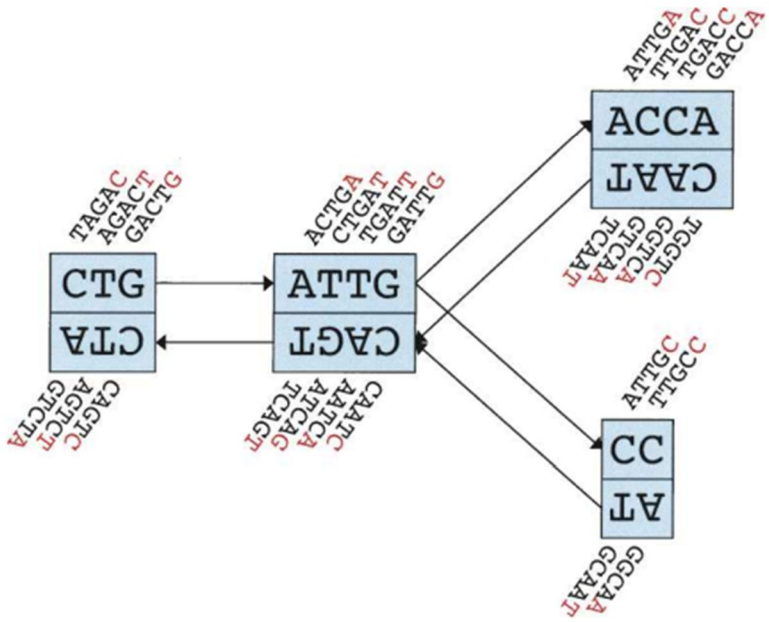
RNA või DNA järjestuse analüüs jaguneb üldiselt kaheks põhisuunaks: *de novo* järjestuse kokkupanek (*assembly*), mida kasutatakse siis, kui referentsgenoom puudub, ja järjestuste joondamine (*alignment*), kus loetud fragmendid võrreldakse juba olemasoleva genoomiga (Li ja Homer, 2010). Lisaks kasutatakse spetsiaalseid tööriistu viiruste identifitseerimiseks, geenieksпressiooni kvantifitseerimiseks ning funktsionaalse annotatsiooni tegemiseks (Wood jt., 2019; Trapnell jt., 2012; Buchfink jt., 2015).

1.6.1. *de novo* järjestamine

de novo järjestamine on järjestusandmete kokkupaneku meetod, mis on genoomika teadusharu keskne meetod, mille abil on võimalik määrata nii uute liikide genoomide järjestusi kui ka indiviidide genoomijärjestusi populatsioonides. Lisaks võimaldab järjestamine analüüsida RNA molekule konkreetsetest proovist ning kasutada DNA järjestust molekulaarbioloogilistes meetodites täpse lugemisena (Pop, 2009; Zerbino ja Birney, 2008). Järjestamise kvaliteet sõltub

lugemite pikkusest ja katvusest, kuid isegi väga lühikestest lugemitest ja kõrge katvusega andmekogumitest on võimalik saada usaldusväärseid tulemusi (Nagarajan ja Pop, 2013). Täielike genoomijärjestuste saamine erinevate liikide puhul on endiselt oluline eesmärk, mis aitab paremini mõista bioloogilist mitmekesisust ning evolutsioonilisi protsesse (Zerbino ja Birney, 2008).

Üks tööriistu *de novo* järjestamiseks on Velvet, mis töötati välja lühikeste järjestuslugemite kokkupanekuks, kasutades de Bruijni graafe. Joonis 5 illustreerib de Bruijni struktuuri graafiliselt. Igas sõlmes on k-mer'ide viimased nukleotiidid esile tõstetud ja nende järjestus moodustab sõlme järjestuse. Igal sõlmel on ka kaksiksõlm, mis esindab pöördkomplementaarseid k-mer'e (ehk vastassuunas loetud ja täiendatud jadade versiooni). See võimaldab arvesse võtta järjestuslugemisi mõlemalt DNA ahelalt. Sõlmede vahelised kaared on näidatud nooltena ja näitavad, kuidas järjestus liigub ühelt sõlmelt teisele (Zerbino ja Birney, 2008).



Joonis 5. de Bruijn struktuur graafiliselt (Zerbino, Birney 2008).

Velveti kasutamiseks on oluline parameetrite (näiteks eelnevalt määratud k-mer'i pikkuse) optimaalne seadistamine, mis mõjutab otseselt järjestuse katvust ja kokkupaneku järjepidevust (Zerbino ja Birney, 2008).

Alternatiivsed tööriistad, nagu SPAdes, ABySS ja SOAPdenovo, on samuti laialdaselt kasutusel ning neid eelistatakse sageli olukordades, kus on tegemist keerukamate või suuremate genoomidega (Bankevich jt., 2012; Simpson jt., 2009; Li jt., 2010). Neid kasutatakse suures

osas genoomika töövoogudes, eriti metagenoomsete või keeruliste andmekogumite korral (Nagarajan ja Pop, 2013).

1.6.2. BLAST-analüüs

BLAST (*Basic Local Alignment Search Tool*) on bioinformaatiline tööriist, mis kasutab lähenemist, mis otseselt läheneb dünaamilise programmeerimise algoritmiga saadavale tulemusele, optimeerides sarnasuse mõõdikut. BLAST suudab tuvastada ka nõrku, kuid bioloogiliselt olulisi järjestuste sarnasusi (Altschul jt., 1990). Meetod töötab, otsides esmalt andmebaasist lühikesi järjestusi, mis on päringujärjestusega piisavalt sarnased. Seejärel kontrollitakse, kas nende tabamus kuulub pikemasse sarnasesse piirkonda, mida võiks käsitleda kui märkimisväärset joondust. Selleks laiendatakse tabamust mõlemas suunas, kuni järjestuste sarnasus enam ei parane (Altschul jt., 1997). BLAST võimaldab mitmeid variante, sealhulgas blastn DNA järjestuste võrdlemiseks ning blastp valkude jaoks, samuti blastx, mis tõlgib nukleotiidse päringu kuues lugemisraamis ja võrdleb seda valguandmebaasiga. Andmebaaside ettevalmistamiseks lokaalseks analüüsiks kasutatakse tööriista makeblastdb, mis teisendab fasta-failid BLASTi indeksiga andmebaasiks (Tao, 2010). BLASTi kasutamine võimaldab kindlaks teha järjestuse päritolu, funktsiooni või evolutsioonilist sugulust (Pearson, 2013).

1.6.3. Järjestusandmete joondamine ja kaardistamine

Üks peamisi referentsgenoomi ülesandeid on toimida sihtmärgina teiste genoomide järjestusandmete joondamisel. Järjestuslugemite joondamine referentsgenoomile tähendab iga sisendjärjestuse positsiooni seostamist vastava homoloogilise positsiooniga referentsgenoomis. Selline joondus võimaldab teha kindlaks, millised piirkonnad järjestatud genoomis vastavad referentsgenoomile, ning annab seeläbi olulist teavet vastava genoomi ülesehituse ja variatsioonide kohta.

Üks joondamistööriistu on BWA (*Burrows-Wheeler Aligner*), mis on optimeeritud lühikeste järjestuste kiireks ja täpseks joondamiseks suurele referentsile (Li ja Durbin, 2009). Pärast lugemite joondamist salvestatakse tulemused üldjuhul SAM (*Sequence Alignment/Map*), mida saab töödelda tööriistaga SAMtools, et eraldada joondusi kindlas genoomipiirkonnas, ühendada ja sorteerida joondusi, määrata ühe nukleotiidi polümorfismid (SNP-d) ja insert-sioonid/deletsioonid ning visualiseerida joondusi (Li jt., 2009). Joondamise tulemuste analüüs võimaldab tuvastada mutatsioone, struktuurilisi erinevusi ja muid geneetilisi variatsioone, mis on olulised nii teadustöös kui ka kliinilistes rakendustes (Paten jt., 2014).

1.6.4. Tüvede määramine metagenoomsetest andmetest

Tüvede määramiseks järjestatud andmetest on loodud GenomeTester tarkvarapakett, mis võrdleb genoome ja tuvastab geneetilisi markereid. Selle üheks ülesandeks on genoomsetes andmetes k-mer'ide (järjestatud nukleotiidijärjestuste pikkusega k) sageduste lugemine. GenomeTester tarkvara loendab antud järjestustes kõik esinevad k-mer'id ja salvestab need binaarvormingus, mis võimaldab andmete hilisemat kiiret töötlemist. Sellise lähenemise eeliseks on võimalus võrrelda kiiresti väga suurt hulka järjestusi ilma vajaduseta neid eelnevalt joondada, mis teeb meetodi sobilikuks ka tüvedevaheliste erinevuste analüüsiks (Kaplinski jt., 2015).

1.6.5. Taksonoomiline klassifitseerimine

Taksonoomiline klassifitseerimine on oluline etapp järjestusandmete analüüsis, mille eesmärk on iseloomustada organismide mitmekesisust, geene ja metaboolseid protsesse (Menzel jt., 2016). Kaasaegsete järjestustehnoloogiate abil saadud suurte metagenoomsete andmehulkade töötlemiseks on välja töötatud mitmeid tööriistu, millest üks on Kraken2 (Wood jt., 2019).

Kraken2 on kiire ja täpne taksonoomilise klassifitseerimise tarkvara, mis kasutab k-mer-põhist meetodit järjestuste määramiseks. Iga järjestuse k-mer võrreldakse eelnevalt koostatud andmebaasis olevate k-mer'idega, millele on määratud taksonoomiline tähis (Wood jt., 2019). Selle tulemusena suudab Kraken2 väga suure efektiivsusega töödelda miljoneid järjestusi, säilitades sealjuures kõrge täpsuse ka madala katvusega proovide korral. Kraken2 sobib hästi tüvede määramiseks, kuna võimaldab täpset ja kiiret järjestuste taksonoomilist jaotamist, kuid selle kasutamine sõltub kasutatavast andmebaasist (Wood jt., 2019).

1.6.6. Fülogeneetiliste puude koostamine ja analüüs

Fülogeneetika kujutab endast kiiresti arenevat teadusvaldkonda, mille raames on välja töötatud sadu spetsialiseeritud tarkvaralahendusi erinevat tüüpi andmete analüüsimiseks. Need tööriistad võimaldavad käsitleda laia valikut uurimisküsimusi ning täita mitmesuguseid teaduslikke eesmärke, näiteks hinnata evolutsioonilisi suhteid (Hall, 2013). Fülogeneetilises analüüsis on oluline järjestuste mitmikjoondamine, milleks kasutatakse näiteks MUSCLE algoritmi (Edgar, 2004). MUSCLE algoritm on kättesaadav näiteks MEGA keskkonnas (Awe jt., 2023). MUSCLE algoritm kasutab järjestuste sarnasuse hindamiseks selliseid kaugusmõõdikut nagu k-mer kaugust joondamata järjestuste ja Kimura kaugust joondatud järjestuste puhul. K-mer põhine hindamine võimaldab kiiret esmase sarnasuse määramist ilma joonduseta, samas kui Kimura mudel korrigeerib kauguse hindamist, arvestades korduvaid mutatsioone ühes

positsioonis. Järjestuste rühmitamiseks kasutatakse UPGMA (*Unweighted Pair Group Method with Arithmetic Mean*) meetodit, mis võimaldab efektiivset ja täpset järjestusjoondust (Edgar, 2004).

Üks populaarsemaid programme suurte andmestike fülogeneetiliseks analüüsimiseks on RAxML (*Randomized Axelerated Maximum Likelihood*). Antud programmi peamiseks tugevuseks on kiire maksimaalse tõenäosuse puuotsingu algoritm, mis suudab leida kõrge tõenäosusskooriga fülogeneetilisi puid. RAxML võimaldab hinnata fülogeneetilise puu toestust bootstrap-meetodi abil. Antud programm toetab DNA- ja valgusjärjestusti, binaarseid, mitmeolekulisi morfoloogilisi ning RNA teisese struktuuri andmeid. RAxML suudab ka automaatselt määrata parima valgu asendusmudeli ning hinnata baaside sagedusi maksimaalse tõepära meetodil (*maximum likelihood*). RAxML võimaldab fülogeneetiliste puude teostamist suurte, kümnete tuhandete taksonitega andmekogumite puhul. Samuti võimaldab programm arvutada marginaalseid eellasolekute seisundeid (*ancestral states*) ning pakub algoritmi puude juurimiseks (Stamatakis, 2014).

2. EKSPERIMENTAALNE OSA

2.1. Töö eesmärgid

Käesoleva magistritöö eesmärgid on:

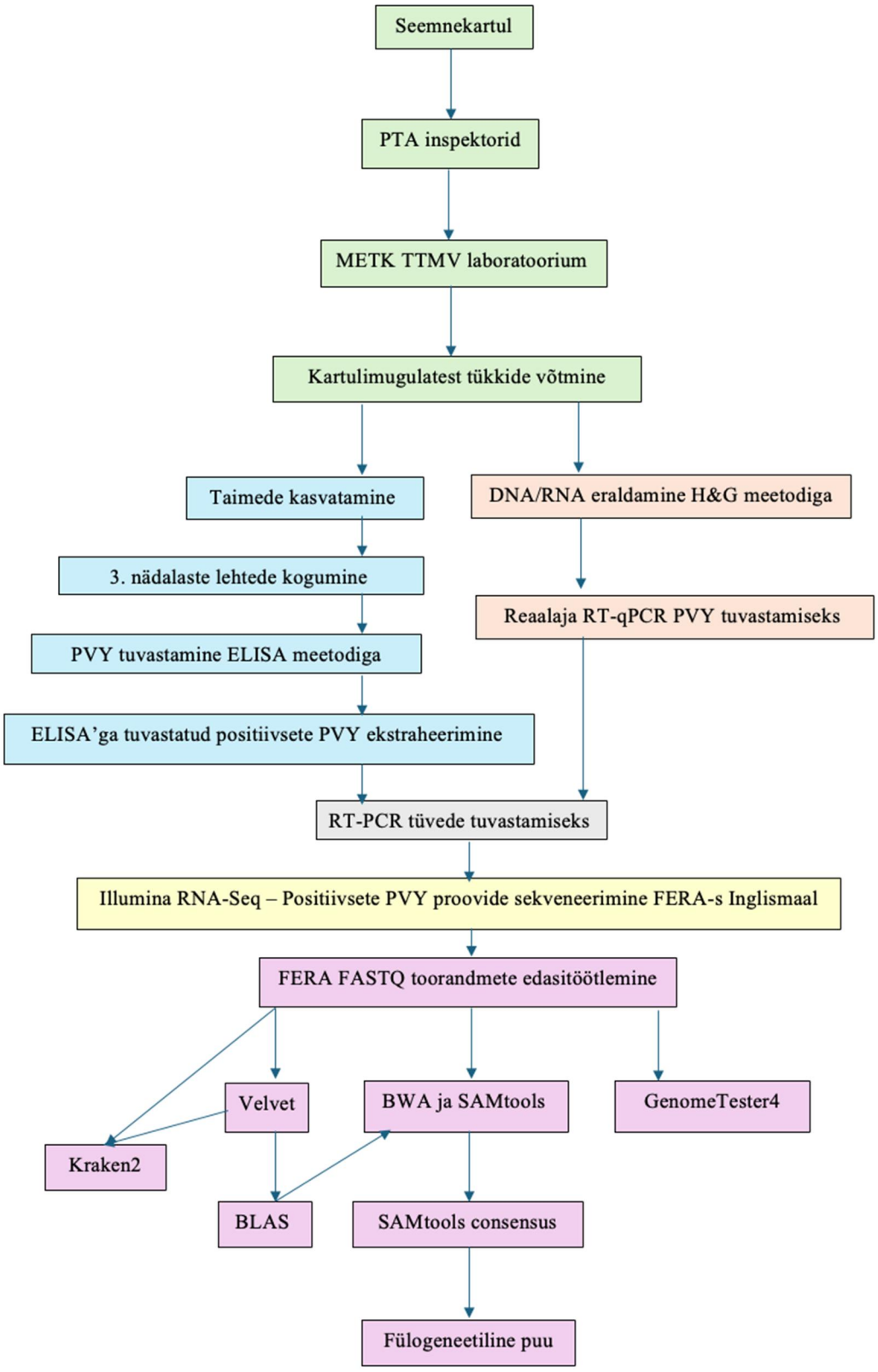
1. Eestis esinevate PVY tüvede järjestuste määramine standardprotseduuril kogutud proovidest;
2. Rekombinantsete tüvede kaardistamine;
3. PVY tüvede fülogeneetiline analüüs;
4. HTS'i hindamine kui toetav meetod erinevate taimeviiruste tuvastamiseks seemnekartulis.

2.2. Materjal ja metoodika

Joonis 6 annab ülevaate skeemina, kuidas kartuli Y-viirust tuvastati antud magistritöö raames ja millistest etappidest koosnes järjestuse analüüsi protsess. Protsess algas sertifitseeritud seemnekartuli proovide kogumisega, mida teostasid Põllumajandus- ja Toiduameti (PTA) inspektorid. Kogutud proovid edastati analüüsimiseks Maaelu Teadmuskeskuse (METK) Taimetervise ja Mikrobioloogia valdkonna (TTMV) laboratooriumisse. Kõik laborisse saabunud mugulad analüüsiti PVY esinemise suhtes. Kuni 2015 aastani oli esmaseks tuvastamismeetodiks ELISA, alates 2015 aastast kasutati esmase määramismeetodina reaalaja RT-qPCR.

Kõikidest proovidest kust tuvastati kartuli Y-viirus aastatel 2011 kuni 2013; 2016 kuni 2019 viidi läbi täiendav RT-PCR PVY tüvede tuvastamiseks. Valik proove, mis esindasid erinevaid variatsioone edastati sekveneerimiseks Inglismaale FERA (The Food and Environment Reasearch Agency) laborisse, kus viidi läbi Illumina RNA-Seq põhine sekveneerimine. Sekveneerimise tulemusena saadud FASTQ toorandmetele tehti esmane analüüs FERAs, kasutades nende poolt väljatöötatud bioinformaatilist analüüsiskeemi. Samuti saadeti METK'i FASTQ toorandmed, millest viidi käesoleva magistritöö raames läbi lisa bioinformaatiline analüüs. Bioinformaatilise andmeanalüüsi eesmärgiks oli määrata viiruse järjestus ning hinnata selle geneetilist varieeruvust ja päritolu, lisaks määrata ka mis teisi organisme antud proovides esineda võiks. Andmetöötluses kasutati järjestuse kokkupanekuks Velvetit ja järjestuste taksonoomiliseks klassifitseerimiseks Kraken2 tööriista. Lisaks rakendati BLAST-analüüsi, et leida joondamiseks sobilik referentsgenoom. Järjestuste täpseks kaardistamiseks ja töötlemiseks kasutati BWA, SAMtools ja GenomeTester4 programme. Seejärel loodi

konsensusjärjestused, mille alusel koostati fülogeneetiline puu, hindamaks saadud PVY genoomide ja erinevate viiruste isolaatide omavahelist geneetilist seost.



Joonis 6. Kartuli Y-viiruse ja järjestuse analüüsi protsessi skeem.

Joonisel kujutatud protsessi eesmärk on saavutada PVY usaldusväärne ja tundlik tuvastamine ning pakkuda põhjalikku molekulaarset ülevaadet viiruse geneetilisest mitmekesisusest. Selline lähenemine toetab seemnekartuli kvaliteedikontrolli ja aitab ennetada ning kaardistada kartuli Y-viiruse levikut kartulikasvatuses.

2.2.1. Kartuli proovid

Kartulil esineva PVY tüvede iseloomustamiseks kasutati inspektorite poolt rutiinse kontrolli käigus kogutud ja laboris analüüsitud seemnekartuli proove. Tabelis 1 on välja toodud katses kasutatud proovid. Kokku analüüsiti 15. erinevat kartulitaime proovi, mis pärinesid järgmistest kartulisortidest: „Lucilla“, „Arielle“, „Red Lady“, „Tiina“, „Albertine“, „Teele“, „Milva“, „Manitou“ ja „Anuschka“. Laboratoorsed esmased analüüsid viidi läbi 2011-2012 ja 2019 Harjumaal, Saku vallas, Sakus endises Põllumajandusuuringute Keskuses (PMK), nüüdses Maaelu Teadmuskeskuses.

Tabel 1. Katsetes kasutatud proovi tähistused, nende sordid ja analüüsimise aasta.

Proovi tähistus	Sort	Analüüsimise aasta
Eesti-1	Lucilla	2019
Eesti-2	Lucilla	2019
Eesti-3	Arielle	2019
Eesti-4	Red Lady	2019
Eesti-5	Tiina	2019
Eesti-6	Red Lady	2019
Eesti-7	Lucilla	2019
Eesti-8	Albertine	2019
Eesti-9	Teele	2019
Eesti-10	Arielle	2011
Eesti-11	Arielle	2011
Eesti-12	Milva	2012
Eesti-13	Manitou	2011
Eesti-14	Anuschka	2011
Eesti-15	Arielle	2011

2011. ja 2012. aastal kasvatati taimed mugulast ette TTMV karantiinkasvuhoones. Kartuli lehed koguti labori töötajate poolt analüüsimiseks siis, kui neil oli väljaarenenud 3 – 4 lehte ja taim oli vähemalt 3. nädala vanune. 2019. aastal viidi analüüsid läbi otse mugulatest. Laborisse saabunud mugulaproovid sisaldasid 100 mugulat, mis jagati 10 alaprooviks ning kust võeti tükid mugula mõlemast otsast.

Lehtedest pressitud mahla ja mugulatükke kasutati totaalse nukleiinhappe eraldamiseks. Kõik proovid säilitati –20 °C juures kuni edasiste analüüsietappideni.

2.1.1. Kartuliviiruste tuvastamise meetodid

Ekstraheerimised kartuli lehtedest ja mugulatest viidi läbi kasutades Maaelu Teadmuskeskuse Laboratooriumite Kvaliteedikäsiraamatu TTMV tööjuhendeid. Lehtedest PVY määramiseks kasutati ELISA (Lisa 1) meetodit ning mugulatest H&G (Lisa 2) ja reaalaja RT-qPCR meetodit. Aastatel 2011-2012 kasutati ELISA meetodit, kuna laboris ei olnud H&G ekstraktsioonimeetod ja reaalaja RT-qPCR juurutatud, mis on oluliselt kiirem ja mida saab kasutada raske materjali puhul nagu on mugul. Eraldatud DNA/RNA säilitati -20 °C juures.

2.1.2. TaqMan Reaalaja pöördtranskriptsiooni polümeraasi ahelreaktsioon

Kartuliproovidest eraldatud DNA/RNA analüüsiti kasutades TaqMan Reaalaja pöördtranskriptsiooni polümeraasi ahelreaktsiooni (RT-qPCR) meetodit. Viiruste tuvastamine viidi läbi vastavalt metoodikale üheetapilise reaktsioonina, kus reaktsioonisegu lõppmaht oli 25 mikroliitrit.

Tabelis 2 on esitatud kasutatud reaktsioonisegu komponendid, mahud ning lõppkontsentratsioonid. Reaktsioonisegu valmistati, kombineerides vajalikud komponendid (puhver, dNTP-d, $MgCl_2$, praimerid, sond, pöördtranskriptaas ja DNA-polümeraasi ensüüm) vastavalt protokollile. Valmis reaktsioonisegu jaotati mikrotiterplaadi süvenditesse ning sellele lisati analüüsitav RNA- või DNA-proov.

Tabel 2. Reaalaja RT-PCR reaktsioonisegu komponendid ja nende maht.

Reaktsioonisegu komponent	Maht reaktsioonisegus (µl)	Lõpp kontsentratsioon reaktsioonisegus	Komponendi tootja
ddH2O	10,38		
10x Gold puhver	2,5	1x	Applied Biosystems
MgCl2	5,5	5,5 mM	Applied Biosystems
dNTP	2,0	800 nM	Promega
Pärisuunaline (F) praimer	1,0	300 nM	Sigma-Aldrich
Vastassuunaline (R) praimer	1,0	300 nM	Sigma-Aldrich
TaqMan sond	0,5	100 nM	Sigma-Aldrich
MMLV (RevertAid)	1,0	0,02 U/µl	Thermo Fisher Scientific
AmpliTaq Gold	0,125	0,025 U/µl	Applied Biosystems
Proovi nukleiinhape	1		

Reaalaja RT-qPCR-il kasutati kahte erinevat praimeripaari (Tabel 3) ja sondi, selleks, et kõige pealt kontrollida ekstraktsiooni õnnestumist tsütokroomi oksidaasi olemasoluga ja seejärel

kasutati spetsiifilisi praimereid kartuli Y-viiruse tuvastamiseks. Tsütokroomi oksidaasi olemasolu kontrollimiseks kasutati kontrollgeeni COX F, COX R praimereid ja COX sondi. PVY tuvastamiseks kasutati PVY F, PVY R praimereid ja PVY sondi.

Tabel 3. Reaalaja RT-qPCR reaktsioonisegus kasutatud praimerid

Märklaud-organism	Praimer	Praimeri järjestus (5' – 3')
COX	COX F	CGTCGCATTCCAGATTATCCA
	COX R	CAACTACGGATATATAAGRRCCRRAACTG
	COX sond	JOE – AGGGGCATTCCATCCAGCGTAAGCA – TAMRA
PVY	PVY F	GGGCTTATGGTTTGGTGCA
	PVY R	CCGTCATAACCCAAACTCCG
	PVY sond	FAM – TGAAAATGGAACCTCGCCAAATGTCA – TAMRA

Tabelis 4 on näha, et TaqMan Reaalaja pöördtranskriptsiooni polümeraasi ahelreaktsiooni esimeseks etapiks oli pöördtranskriptsioon ehk RNA pealt sünteesiti DNA, mis viidi läbi 48 °C juures 30 minutit. Seejärel toimub eelkuumutamine 10 minutit 98 °C juures, et lõhustuda vesiniksidemed ja moodustada üksikahelaline DNA. Sellele järgnes denaturatsioon 95 °C 15 sekundit ning praimerite seondumine ja DNA süntees 60 °C juures 1 minut. Denaturatsiooni, praimerite seondumise ja DNA sünteesi etappe korrati 40 korda. Igale analüüsile lisati juurde ka positiivne ja negatiivne kontroll ning üks kuni kaks reaktsioonisegu kontrolli, et veenduda analüüsi kvaliteedis.

Tabel 4. Reaalaja RT-qPCR reaktsiooni etapid.

Reaalaja RT-PCR reaktsiooni etapp	Temperatuur (°C)	Etapi aeg	Tsüklite arv
Pöördtranskriptsioon	48	30 minutit	1
Eelkuumutamine	95	10 minutit	1
Denaturatsioon	95	15 sekundit	40
Praimerite seondumine ja DNA süntees	60	1 minut	

Reaalaja RT-qPCR analüüsid viidi läbi QuantStudio™ 5 Real-Time PCR System masinaga ja tulemused mõõdeti antud masina tarkvaraga. Tulemused esitati Ct väärtustena, mis näitab reaalaja RT-qPCR programmi tsüklite tundlikkusläve, kus on fluorestsents tuvastatav. Tundlikkusläveks nimetatakse joont, mis määrab amplifikatsioonikõvera Ct väärtuse ning antud meetodi puhul on selle ülemiseks läveks 40. Kui 40 tsükli jooksul fluoresentsis tõusu ei tuvastata ei sisalda proov märklaudorganismi. Reaalaja RT-qPCR testides kasutatakse üldjuhul Ct tundlikusläve piirina 35 – 40 tsüklit, seega kõik, mis jäävad alla nende väärtuse on positiivsed. Endises PMK ja nüüdses METK TTMV laboratooriumis on määratud selleks piiriks 36.

2.2.2. Kartuli Y-viiruse positiivsete proovide sekveneerimine

Välja valitud kartuli Y-viiruse suhtes positiivseks osutunud proovide ekstraktid saadeti Inglismaale sekveneeriti 2020. aastal Ühendkuningriigis The Food and Environment Research Agency (FERA) asutuses. Proovid valmistati ette ja järjestati vastavalt Fowkes jt (2021) poolt kirjeldatud protokolli.

Esmalt valmistati iga RNA–proovi põhjal kasutades Illumina TruSeq Complete Plant Leaf komplekti ribosoomivabad, indekseeritud sekveneerimisraamatukogud. Protokoll sisaldas rRNA eemaldamist, et suurendada viiruslike järjestuste tuvastamise tundlikkust. Igale proovile lisati unikaalne indeks, et hiljem eristada proove üksteisest. Seejärel segati kõik raamatukogud kokku ühtseks seguks ja sekveneeriti Illumina MiSeq platvormil, kasutades V3 2 x 300 bp paariotsalist sekveneerimiskomplekti.

Sekveneerimise järgselt eraldati andmed proovide kaupa vastavalt indeksitele, kasutades MiSeq seadme sisseehitatud tarkvara. Saadud lugemid töödeldi edasi bioinformaatilise tööriista Angua abil. Angua on bioinformaatiline töövoog, mis analüüsib Illumina paired-end sekveneerimisandmeid viiruste tuvastamiseks. Töö algab andmete puhastamise ja kvaliteedikontrolliga (FastQC, MultiQC, Bbduk), millele järgneb transkriptide koostamine programmiga Trinity. Vähemalt 200 baasi pikkused järjestused võrreldakse nt andmebaasiga (BlastN) ja analüüsitakse MEGANi abil, et tuvastada viiruslik päritolu. Sarnased järjestused eemaldatakse klastrimisega (Mmseqs2) ning järjestusi kontrollitakse täiendavalt valgu tasemel (diamond-BlastX). Tulemused visualiseeritakse MEGANis ja koostatakse kokkuvõtlikud koostamisstatistikad (Angua-nf, 2024).

Käesolevas töös kasutati üksnes FERA laborist saadud toorandmeid ehk FASTQ formaadis järjestusfaile, mille põhjal viidi läbi edasine andmeanalüüs Eestis. Sekveneerimise eksperimentaalne osa ise viidi läbi täielikult FERA laboris.

2.2.3. *de novo* järjestuse kokkupanek

Edasiseks analüüsiks viiruste järjestuste rekonstrueerimise jaoks kasutati Velvet assemblerit, et teha *de novo* järjestuse kokkupanek. Analüüs viidi läbi kahes etapis, kus kasutati käsureal *velveth* ja *velvetg* programme.

Analüüsitava genoomse materjali kokkupaneku esimene samm viidi läbi Velvet tarkvarapaketi osaga *velveth*, mille ülesandeks on luua esmased andmestruktuurid. Sisendina kasutati eelnevalt kvaliteedikontrolli läbinud (QC-*filtered* ehk kvaliteedi kontrolli läbinud) paarisotsalisi

järjestuslugemeid FASTQ formaadis. Lugemid olid töödeldud Illumina MiSeq süsteemi tarkvara abil, mis hõlmas adapterite eemaldamist, madala kvaliteediga aluspaaride trimmimist ning muude järjestusmüra allikate elimineerimist. Selle põhifunktsioon on k-mer põhiste andmestruktuuride, nagu *Roadmaps* ja *Sequences* failide genereerimine, mis on aluseks järgnevatele *de novo* järjestuse kokkupaneku sammudele. Oluline parameeter selles etapis on k-mer pikkus, mis mõjutab oluliselt lõplikku assembleri kvaliteeti. Velvet kasutab eri pikkustega k-mer'e.

Teises etapis käivitati *velvetg*, mille sisendiks on *velveth*-i poolt genereeritud kaust. *Velvetg* kasutab seal salvestatud andmestruktuure kontiigide loomiseks. Selle protsessi käigus arvutatakse katvused ja ühendatakse sobivad järjestused kokku. Väljundfailideks tekivad *velvetg* käsu tulemusel *contigs.fa*, *stats.txt* ja *Graph2*. *Contigs.fa* failis sisalduvad järjestatud kontiigid FASTA-formaadis, mis sobivad edasiseks analüüsiks nagu näiteks BLAST. *Stats.txt* annab ülevaate kokkupaneku statistilistest näitajatest nagu on kontiigide arv, nende pikkus ja katvus ning N50. *Graph2*-te kasutatakse kokkupaneku sisemiste struktuuride salvestamiseks.

Kasutatud RNA-Seq andmestik, mis põhineb kogu proovist pärineva RNA sekveneerimisel, võimaldas koguda mitte ainult uuritava kartuli Y-viirus järjestusi, vaid ka avastada teisi RNA-viiruseid, mis proovides võisid esineda. RNA-Seq meetod ei ole viiruse-spetsiifiline ning hõlmab kõiki proovis esinevaid aktiivselt transkribeeritavaid RNA molekule, mistõttu saab tuvastada ka kaasinfektsioone või varjatud viiruseid, mis ei pruugi olla eelnevalt teada.

2.2.4. BLAST-analüüs ja referentsgenoomi valik

Velvet assembleriga saadud kokkupandud järjestused (kontiigid) võeti edasiseks taksonoomiliseks määramiseks BLAST-analüüsile, kasutades selleks kohapeal koostatud viirusetüvede andmebaasi. See andmebaas loodi Geneious Prime tarkvara abil, kuhu koondati kõik teadaolevad kartuli Y-viiruse (PVY) täispikad genoomijärjestused, mis olid saadaval NCBI GenBank andmebaasis seisuga 03.04.2025. Kokku koguti 705 PVY genoomijärjestust, mis eksporditi FASTA-formaadis ja vormistati BLAST-iga ühilduvaks lokaalseks andmebaasiks. Selleks kasutati NCBI BLAST+ käsureatööriista *makeblastdb*, millega loodi vajalik andmebaasistruktuur ning viidi läbi andmebaasi indekseerimine.

BLAST-analüüsiks kasutati Velvetiga genereeritud kontiig-faile FASTA-formaadis, mida võrreldi eelnevalt loodud lokaalse PVY andmebaasiga. Analüüsi eesmärk oli tuvastada iga kontiigi suurim järjestuslik sarnasus teadaolevate PVY tüvedega. Seejärel valiti edasiseks järjestuste kaardistamiseks välja need referentsgenoomid, millel oli BLAST-analüüsis kõrgeim

katvus ja identiteet, et tagada võimalikult täpne ja usaldusväärne viiruse genoomi rekonstruktsioon.

Kuna PVY genoomid on äärmiselt varieeruvad ja sageli esinevad rekombinantseid tüvesid, ei olnud piisav lähenemine kasutada ühtainsat referentsgenoomi kogu joondamise protsessis. Sellisel juhul jäi kontiigide kaetus sageli osaliseks või tekkis mittespetsiifiline vastavus. Seetõttu viidi läbi BLAST-analüüs iga proovi vastu eraldi, mis võimaldas tuvastada tüvepõhise maksimaalse katvusega vasteid, ning tagas seeläbi täispikkade genoomide usaldusväärse rekonstrueerimise. Eriti oluline oli see juhul, kui proovides esinesid rekombinandid või koinfektsioonid, kus erinevad genoomipiirkonnad võivad pärineda erinevatelt tüvedelt. Selline lähenemine suurendas analüüsi spetsiifilisust ja võimaldas hiljem korrektsemat tüvede klassifitseerimist.

2.2.5. Järjestusandmete kaardistamine referentsgenoomile

BLAST-analüüsil sobiva referentsgenoomi valimise põhjal viidi läbi järjestusandmete kaardistamine, kasutades BWA ja SAMtools tööriistu. Sisendfailidena kasutati puhastatud järjestusandmeid FASTQ-formaadis ning valitud PVY referentsgenoomi FASTA-faili. Kõigepealt indekseeriti referentsgenoom *bwa index* käsuga ning seejärel viidi läbi järjestuslugemite kaardistamine *bwa mem* abil, mille tulemuseks olid SAM-formaadis joondusfailid. SAM-failid konverteeriti BAM-formaati käsuga *samtools view*, seejärel sorteeriti (*samtools sort*) ja indekseeriti (*samtools index*). Saadud sorteeritud ja indekseeritud BAM-failid kasutati edasiseks analüüsiks, sealhulgas katvuse hindamiseks ja SNP-variantide tuvastamiseks. Joondamise protsess tagas, et järjestusandmed oleksid kvaliteetselt ja täpselt seotud valitud referentsgenoomiga, mis on edasiste analüütiliste sammude usaldusväärsuse aluseks. Lisaks sellele koostati Eesti enda referentsgenoom samtools consensus meetodi abil, mis võimaldas luua konsensusjärjestuse olemasolevate andmete põhjal.

2.2.6. Kartuli Y-viiruse tüvede määramine

Proovis viirusetüvede rekombinatsioonikohtade positsioneerimiseks järjestuse ulatuses ja koinfektsiooni tuvastamiseks kasutati GenomeTester4 tarkvarapaketti, täpsemalt funktsiooni *gmer_counter*, mis võimaldab kvantitatiivselt hinnata lugemite jaotumist erinevate PVY tüüpide (PVY-N ja PVY-O) vahel. Selles analüüsis kasutati viirusetüüpide võrdlusena järgmisi genoomiviiteid: MW811190.1 (PVY-N) ja KY847925.1 (PVY-O), mis valiti nende kõrge kvaliteedi ja täieliku järjestuse alusel. Selleks koostati *gmer_counter*i andmebaas, mis sisaldas mõlemat tüve eristavaid k-mer'e kogu järjestuse ulatuses. *gmer_counter*i andmebaasis tohib iga

k-mer olla vaid korra, seetõttu samu k-mer'e sisaldavad genoomi osad jäävad graafilise liidesest välja ning tulemus on indikatiivne koinfektsiooni ja rekombinantide osas. Analüüsiks kasutati *gmer_counter*i poolt arvutatud tüvespetsiifiliste k-mer sagedusi FASTQ sisend failis. Selle abil oli võimalik täpselt määrata, milline osa lugemitest kattus N-tüüpi või O-tüüpi viirustüvedele iseloomulike järjestustega. Analüüsi tulemusena saadud andmed näitasid kvantitatiivselt tüvede jaotust ning võimaldasid hinnata ka potentsiaalseid koinfektsioone, kus ühes proovis esinevad korraga mõlema tüve omadused.

Lisas 3 on välja toodud *gmer_counter*i andmebaasi faili struktuur. Esimene tulp tähistab positsiooni number, teine näitab mitu k-mer tulpa on ja viimased kaks on unikaalsed k-mer'id andmestikus. Samasugused nukleotiidipositsioonid standardiseeritakse ning insertsioonid ja deletsioonid eemaldatakse.

Tulemuste paremaks tõlgendamiseks loodi kohandatud R-skript, mis visualiseeris tüvespetsiifiliste lugemite jaotuse kogu genoomi ulatuses. Graafiline väljund võimaldas esile tuua PVY-N ja PVY-O tüüpide positsioonid ning nendevahelised üleminekud, mis on eriti oluline rekombinantide ja segainfektsioonide tuvastamisel. Selline visualiseerimine andis selge ülevaate, millised genoomipiirkonnad kuulusid kummalegi tüvele ja kuidas need proovis kattusid. Eestis esinevad tüved määrati kasutades 2017. aastal ilmunud artiklis oleva joonise järgi (Joonis 4).

2.2.7. Kartuli Y-viiruse fülogeneetiline analüüs

Kartuli Y-viiruse järjestuste evolutsioonilise päritolu ning tüvedevaheliste fülogeneetiliste seoste määramiseks konstrueeriti fülogeneetiline puu. Analüüsis kasutati Eesti proovide genoomijärjestusi (7 täisgenoomi) ja genoomide järjestusi, mis valiti varasema joondamise etapis (5 genoomi). Lisaks kaasati iga proovi kohta kolm kõige geneetilisemalt sarnast konsensusjärjestust (16 genoomi), et tagada esinduslikkus tüvede vaates ning maksimaalne informatiivsus puu ülesehitamisel. Ülejäänudest mitteesindatud tüvede gruppidest lisati veel kuni 5 genoomi Green jt. (2017) artiklist, kokku 38 isolaati. Fülogeneesi puu koostamiseks kokku kasutati 66. isolaati.

Uurimuses kasutatud proovidest genereeriti täispikad konsensusjärjestused, kasutades Samtools consensus funktsiooni, mis põhines proovide lugemite kaardistusel valitud referentsgenoomide vastu. Need konsensusjärjestused esindavad proovi kõige tõenäolisemat PVY genoomi, võttes arvesse variatsioonide ja kaetuse mustrit kogu järjestuses.

Kõik järjestused, nii oma proovide konsensused kui ka valitud referentsgenoomid, viidi joondamiseks MEGA 11 tarkvarasse, kus kasutati mitmikjoondustööriista MUSCLE vaikeseadetega. MUSCLE võimaldas täpset ja kiiret nukleotiidijärjestuste joondamist, säilitades maksimaalse järjestusliku homoloogilisuse ja optimeerides järgneva fülogeneetilise puu koostamise aluseks oleva andmestiku kvaliteeti.

Joondatud järjestuste põhjal viidi MEGA11-s esmalt läbi mudelivalik, mille tulemusel määrati andmestikule sobivaim evolutsiooniline mudel (GTR+G+I) ning selle põhjal koostati algne fülogeneetiline puu maksimaalse tõepära meetodil (Maximum Likelihood, ML), kasutades 1000 bootstrappi. Kuigi MEGA võimaldas esmase puu koostamist ja mudeli hindamist, viidi lõplik fülogeneetiline analüüs läbi spetsialiseeritud tarkvaras RAxML,

RAxML analüüs viidi läbi kasutades joondatud FASTA-faili, mis sisaldas Eesti konsensusgenoome, joondamiseks kasutatud järjestusi, konsensusjärjestustele geneetiliselt sarnaseid järjestusi ning genoome teistest tüvedest. Analüüsi tulemusena koostati evolutsiooniline puu, mis visualiseeris uuritud proovide fülogeneetilist paigutust PVY tüvede vahel, võimaldades hinnata, millistest referentsrühmadest Eesti proovid geneetiliselt pärinevad.

2.2.8. Taksonoomiline klassifitseerimine RNA-viiruste tuvastamiseks

Lisaks PVY tuvastamisele viidi proovides läbi laiem RNA-viiruste määramine, kasutades Kraken 2 tarkvara. Analüüsi sisendiks võeti puhastatud järjestusandmed FASTQ-formaadis, mis analüüsiti Kraken 2 abil, kasutades Kraken 2 enda viiruste andmebaasi „Viral“ (28.12.2024), mis sisaldas erinevate teadaolevate RNA-viiruste täisgenoomseid järjestusi. Kraken 2 määras iga järjestuslugemi taksonoomilise kuuluvuse ning koostas detailse ülevaate sellest, milliseid RNA-viiruseid proovides esines lisaks PVY-le. Tulemused võimaldasid tuvastada võimalikud koinfektsioonid ja hinnata viiruskoosluste mitmekesisust igas proovis. Tulemuste visualiseerimiseks kasutati Sankey diagrammi (Bogart, 2014).

2.3. Tulemused

2.3.1. Eestis esinevad kartuli Y-viiruse tüved

Tabelis 5 on esitatud RNA-sekveneerimise tulemuste peamised kvantitatiivsed näitajad, mille alusel hinnati kartuliproovide reaalaja RT-qPCR Ct väärtust, PVY viiruse genoomi katvust referentsile joondamisel, kontiigide kvaliteeti ning sekveneerimissügavust. Iga rea kohta on toodud proovi tähistus, sekveneerimisel saadud lugemite koguarv, PVY viiruse järjestustega kaardistunud lugemite arv, kontiigide arv (PVY-ga seonduvad järjestused pärast

assambleerimist), kontiigi keskmine pikkus, keskmine järjestussügavus, genoomi katvuse protsent ning lõpliku viirusgenoomi pikkus.

Lugemite arv oli vahemikus 214 104 ("Eesti-15") kuni 4 222 956 ("Eesti-3"), kuid ainuüksi suur lugemite arv ei olnud määrav PVY genoomi rekonstrueerimise edukuseks. Oluline oli ka see, kui suur osa lugemeid kaardistus PVY referentsgenoomile ning kui täielikult katavad need lugemid viiruse kogu genoomi. Näiteks proovis "Eesti-4" oli küll väga hea katvus (100%), kuid analüüs ei tuvastanud BLAST'i ega ka Kraken2 analüüsis kartuli Y-viirust ning see proov jäeti PVY tüveanalüüsist välja. Kraken2 analüüsis selgus, et proovis esines hoopiski kartuli S-viirus (PVS), millest saadi ka kokku täisgenoom (8445 nukleotiidi).

Tabel 5. Kartuli Y-viiruse genoomijärjestuse rekonstrueerimise tulemused.

Proovi tähistus	Reaalaja RT-qPCR Ct väärtus	Lugemite koguarv	Kaardistunud lugemite arv	Kontiigide arv	Kontiigi keskmine pikkus	Keskmine järjestussügavus (katvus)	Genoomi katvuse protsent (%)	Genoomi pikkus (nt arv)
Eesti-1	29,4	1623000	3	2	147,00	1,79	10,05	2268
Eesti-2	24,6	1591996	388	7	1193,29	7,47	89,50	9196
Eesti-3	27,0	4222956	13	8	169,75	1,84	16,52	7944
Eesti-4	40,0	2037595	3629	1	8445,00	76,80	100,00	8445
Eesti-5	32,3	790924	15	8	126,75	1,44	11,29	8094
Eesti-6	40,0	2357852	82	8	401,75	4,80	33,45	9595
Eesti-7	29,7	1130988	13	9	87,78	1,58	13,17	5240
Eesti-8	28,8	927234	13	8	181,00	1,56	18,48	7686
Eesti-9	19,1	1445967	16120	1	9688,00	337,16	100,00	9688
Eesti-10	12,1	982576	64690	1	9694,00	1416,94	100,00	9694
Eesti-11	19,9	2504611	16388	1	9686,00	273,60	100,00	9686
Eesti-12	11,7	1404645	23171	1	9687,00	337,05	100,00	9687
Eesti-13	15,2	1200310	80140	1	9580,00	1432,95	100,00	9580
Eesti-14	16,5	1067161	2048	1	9675,00	27,79	99,90	9675
Eesti-15	27,6	214104	11	10	70,90	1,09	6,82	8201

PVY tüve määramiseks sobisid proovide hulgast ainult need, kus genoomi katvuse protsent oli piisav (≥99%) ning kontiigi pikkus lähenes PVY eeldatavale täisgenoomi pikkusele (~9600–9700 nukleotiidi). Sellistele kriteeriumitele vastasid proovid "Eesti-9" kuni "Eesti-14", milles PVY genoom rekonstrueeriti terviklikult (katvus 99,9–100%) ja kontiigi pikkused ulatusid 9580 kuni 9694 nukleotiidini. Nende proovide sekveneerimissügavus jäi vahemikku 27,79 – 1432,95.

Ülejäänud proovides, sealhulgas "Eesti-1", "Eesti-2", "Eesti-3", "Eesti-5" kuni "Eesti-8" ning "Eesti-15", oli PVY spetsiifiliste kontiigide katvus madal (alla 90%), kontiigid lühikesed ning genoomi katvus fragmentaarne, mistõttu ei võimaldanud need proovid PVY tüvede usaldusväärset tuvastamist.

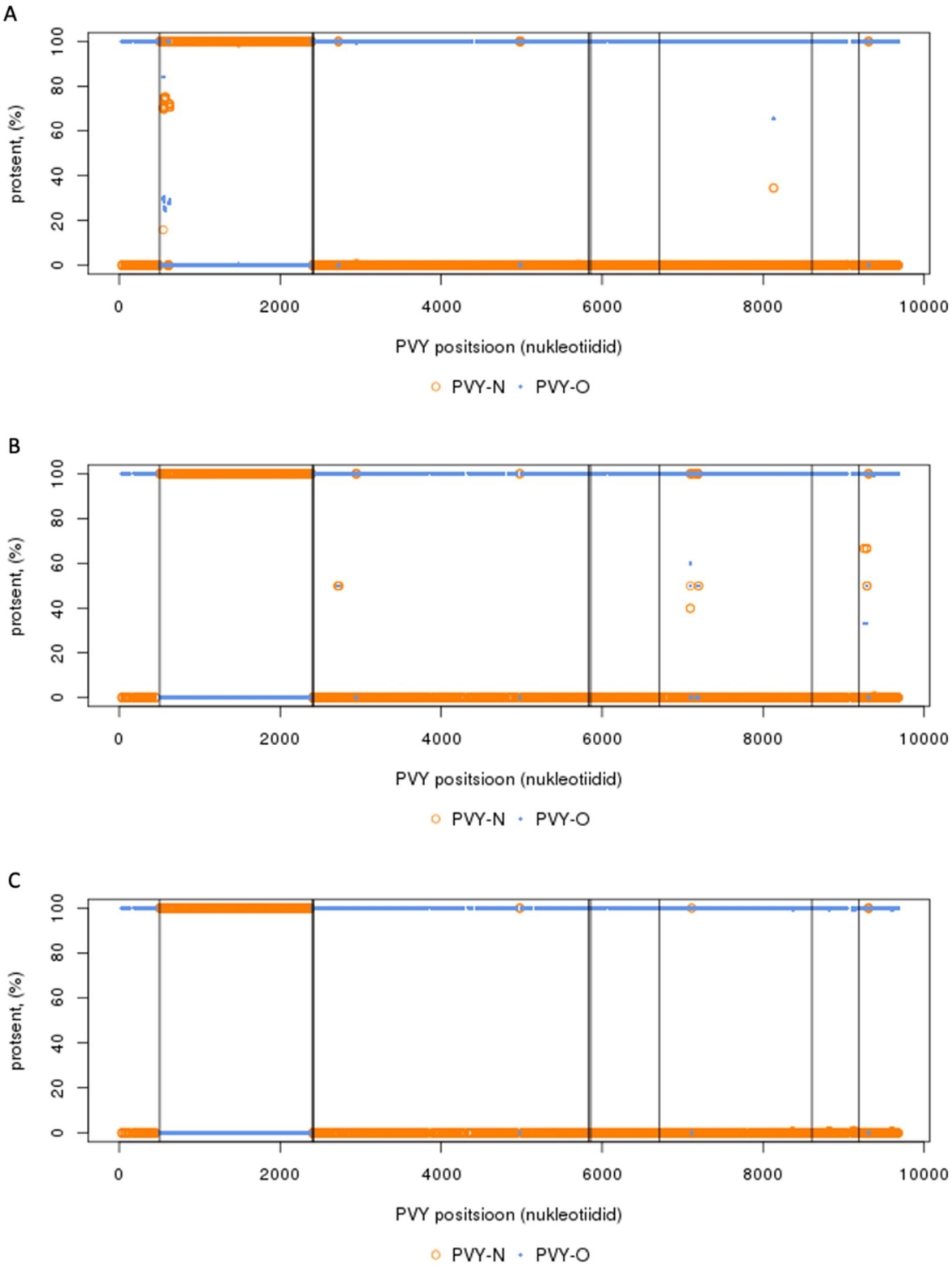
Analüüsi tulemusel selgus, et kuues proovis ("Eesti-9" kuni "Eesti-14") oli viiruslikku RNA-d piisaval hulgal ning andmekvaliteet kõrge, mis võimaldas PVY genoomide täielikku rekonstrueerimist ja usaldusväärset tüvede määramist. Ülejäänud üheksas proovist ei olnud võimalik tuvastada kindlat PVY tüve.

Tüvede määramisel kasutati rekombinatsiooni ja koinfektsiooni analüüsi, mis põhineb kaardistatud järjestuste katvusel teadaolevate PVY tüvede PVY-N ja PVY-O referentsgenoomidega. Tuvastati, millistes genoomipiirkondades esineb suurem vastavus kumbagi tüve tüüpiliste järjestustega ning kui suur osa viirusgenoomist kuulub kummalegi.

Joonisel 7 on kujutatud PVY rekombinatsiooni ja koinfektsiooni analüüsitulemus proovi „Eesti-9“, „Eesti-10“ ja „Eesti-12“ põhjal. Graafik näitab viiruse genoomi positsioonide kaupa PVY-N (oranž) ja PVY-O (sinine) tüvede järjestuse katvust esinemisprotsendi alusel. X-teljel on esitatud PVY genoomi positsioonid (0–10000 nt) ning Y-teljel vastava piirkonna esinemissagedus valimis, väljendatuna protsentides.

Tulemused näitavad selget rekombinatsioonimustrit: PVY-O järjestused domineerivad genoomi 5'-otsa piirkonnas (ligikaudu kuni 500 nukleotiidi), millele järgneb PVY-N 500-2400 nukleotiidi ning seejärel taas PVY-O spetsiifilised järjestused genoomi 3'-otsa ulatuses (alates 2400-9700 nukleotiidi). Selline muster on tüüpiline $PVY^{N\text{-}Wi}$ rekombinanttüvedele.

Proovide puhul viidi sekveneerimisandmete töötlus läbi esmalt nii, et puhastatud lugemid koondati viirusliku genoomi koostamiseks Velvet assembleri abil, kus optimeeriti k-mer pikkust ja katvuse läve sobivaimaks PVY genoomi rekonstrueerimiseks. Velveti väljundist saadi kontiigid. Kontiigi kvaliteeti ja kaetust hinnati Samtools depth käsu abil, mille tulemused kinnitasid keskmiseks katvuse sügavuseks proovil „Eesti-9“ 337,16×, „Eesti-10“ 1416,94× ja „Eesti-12“ 337,05×. Saadud kontiige võrreldi NCBI BLASTn abil andmebaasis olevate PVY järjestustega. Proovil „Eesti-9“ tuvastati 93,22% nukleotiidiline identsus isolaadiga AM113988, mida kasutati edasiseks joondamiseks. Proovidel „Eesti-10“ ja „Eesti-12“ oli vastavalt kõige parem nukleotiidiline identsus isolaatidega OL472145 (identsusega 100%) ja ON604839 (identsusega 100%).



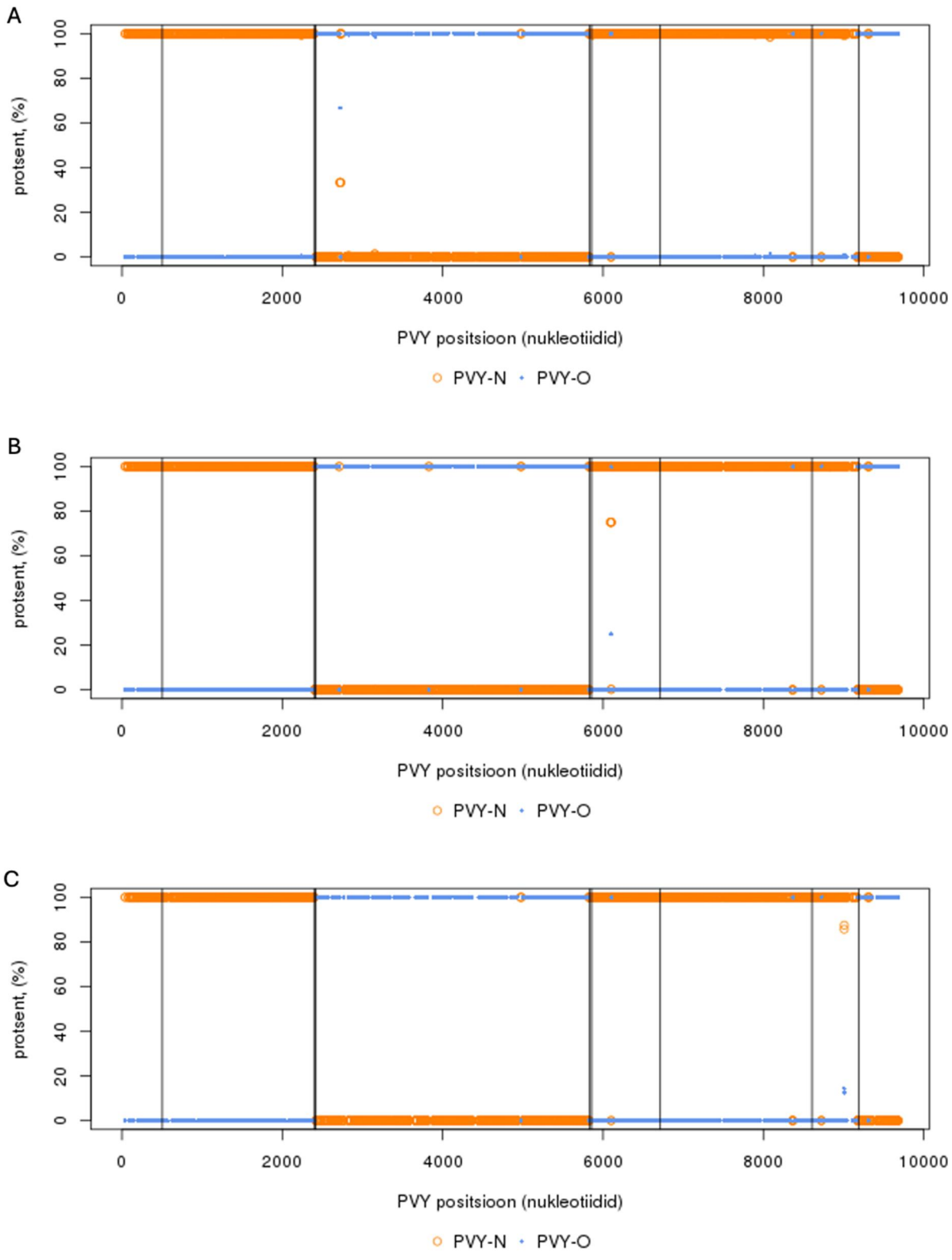
Joonis 7. Proovide „Eesti-9“ (A), „Eesti-10“ (B) ja „Eesti-12“ (C) $PVY^{N\text{-}Wi}$ tüvede määramise graafikud. Vertikaalsed jooned tähistavad joonisel enam levinud rekombinatsiooni kohti. Üksikud punktid joonisel näitavad genoomi olevaid üksikuid mutatsioone.

Täiendavalt ilmnes Samtools tview visualiseerimise käigus, et proov „Eesti-9“ sisaldas tegelikult kahte erinevat, kuid samast tüvest ($PVY^{N\text{-}Wi}$) pärit viirusgenoomi. Selle

koinfektsiooni tuvastamiseks analüüsiti variatsioonipunkte Samtools mpileup abil. Mõnes genoomi positsioonis ilmnesid selged alleelidominantsused, kus üks nukleotiid esines oluliselt sagedamini kui teine. Selline olukord viitab olemasolevale geneetilisele variatsioonile viirusepopulatsioonis, tõenäoliselt kahe lähedase genoomivariandi samaaegsele esinemisele proovis. Antud proovi PVY täisgenoomide kogupikkuseks saadi 9702 nukleotiidi. Proov „Eesti-10“ tulemus SAMtools view'ga näitas, et genoomi 5'-osas esines ulatuslikult määramatuid nukleotiide (N), mis viitas sellele, et avalikus andmebaasis olev järjestus ei sobinud hästi kohaliku isolaadi struktuuriga. Lisaks selgus SAMtools consensus käsuga, et genoomi pikkus ei vasta PVY järjestuse pikkusele. Seejärel BLAST'iti 5' otsast 150. aluspaari pikkune jupp ja saadi tulemuseks, et tegu on taimegeeniga, mitte kartuli Y-viiruse genoomiga. Genoomi pikkuseks saadi 9694 nukleotiidi. „Eesti-12“ proovis visuaalsel analüüsil ilmus ühtne katvus kogu genoomi ulatuses ja tulemuseks saadi üheainsa kontiigina täielik viirusgenoom pikkusega 9687 nukleotiidi.

Joonisel 8 visualiseeritud proovid „Eesti-11“, „Eesti-13“ ja „Eesti-14“ näitavad selgelt rekombinantset ülesehitust, milles domineerivad PVY-N järjestused genoomi 5′-osas (ligikaudu kuni 2400 nukleotiidi), mille järel toimub üleminek PVY-O järjestusele. See muutus on visualiseeritud graafikul oranži (PVY-N) signaali langusena ja sinise (PVY-O) signaali tõusuna 100% juures. Alates ~2400 nukleotiidi positsioonist püsib PVY-O järjestuste esinemine kuni 5800 nukleotiidini. Positsioonil 5800 – 9200 nukleotiidi on esindatud PVY-N ning alates 9300 nukleotiidi kuni genoomi 3′-otsani PVY-O tüüpi järjestused.

Antud joonisel illustreeritud mustri vahetus viitab tugevalt sellele, et tegemist on PVY rekombinanttüvega, mille genoom koosneb kahe erineva tüve PVY-N ja PVY-O segmentidest. Selline graafiku struktuur on iseloomulik $PVY^{NTN(a)}$ tüüpi rekombinanttüvedele.



Joonis 8. Proovide „Eesti-11“ (A), „Eesti-13“ (B) ja „Eesti-14“ (C) $PVY^{NTN(A)}$ tüvede määramise graafikud. Vertikaalsed jooned tähistavad joonisel enam levinud rekombinatsiooni kohti. Üksikud punktid joonisel näitavad genoomi olevaid üksikuid mutatsioone.

Proovide „Eesti-11“, „Eesti-13“ ja „Eesti-14“ sekveneerimisandmete töötluse põhjal määrati keskmiseks katvuseks vastavalt 273,60×, 1432,95× ja 27,79×. Esmase kaardistamise

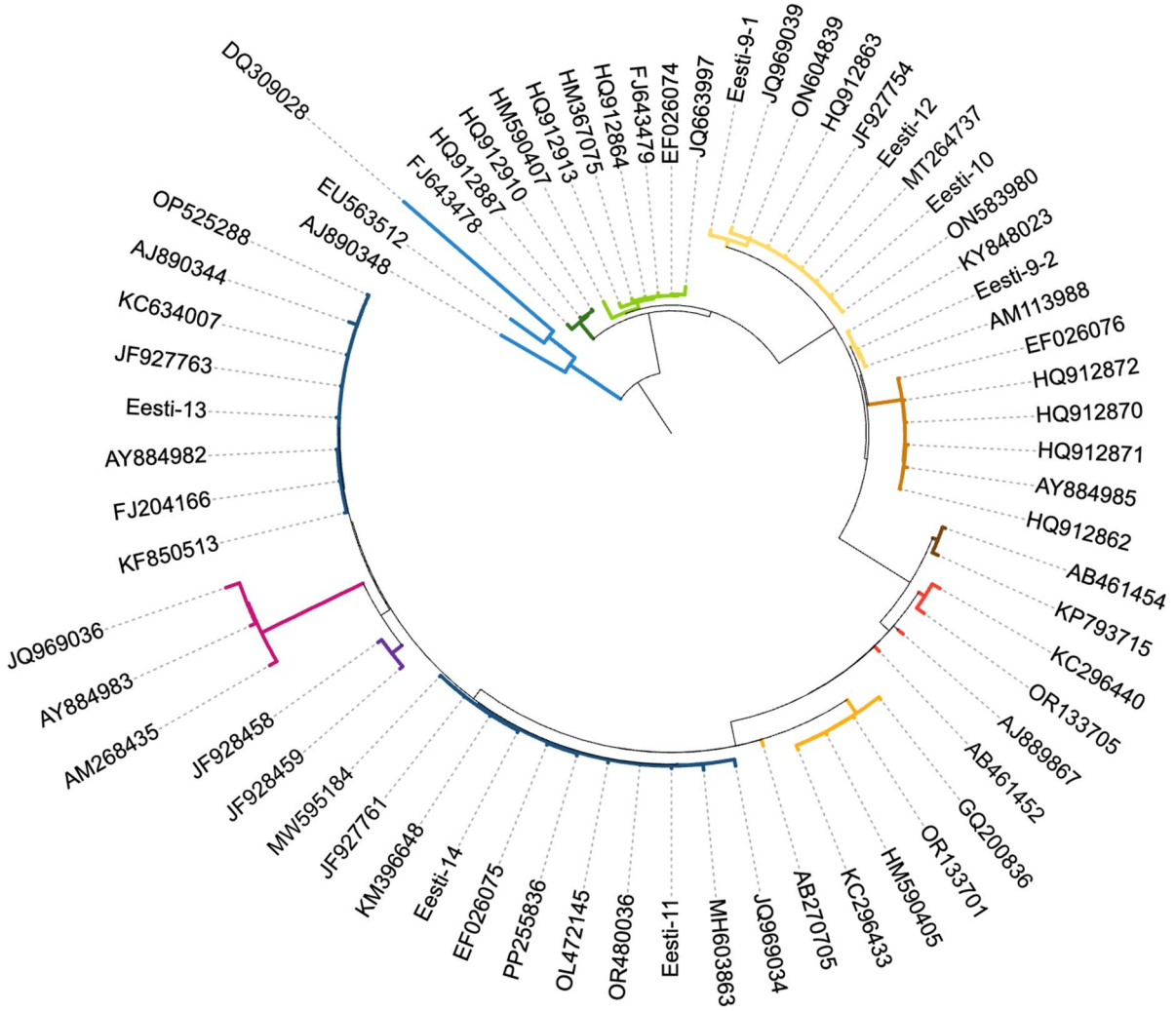
tulemusena saadi parimateks isolaatideks proovil „Eesti-11“ PP255836 (identsusega 100%), proovil „Eesti-13“ MH603863, millega saavutati 90,65% identsus ning proovil „Eesti-14“ PP255836 isolaat identsusega 99,78%. Samtools tview visuaalne inspekteerimine näitas, et proovide „Eesti-11“ ja „Eesti-13“ genoomid oli kaetud ühtlase sügavusega ning ei ilmnenud suuri kõikumisi ega katkestusi kaetuses. Proovi „Eesti-14“ genoomi katvus ei olnud täielik ning Samtools mpileup käsuga läbi viidud variatsioonipunktide analüüs võimaldas eemaldada suurema osa määramatuid nukleotiide, kuid siiski jäid üksikud määramata nukleotiidid (N), mis viitavad lokaalsetele vähese katvusega aladele. Genoomi katvuse protsendiks saadi 99,9%. Konsensuse koostamisel saadi viirusgenoomid pikkusega 9686, 9580 ja 9675 nukleotiidi.

2.3.2. Kartuli Y-viiruse tüvede fülogeneetiline analüüs

Kogu genoomi järjestuste alusel koostatud maksimaalse tõepära meetodil genereeritud fülogeneetiline puu näitab, et kartuli Y-viiruse genoomide hulgas on selgelt eristuvad geneetilised rühmad, mis vastavad enamlevinud PVY tüvedele: C, E, O, O5, Eu-N, N:O, N-Wi, NTN(a), SYR-I, SYR-II ja SYR-III. Tüvede tähistamiseks kasutati erinevaid värvikoode, mis võimaldas visuaalselt jälgida nende fülogeneetilisi suhteid ja geograafilist mustrit. Tüvede harupikkused on välja toodud lisa 4.

Puustruktuur peegeldab PVY genoomide selget jagunemist mitmesse klastrisse (Joonis 9). Fülogeneetiline analüüs viitab sellele, et Eesti kartuli Y-viiruse (PVY) genoomijärjestused on pigem piiratud mitmekesisusega, kuuludes vaid kahe peamise rekombinantse tüve PVYN-Wi ja PVYNTN(a) alla.

Eesti PVYN-Wi tüve isolaadid näivad moodustavat kaks väiksemat rühma. Ühed järjestused on omavahel üsna sarnased ja paiknevad fülogeneetilises puus lähestikku, samas kui teised jäävad sellest veidi kaugemale. Selline jaotus võib viidata kergele sisemisele varieeruvusele, ent üldine geneetiline erinevus jääb väikeseks. PVYNTN(a) tüve esindavad Eesti järjestused paiknevad samuti küllaltki kompaktselt, andes mõista, et nende vahel puuduvad suured genoomilised lahknevused. Kogu populatsiooni ulatuses ei näi Eesti järjestused moodustavat iseseisvaid või tugevalt eristuvaid harusid, vaid pigem liituvad olemasolevate rahvusvaheliste klastritega.



Joonis 9. Kartuli Y-viiruse evolutsioonile päritolu visualiseerituna fülogeneetilisel puul. Sinine värv tähistab C tüve, lilla E tüve, heleroheline O tüve, tumeroheline O5 tüve, pruun N:O tüve, roosa Eu-N tüve, kollane N-Wi tüve, tumesinine NTN(a) tüve, oranž SYR-I tüve, punane SYR-II tüve ja tumepruun SYR-III tüve.

Rahvusvahelises võrdluses asetuvad Eesti PVY isolaadid selliste tüvede sekka, mille läheduses on esindatud järjestused mitmetest erinevatest geograafilistest piirkondadest. „Eesti-10“ ja „Eesti-12“ kuuluvad PVY^{N-Wi} tüvesse, paiknedes samas klastris koos isolaadiga ON583980 (Kazakhstan) ja MT264737 (Iirimaa). „Eesti-9-1“ ja „Eesti-9-2“ kuuluvad samuti rekombinantse PVY^{N-Wi} tüve alla, kuid need on fülogeneetiliselt tihedalt seotud järjestustega JQ969039 (Belgia) ja AM113988 (Saksamaa). „Eesti-11“ paikneb PVY$^{NTN(a)}$ tüves, jagades fülogeneetilist haru järjestustega OR480036 (Venemaa) ja MH603863 (Lõuna-Koera). „Eesti-13“ ja „Eesti-14“ esindavad samuti PVY$^{NTN(a)}$ tüüpi. Need järjestused paiknevad fülogeneetiliselt lähestikku järjestustega JF927763 ja KM396648, mis pärinevad vastavalt Poolast ja Sloveeniast.

2.3.3. Teised kartulit nakatavad taimeviirused Eestis

Sankey diagrammil kujutatud andmestik pärineb 15-st Eesti proovist, mille analüüs viidi läbi taksonoomilise klassifitseerimise tööriistaga Kraken2, kasutades Krakeni „Viral“ andmebaasi (Joonis 10). Igast proovist võeti need viirus- või bakteriofaagitaksonid, mille lugemite arv ületas 100. Diagramm illustreerib nende suhteid, kuvades lugemite jaotust proovide ja tuvastatud organismide vahel, võimaldades visuaalselt hinnata nii kvantitatiivseid erinevusi kui ka taksonilist mitmekesisust proovide lõikes. Iga „Eesti“ proovi järel olev arv tähistab nende järjestuslugemite koguarvu, need arvud ei kajasta kogu proovist saadud järjestusandmestikku (FASTQ lugemite koguarvu), vaid üksnes seda osa, mis leidis taksonoomilise vaste andmebaasis. Vastavus puudub lugemitel, mille järjestused olid kas liiga mittespetsiifilised, pärinesid andmebaasis kajastamata organismidest, või olid liiga lühikesed ja madala kvaliteediga. Seetõttu erineb diagrammil toodud lugemite arv märgatavalt algsete järjestusfailide kogumahust.

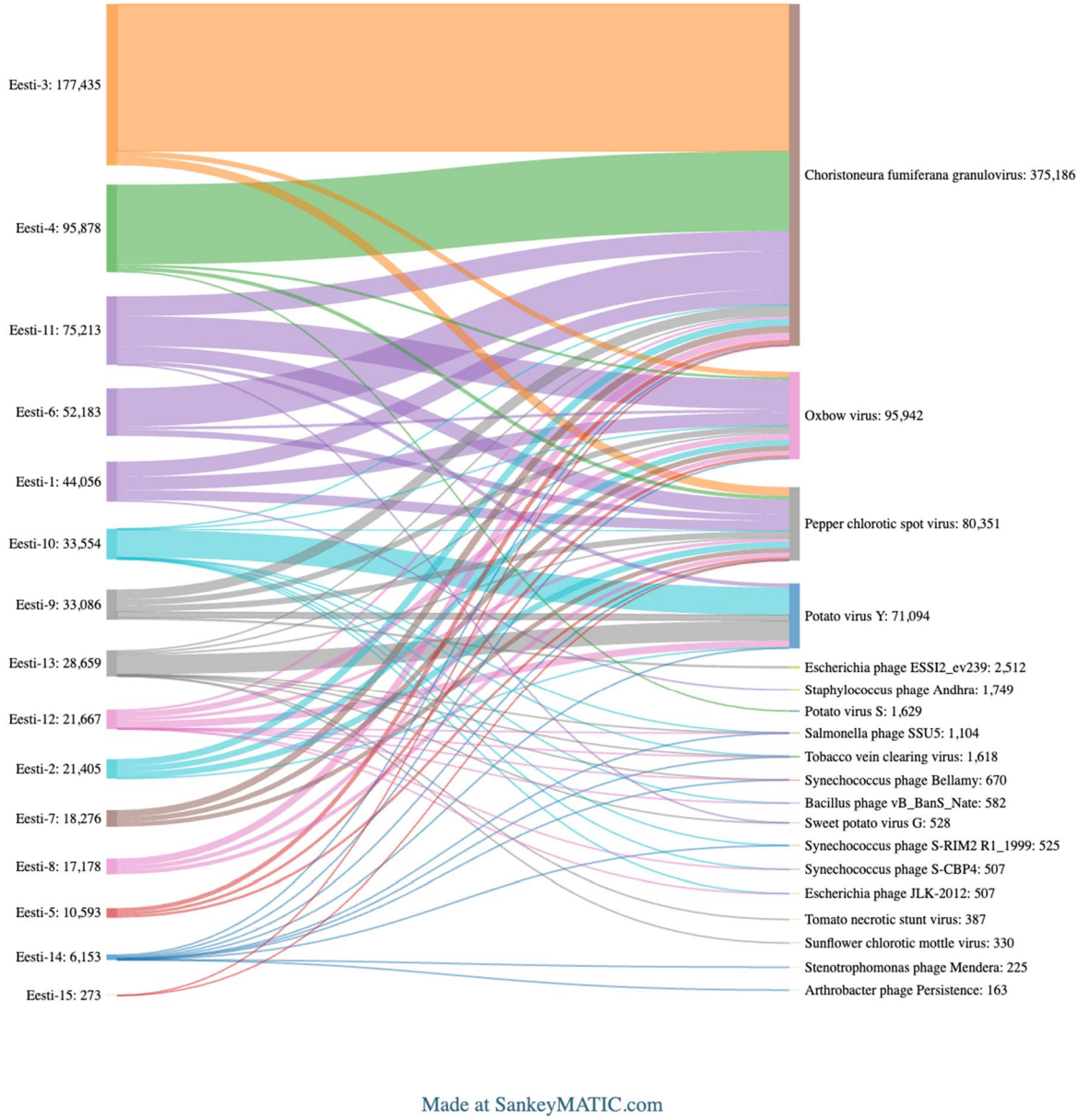
Analüüsi tulemustest ilmnes, et kõige sagedamini ja suurimas mahus esines *Choristoneura fumiferana granulovirus*, mis domineeris eriti tugevalt proovides Eesti-3 ja Eesti-4, vastavalt 177 435 ja 95 878 taksoniga seostatud järjestuslugemiga. Nimetatud viirusjärjestused moodustasid neis proovides selge enamuse kõigist klassifitseeritud lugemitest.

Lisaks granuloviirusele tuvastati mitmes proovis ka teisi viiruseid ja bakteriofaage. *Oxbow virus*, *Pepper chlorotic spot virus* ja *Potato virus Y* olid märkimisväärse esinemissagedusega paljudes proovides. PVY esinemine proovis on loomulik, kuna sekveneerimisele saadeti proovid, kus varasemalt oli tuvastatud PVY. Näiteks Eesti-11, Eesti-6, ja Eesti-1 proovid (lugemite arvuga vastavalt 75 213, 52 183, 44 056) sisaldasid suurel hulgal organime. Sarnast koosseisu täheldati ka proovides Eesti-10 ja Eesti-9.

Bakteriofaagid olid mitmes madalama üldlugemite arvuga esinevaks taksonigrupiks. Näiteks proovides Eesti-1, Eesti-2, Eesti-7, Eesti-8 ja Eesti-12 täheldati bakteriofaagide esinemist, sealhulgas taksonid nagu *Escherichia phage ES12_ev239*, *Salmonella phage SSU5*, *Staphylococcus phage Andhra* ja mitmed *Synechococcus*-liigid.

Proovid Eesti-5, Eesti-14 ja Eesti-15 sisaldasid võrreldes teistega märkimisväärselt vähem taksonoomiliselt määratud lugemeid – vastavalt 10 593, 6 153 ja 273. Eesti-5 proovis tuvastati vähesel määral taimeviiruseid, samas kui Eesti-14 proovis domineerisid keskkonnaspetsiifilised taksonid, näiteks *Synechococcus*’e faagid. Eesti-15 puhul oli tegemist kõige madalama viiruskoormusega prooviga, kus esines üksikuid järjestusi. Selle alusel ei saa

teha sisulisi järeldusi viiruskoosseisu kohta, kuid võimalik on oletada madalat viiruserikkust või vähest kontaminatsiooni.



Joonis 10. „Eesti 1“ – „Eesti 15“ proovides esinevad organismid kujutatuna Sankey diagrammil.

Diagrammil kujutatud andmestik näitab märkimisväärset varieeruvust taksonilises koostises Eesti metagenoomsete proovide lõikes. *Choristoneura fumiferana granulovirus* osutus kõige sagedamini esinevaks ja selle viiruse järjestusi tuvastati enim, kuid mitmetes proovides esinesid ka kompleksemad kooslused, kus viiruslikud ja bakteriofaagide järjestused jagasid omavahelise kvantitatiivse osakaalu. Levinumad viirused, nagu *Oxbow virus*, *Pepper chlorotic spot virus* ja *Potato virus Y*, olid esindatud mitmetes proovides, kuid varieeruva osakaaluga. Lisaks oli huvitavaks leiuks taimeviirus *Tobacco vein clearing virus*, mida sisaldasid proovid „Eesti-10“, „Eesti-12“-„Eesti 14“. Samas oli bakteriofaagide levik ulatuslik, eriti väiksema

viiruskoormusega proovides, viidates mikroobikoosluste aktiivsele rollile ja prokariootsete peremeeste laialdasele olemasolule keskkonnas. Need erinevused peegeldavad nii ökoloogilisi kui ka võimalikke geograafilisi eripärasid, mis kujundavad viirus- ja faagikooslusi looduslikes metagenoomilistes proovides.

2.4. Arutelu

Käesoleva magistritöö eesmärk oli uurida Eestis esinevate kartuli Y-viiruse (PVY) tüvede geneetilist mitmekesisust, sh nende järjestuste kokkupanekut, rekombinantsete vormide võimalikku esinemist, fülogeneetilist päritolu ning ka teiste RNA-viiruste olemasolu kartuliproovides, kasutades metagenoomset järjestust ja bioinformaatilisi tööriistu. Tulemused kinnitavad, et PVY populatsioon Eestis on geneetiliselt mitmekesine, sisaldades mitmeid erinevaid liine ja rekombinantseid mustreid. Lisaks näitas metagenoomne analüüs, et kartuliproovid sisaldavad mitmesuguseid teisi viraalseid järjestusi ja bakteriofaagide järjestusi, mis viitab proovide kõrgele mitmekesisusele ning metoodika tundlikkusele.

Esimese eesmärgi täitmisel koostati PVY täielik genoom kuues proovis („Eesti-9-1“ – „Eesti-14“). Kõige parema tulemuse andis proov „Eesti-10“, milles rekonstrueeriti PVY genoom üle 9694 nukleotiidi pikkusena, kõrge katvusega (1416,94×). Järjestusanalüüs näitas, et see genoom kuulus $PVY^{N\text{-}Wi}$ rekombinanttüve hulka, mida iseloomustab kolmes piirkonnas esinev järjestuse päritolu vaheldumine PVY-N ja PVY-O tüvede vahel. Sellise rekombinantse struktuuriga PVY genoom on kirjanduses hästi dokumenteeritud – näiteks Green jt. (2017) on kirjeldanud täpselt samasugust rekombinatsioonimustrit rekombinantses tüves nagu $PVY^{N\text{-}Wi}$, millel on oluline mõju nende levikuvõimele ja virulentsusele.

Teistes proovides, nagu „Eesti-1“ – „Eesti-8“ ja „Eesti-15“, väljaarvatud „Eesti-4“, oli võimalik kokku panna lühemaid kontiige, mis joondusid PVY genoomile, kuid täielikku genoomi ei õnnestunud saada. Selgus, et suure üldise järjestuslugemite arvu olemasolu ei taga iseenesest PVY katvust – olulisem oli spetsiifiliste PVY-seostuvate lugemite arv, mis sõltus viiruse tegelikust kogusest proovis ning RNA kvaliteedist. Võrreldes kirjandusega, kus näiteks Quenouille jt. (2013) kirjeldavad PVY genoomide suurt geneetilist varieeruvust ja varieeruvat levikut sõltuvalt geograafiast ja rekombinatsioonist näitab käesolev töö, et Eestis on levinud peamiselt kaks genoomi tüüpi: $PVY^{NTN(a)}$ ja $PVY^{N\text{-}Wi}$.

Töö tulemused näitasid selget seost reaalaja RT-qPCR analüüsis saadud tsüklite arvu (Ct väärtuse) ja PVY täisgenoomide edukalt kokkupandud järjestuste vahel. Kõigis proovides, millel oli Ct väärtus alla 20, õnnestus PVY täielik genoom rekonstrueerida kõrge katvusega.

See viitab sellele, et Ct väärtus <20 on usaldusväärne indikaator viiruse piisava kontsentratsiooni kohta proovis, mis on sobiv järjestamiseks ja edasiseks bioinformaatiliseks analüüsiks. Kuigi enamik täisgenoome saadi lehematerjalist, õnnestus kaks täielikku PVY genoomi kokku panna ka seemnemugulatest pärit proovist „Eesti-9", mille madalad Ct väärtused kinnitasid piisavat viiruse sisaldust. Seega on reaalaja RT-qPCR oluline etapp, mille abil saab hinnata metagenoomseks järjestamiseks sobivaid proove ning suurendada täisgenoomide rekonstrueerimise tõenäosust.

Teise eesmärgi raames hinnati Eestis esinevate PVY tüvede rekombinantset olemust ja tüübilist kuuluvust, kasutades kuut proovi, millest õnnestus rekonstrueerida viiruse täisgenoomid. Kõigis kuues proovis oli genoomne katvus piisav, et määrata täpselt rekombinatsioonimustrid. Analüüs näitas, et järjestuste üleminekualad PVY-N ja PVY-O päritolu vahel kattusid hästi varem dokumenteeritud rekombinatsioonipiirkondadega. Kõik kuus järjestust vastasid teadaolevatele rekombinantsete PVY tüüpide mustritele, võimaldades igale proovile omistada konkreetse tüve klassifikatsiooni. Tulemused kinnitavad, et Eestis ringlevad peamiselt rekombinantsed PVY vormid, mille genoomne ülesehitus sarnaneb laialt levinud tüüpidega, nagu $PVY^{NTN(a)}$ ja $PVY^{N\text{-}Wi}$. See on kooskõlas rahvusvahelise kirjandusega, kus sellised rekombinandid on viimasel aastakümnel kujunenud domineerivateks tüvedeks paljudes Euroopa ja lähiregioonide PVY populatsioonides (Karasev ja Gray, 2013; Quenouille jt., 2013). Kuna kõigi kuue täisgenoomi järjestus andis selge ja usaldusväärse rekombinatsioonimustri, ei viidanud tulemused uute seni kirjeldamata rekombinantide olemasolule. Seevastu kinnitati, et Eestis levivad PVY tüved esindavad teadaolevaid rekombinantseid liine, mille esinemine viitab kas varasemale sissetoomisele või lokaalsele püsivusele. Antud töös koostatud genoomid ja nende rekombinatsioonianalüüs on esmakordne Eestis esinevate tüvede põhjalik identifitseerimine. Töö tulemused toetavad arusaama, et PVY geneetiline varieeruvus Eestis on kujunenud vastavuses rahvusvaheliste levikumustritega ning ei viita hetkel uute rekombinantsete PVY variantide esinemisele kohalikus populatsioonis.

Kolmanda uurimiseesmärgi täitmiseks viidi läbi fülogeneetiline analüüs, milles käsitleti seitsme Eesti PVY isolaadi „Eesti-9-1", „Eesti-9-2", „Eesti-10", „Eesti-11", „Eesti-12", „Eesti-13" ja „Eesti-14" kogu genoomi järjestusi. Maksimaalse tõepära meetodil koostatud fülogeneetiline puu võimaldas hinnata nende järjestuste omavahelisi seoseid ning asetust globaalses kontekstis. Tulemused viitavad sellele, et kõik analüüsitud Eesti isolaadid kuuluvad vaid kahe teadaoleva rekombinantse päritoluga tüve $^{PVYN\text{-}Wi}$ ja $PVY^{NTN(a)}$ alla, mis võib osutada pigem piiratud tüvede esinemisele uuritud valimis. $PVY^{N\text{-}Wi}$ ja $PVY^{NTN(a)}$ on Eesti isolaatide vahel siiski mõningast varieeruvust. $PVY^{N\text{-}Wi}$ tüüpi järjestuste puhul näib esinevat kaks

väiksemat gruppi: osa järjestusi paiknevad omavahel lähestikku, mis võib viidata geneetilisele sarnasusele, samas kui teised jäävad veidi kaugemale, andes alust oletada väikest sisemist mitmekesisust selles rühmas. $PVY^{NTN(a)}$ tüüpi järjestused paistavad olevat omavahel sarnasemad, mis võib viidata väiksemale genoomilisele erinevusele nende vahel. Eesti isolaadid ei moodusta seejuures eraldiseisvat või selgelt eristuvat haru, vaid näivad liituvat pigem olemasolevate rahvusvaheliste klastritega. Selline muster võib viidata sellele, et Eesti PVY järjestused jagavad geneetilisi jooni isolaadidega, mida on varem tuvastatud erinevates maailma piirkondades. Eesti PVY^{N-Wi} järjestused paiknevad lähestikku järjestustega, mida on leitud Euroopast. $PVY^{NTN(a)}$ tüve Eesti järjestused jagavad harusid varem Venemaal ja Lõuna-Koreas tuvastatud järjestusega. Taolisi võimalikke levikumustreid toetavad ka varasemad uuringud, mis on rõhutanud PVY leviku seotust rahvusvahelise seemnekartuli kaubanduse ning mehaanilise levikuga inimtegevuse kaudu (Adams jt., 2009; Tsedaley, 2015). Selliseid seoseid on täheldatud ka piirkondades, kus rekombinantsed PVY vormid on saanud domineerivaks ja viiruste liikumine on olnud aktiivne (Karasev ja Gray, 2013). Kuid siiski peab rõhutama, et need tulemused viitavad võimalusele, et Eesti PVY järjestuste kujunemisele on mõjutanud mitmesuunaline levik, kus sarnased rekombinantsed viirusvormid võivad olla Eestisse jõudnud erinevate kanalite kaudu. Samas tuleb rõhutada, et antud järeldused põhinevad piiratud arvu järjestuste analüüsil ning ei võimalda teha lõplikke üldistusi Eesti PVY populatsiooni kohta. Tulemuste kinnitamiseks ja mitmekesisuse ulatuse hindamiseks oleks vajalik läbi viia laiaulatuslikum uuring, mis hõlmaks suuremat arvu proove eri geograafilistest piirkondadest ja aastatest.

Neljas eesmärk keskendus HTS'i hindamisele erinevate seemnekartulit nakatavate taimeviiruste ja viiruselaadsete osakeste esinemise määramisele RNA-Seq andmete taksonoomilise klassifitseerimise alusel. Tulemuste põhjal ilmnes, et lisaks PVY-le oli Eesti proovides rohkelt teiste organismide lugemeid, sealhulgas *Choristoneura fumiferana granulovirus*, *Oxbow virus* ja *Pepper chlorotic spot virus*. *Choristoneura fumiferana granulovirus* on peremeesspetsiifiline putukaviirus, mille esinemine taimematerjalis on tõenäoliselt seotud proovidesse sattunud keskkonna DNA või RNA jääkidega, näiteks taimepinnal või koe sees elanud putukate RNA jääkidega (Forté jt., 1999; Bah jt., 1999). Sarnaseid juhtumeid, kus loomse või putukse päritoluga viirused domineerivad taimsetes metagenoomiproovides, on kirjeldatud ka varasemates metagenoomilistes uuringutes, eriti kui sekveneerimiseks kasutatakse kogu RNA-d ilma taimespetsiifiliste fraktsioonide rikastamiseta (Wood jt., 2019). Samuti tuvastati *Oxbow virus järjestusi*, see viirus on teadaolevalt seotud närilistega ja isoleeritud Ameerika Ühendriikides *Oryzomys palustris* isenditelt (Kang jt.,

2009). Kuna nende näol ei ole tegemist taimeviirustega, siis nende edasist analüüsi ei teostata. *Pepper chlorotic spot virus*, mida Zheng jt. (2017) kirjeldasid uue isolaatina Hiinas Yunnani provintsis, põhjustab süsteemseid sümptomeid *Solanaceae* taimedes, nagu paprika ja tomat. Teaduskirjanduse põhjal ei saa väita, et antud viirus oleks seotud kartuliga, sest seda ei ole kartulil varasemalt dokumenteeritud ega ka kirjeldatud. Selle viiruse esinemine kartuliproovides Eestis ei viita tõenäoliselt tegelikule nakatumisele, vaid pigem RNA jääkainetele.

Tuleb siiski rõhutada, et kuigi Kraken2 klassifikatsioonitulemused näitasid mõnede nimetatud viiruste puhul suurt lugemite arvu, ei joondu need lugemid vastava viiruse referentsgenoomidele. Seetõttu ei saa ainult nende klassifitseerimiste põhjal järeldada tegelikku infektsiooni olemasolu. Need leiud toetavad kirjanduses esitatud seisukohti, mille kohaselt metagenoomiline sekveneerimine, eriti kogu RNA põhjal, toob esile väga laia taksonoomilise spektriga viirusjärjestusi, millest paljud ei ole peremeestaimedega seotud (Menzel jt., 2016; Adams jt., 2009). Seetõttu on selliste tulemustega töötades äärmiselt oluline kriitiline tõlgendamine, kus eristatakse tegelikku bioloogilist tähendust potentsiaalsest tehnilisest või keskkondlikust müraallikast.

Taimede tegelike patogeenide seas esines andmestikus ka viiruseid, nagu kartuli S-viirus ja *Tobacco vein clearing virus*, mis on olulised ja otseselt seotud kartuli nakatamisega. Aastal 2023 leidsid Ewan jt. samuti, et varasemalt mitte kirjeldatud kui kartulit nakatav viirus *Tobacco vein clering virus* esines neil proovis. Lisaks tuvastati mitmeid bakteriofaage, sealhulgas *Salmonella phage SSU5* ja *Synechococcus phage Bellamy*. On hästi teada, et faagid metagenoomsetes RNA-sekveneerimise andmestikes võivad toimida ökoloogiliste koosluste markeritena, eriti madala viiruskoormusega keskkondades (Paez-Espino jt., 2016; Roux jt., 2015).

Magistritöö tulemustest saab järeldada, et Eesti PVY populatsioonis on esindatud kaks rekombinantset genoomistruktuuri ($PVY^{NTN(a)}$ ja PVY^{N-Wi}), mis on sarnased rahvusvaheliselt levinud tüvedele. Tüvede fülogeneetiline jaotus peegeldab tugevaid seoseid nii Euroopa kui ka Lähis-Ida ja Lõuna-Aasia viiruspopulatsioonidega, viidates globaalsele levikuvõrgustikule. Lisaks PVY-le esines metagenoomsetes proovides märkimisväärne hulk teiste viiruste ja bakteriofaagide järjestusi, mis osaliselt võivad pärineda keskkonnast või olla määratud andmebaaside k-mer katvuste põhjal. Seetõttu on oluline, et taksonoomilise määramise tulemusi tõlgendataks koos bioloogilise konteksti ja proovide töötlusmeetoditega.

Kraken2 ja RNA-seq põhine metagenoomne analüüs osutus sobivaks tööriistaks viiruskoosluste määramiseks, kuid vajab alati toetavat valideerimist. Töö toob esile vajaduse laiaulatuslikuks viiruste seireks kartulikasvatuses ning rõhutab molekulaarse analüüsi rolli viiruse evolutsiooni ja leviku mõistmisel. Eesti PVY populatsiooni ja üldise viroomi mõistmine annab olulise sisendi taimede tervise jälgimisele ja haiguste ennetusele ka laiemas põllumajanduslikus kontekstis.

KOKKUVÕTE

Antud magistritöö oli jätk varasemale bakalaureusetööle (Stražev, 2023). Töö eesmärk oli uurida Eestis esinevate kartuli Y-viiruse (PVY) tüvede geneetilist mitmekesisust, hinnata nende rekombinantset olemust, fülogeneetilist päritolu ning määrata teiste RNA-viiruste ja viiruselaadsete osakeste esinemist kartuliproovides, kasutades RNA-sekveneerimist ja bioinformaatilisi analüüsimeetodeid.

Analüüsi käigus tuvastati kuues proovis PVY täisgenoomid millest selgus, et Eestis ringlevad peamiselt kaks teadaolevat rekombinantset tüve: $PVY^{NTN(a)}$ ja PVY^{N-Wi}. Genoomide rekombinatsioonimustrid vastasid varem kirjeldatud musterpiirkondadega, viidates nende tüvede rahvusvahelisele päritolule ja püsivusele lokaalses populatsioonis.

Fülogeneetilise analüüsi tulemused näitasid, et seitsme Eesti PVY isolaadi genoomijärjestused kuuluvad peamiselt kahe rekombinantse tüve PVY^{N-Wi} ja $PVY^{NTN(a)}$ alla. Eesti isolaatide vahel esines mõningast geneetilist varieeruvust, eriti PVY^{N-Wi} tüve puhul, kus eristusid kaks väikest gruppi, samas kui $PVY^{NTN(a)}$ järjestused olid omavahel suhteliselt sarnased. Eesti isolaadid ei moodustanud eraldiseisvat haru, vaid paiknesid koos teiste rahvusvaheliste järjestustega, mis viitab nende geneetilisele seotusele eri maailma piirkondade viirusvormidega. Tulemused põhinevad siiski piiratud valimil ning laiemate järelduste tegemiseks on vajalik suurem ja mitmekesisem proovide kogum. Samuti kinnitas uuring, et genoomi täielikuks rekonstrueerimiseks on kriitilise tähtsusega algmaterjali kvaliteet ja selles sisalduva viiruse kontsentratsioon. RT-qPCR analüüside Ct väärtused alla 20 osutusid usaldusväärseks indikaatoriks, mille alusel oli võimalik edukalt viia läbi genoomne järjestamine ja PVY täisgenoomi koostamine.

Taksonoomilise klassifitseerimise käigus tuvastati lisaks PVY-le ka mitmeid teisi viiruslikke ja bakteriofaagide järjestusi, sealhulgas *Choristoneura fumiferana granulovirus*, *Oxbow virus* ja *Pepper chlorotic spot virus*. Kuigi Kraken2 klassifikatsioon andis antud viiruste jaoks suure lugemite arvu, siis joondamine nende vastavate referentsgenoomidega ei andnud head katvus kogu genoomi ulatuses, mistõttu ei saa nende esinemist tõlgendada kui tõelist infektsiooni. Tulemused rõhutavad vajadust kriitiliseks tõlgenduseks.

Magistritöö tulemused kinnitavad, et metagenoomne järjestamine koos bioinformaatilise töötlusega on võimas tööriist PVY ja muude viiruste tuvastamisel, kuid selle rakendamine eeldab hoolikat andmete tõlgendamist ning täiendavat valideerimist. Töö annab olulise panuse Eesti PVY populatsiooni iseloomustamisse ja viiruskoosluste mõistmisse kartulikasvatuses,

kuid lõplike järelduste hindamiseks oleks vaja läbi viia edasised uurimised, et kinnitada saadud tulemusi.

SUMMARY

This master's thesis is a continuation of the earlier bachelor's thesis (Stražev, 2023). The aim of the work was to investigate the genetic diversity of potato virus Y (PVY) strains occurring in Estonia, assess their recombinant nature, phylogenetic origin, and identify the presence of other RNA viruses and virus-like particles in potato samples using RNA sequencing and bioinformatics analysis methods.

During the analysis, full PVY genomes were detected in six samples, revealing that mainly two known recombinant strains circulate in Estonia: $PVY^{NTN(a)}$ and PVY^{N-Wi}. The recombination patterns of the genomes corresponded to previously described hotspot regions, indicating the international origin of these strains and their persistence in the local population.

The results of the phylogenetic analysis showed that the genome sequences of seven Estonian PVY isolates mainly belong to two recombinant strains, PVY^{N-Wi} and $PVY^{NTN(a)}$. Some genetic variation was observed between the Estonian isolates, especially within the PVY^{N-Wi} strain, where two smaller groups could be distinguished, whereas the $PVY^{NTN(a)}$ sequences were relatively similar to each other. The Estonian isolates did not form a distinct clade but clustered with other international sequences, indicating their genetic relatedness to virus variants found in different regions worldwide. However, these results are based on a limited sample size, and broader conclusions require a larger and more diverse set of samples.

The study also confirmed that the quality of the starting material and the viral concentration it contains are critical for complete genome reconstruction. RT-qPCR analyses with Ct values below 20 proved to be a reliable indicator, enabling successful genome sequencing and full PVY genome assembly.

During taxonomic classification, several other viral and bacteriophages sequences were identified in addition to PVY, including *Choristoneura fumiferana granulovirus*, *Oxbow virus*, and *Pepper chlorotic spot virus*. Although Kraken2 classification yielded a high number of reads for these viruses, they did not cover their respective reference genomes well, so their presence cannot be interpreted as a true infection. These findings emphasize the need for critical interpretation.

The results of the master's thesis confirm that metagenomic sequencing combined with bioinformatics processing is a powerful tool for detecting PVY and other viruses but requires careful data interpretation and further validation. This work provides an important contribution

to the characterization of the Estonian PVY population and the understanding of viral communities in potato cultivation. However, further investigations are needed to confirm the results and to support definitive conclusions.

TÄNUSÕNAD

Soovin tänada oma juhendajaid Piret van der Sman ja Tarmo Puurand, kes tegid võimalikuks antud töö läbi viimise. Tänusõnad lähevad veel ka Maaelu Teadmuskeskuse Taimetervise ja mikrobioloogia valdkonna laboratooriumile ja selle töötajatele, kes antud magistritöös viisid läbi kartulist tehtavad analüüsid. Tänan Tartu Ülikooli, kelle tehnilised ressursid võimaldasid teha bioinformaatilist andmeanalüüsi. Lisaks tänusõnad FERA'le, kes viis läbi sekveneerimise ja kellelt pärinevad toorandmed.

KASUTATUD KIRJANDUS

Adams, I. P., Glover, R. H., Monger, W. A., Mumford, R., Jackeviciene, E., Navalinskiene, M., Samuitiene, M., Boonham, N. (2009). Next-generation sequencing and metagenomic analysis: a universal diagnostic tool in plant virology. – *Molecular Plant Pathology*. Nr. 10(4), 537–545. https://doi.org/10.1111/j.1364-3703.2009.00545.x (06.05.2025)

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., Lipman, D. J. (1990). Basic local alignment search tool. – *Journal of Molecular Biology*. Nr. 215(3), 403–410. https://doi.org/10.1016/S0022-2836(05)80360-2 (07.05.2025)

Altschul, S. F., Madden, T. L., Schäffer, A. A., Zhang, J., Zhang, Z., Miller, W., Lipman, D. J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. – *Nucleic Acids Research*. Nr. 25(17), 3389–3402. https://doi.org/10.1093/nar/25.17.3389 (07.05.2025)

Angua-nf. (2024). *Fera Bioinformatics GitLab Repository*. Fera Science Ltd. https://gitlab.fera.co.uk/fera_bioinformatics/genomics/angua-nf (27.05.2025)

Aphids as Crop Pests. (2017). /Ed. H. F. van Emden, R. Harrington. Wallingford: CAB International. [on-line] ebrary https://books.google.ee/books?hl=en&lr=&id=tH46DwAAQBAJ&oi=fnd&pg=PA323&dq=transmission+of+plant+viruses&ots=QAZ3qMxlyq&sig=YR1v2_yOduNOk92qWqqoU0y8Fwk&redir_esc=y#v=onepage&q=transmission%20of%20plant%20viruses&f=false (25.04.2025)

Awe, O. I., En najih, N., Nyamari, M. N., Mukanga, L. B. (2023). Comparative study between molecular and genetic evolutionary analysis tools using African SARS-CoV-2 variants. – *Informatics in Medicine Unlocked.* Nr. 36, 101143. https://www.sciencedirect.com/science/article/pii/S2352914822002805 (26.05.2025)

Bah, A., Lucarotti, C. J., Arella, M., Guertin, C. (1999). *Choristoneura fumiferana* granulovirus: sequence analysis and 5′ characterization of ORF891. – *Archives of Virology, 144*(4), 737–746. https://doi.org/10.1007/s007050050539 (20.05.2025)

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., Lesin, V. M., Nikolenko, S. I., Pham, S., Prjibelski, A. D., Sirotkin, A. V., Vyahhi, N., Tesler, G., Alekseyev, M. A., Pevzner, P. A. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. – *Journal of Computational Biology*. Nr. 19(5), 455–477. https://doi.org/10.1089/cmb.2012.0021 (07.05.2025)

Beals, K. A. (2019). Potatoes, nutrition and health. – *American Journal of Potato Research*. Nr. 96(2), 102–110. https://doi.org/10.1007/s12230-018-09705-4 (25.04.2025)

Bentley, D. R., Balasubramanian, S., Swerdlow, H. P., Smith, G. P., Milton, J., Brown, C. G., Hall, K. P., Evers, D. J., Barnes, C. L., Bignell, H. R., Boutell, J. M., Bryant, J., Carter, R. J., Cheetham, R. K., Cox, A. J., Ellis, D. J., Flatbush, M. R., Gormley, N. A., Humphray, S. J., Irving, L. J., ... Smith, A. J. (2008). Accurate whole human genome sequencing using reversible terminator chemistry. – *Nature.* Nr. 456, 53–59. https://www.nature.com/articles/nature07517 (27.05.2025)

Bioreba AG. (2023). *ELISA test procedure: English, French, German, Spanish and Polish* [PDF]. https://www.bioreba.ch/saas/CustomUpload/374O357O340O370O356O369O350O321O360O366O369O356O353O352O350O320O326O/ELISA_Test_procedure_e_f_d_es_und_pl_120523.pdf (27.05.2025)

Biotechnology in Agriculture and Forestry 3. (1987)./Ed. Y. P. S. Bajaj. Springer-Verlag Berlin Heidelberg. [on-line] ebrary (24.05.2025)

Birch, P. R. J., Bryan, G., Fenton, B., Gilroy, E. M., Hein, I., Jones, J. T., Prashar, A., Taylor, M. A., Torrance, L., Toth, I. K. (2012). Crops that feed the world 8: Potato: Are the trends of increased globaal production sustainable?. – *Food Security*. Nr. 4.

https://research-portal.st-andrews.ac.uk/en/publications/crops-that-feed-the-world-8-potato-are-the-trends-of-increased-gl (24.05.2025)

Bogart, S. (2014). *SankeyMATIC: Build a Sankey diagram*. https://sankeymatic.com/ (19.05.2025)

Boonham, N., Kreuze, J., Winter, S., Van der Vlugt, R., Bergervoet, J.,Tomlinson, J., Mumford, R. (2014). Methods in virus diagnostics: from ELISA to next generation sequencing. – *Virus Res*. Nr. 186, 20–31. https://doi.org/10.1016/j.virusres.2013.12.007 (06.05.2025)

Brunt, A. A., Grabtee, K., Dallwitz, M. J., Gibbs, A. J., Watson, L. (1996). Viruses of Plants. New York: CABI Publishing. 1034 lk. (25.04.2025)

Buchfink, B., Xie, C., Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. – *Nature Methods*. Nr. 12, 59–60. https://doi.org/10.1038/nmeth.3176 (06.05.2025)

Bustin, S. A. (2005). Real-time reverse transcription PCR. – *In Encyclopedia of Diagnostic Genomics and Proteomics* (p. 1131). https://www.gene-quantification.de/bustin-qrt-pcr-review-2005.pdf (26.05.2025)

Bustin, S. A. (2007). RT-qPCR and molecular diagnostics: No evidence for measles virus in the GI tract of autistic children. *Anglia Ruskin University*. https://www.researchgate.net/publication/242579832_RT-qPCR_and_molecular_diagnostics_no_evidence_for_measles_virus_in_the_GI_tract_of_autistic_children (26.05.2025)

Cambio. (s.a). Lab Reagents: Smart Buffers and Reagents. [veebileht] https://www.cambio.co.uk/1283/77/products/potassium-acetate/ (06.05.2025)

Chan, M.-S., Jeger, M. J. (1994). An Analytical Model of Plant Virus Disease Dynamics with Roguing and Replanting. – *Journal of Applied Ecology*. Nr. 3. [e-ajakiri] https://www.jstor.org/stable/pdf/2404439.pdf (25.04.2025)

Chikh Ali M., Maoka T., Natsuaki K. T. Natsuaki T. (2010). The simultaneous differentiation of Potato virus Y strains including the newly described strain PVY^{NTN-NW} by multiplex PCR assay. – *Journal of Virological Methods*. Nr. 165: 15 – 20. https://pubmed.ncbi.nlm.nih.gov/20025905/ (13.05.2025)

Chikh Ali, M., Maoka, T., Natsuaki, K. T. (2007). The occurrence and characterization of new recombinant isolates of *PVY* displaying shared properties of PVY^{nW} and PVY^{ntn}. – *Journal of Phytopathology*. Nr. 155(7–8), 409–415. https://onlinelibrary.wiley.com/doi/10.1111/j.1439-0434.2007.01251.x (25.05.2025)

Compendium of Potato Diseases. (1981). /Ed. W. J. Hooker. USA: The American Phytopathological Society. [on-line] ebrary (25.04.2025)

Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. – Nucleic Acids Research. Nr. 32(5), 1792–1797. https://doi.org/10.1093/nar/gkh340 (07.05.2025)

Elwan, E. A., Rabie, M., Abdel Aleem, E. E., Fattouh, F. A., Kagda, M. S., Zaghloul, H. A. H. (2023). Exploring virus presence in field-collected potato leaf samples using RNA sequencing. – *Journal of Genetic Engineering & Biotechnology*. Nr. 21, 106. https://pmc.ncbi.nlm.nih.gov/articles/PMC10589165/ (29.05.2025)

Fauquet, C. M., Mayo, M. A., Maniloff, J., Desselberger, U., Ball, L. A. (2005). Virus Taxonomy. London: Elsevier Academic Press. (25.04.2025)

Fawcett, G. L., Eterovic, A. K. (2017). Chapter Four - Identification of Genomic Somatic Variants in Cancer: From Discovery to Actionability. – *Advances in Clinical Chemistry*. Nr. 78, 123–162. https://doi.org/10.1016/bs.acc.2016.07.006 (22.05.2025)

Fitzpatrick, A. H., Rupnik, A., O'Shea, H., Crispie, F., Keaveney, S., Cotter, P. (2021) High Throughput Sequencing for the Detection and Characterization of RNA Viruses. – Front. Microbiol. Nr. 12:621719.

https://www.frontiersin.org/journals/microbiology/articles/10.3389/fmicb.2021.621719/full (24.05.2025)

Food and Agriculture Organization of the United Nations. (s.a). Crops and livestock products. [veebileht] https://www.fao.org/faostat/en/#data/QCL (24.04.2025)

Forté, A. J., Guertin, C., Cabana, J. (1999). Pathogenicity of a granulovirus towards *Choristoneura fumiferana* (Lepidoptera: Tortricidae). – *The Canadian Entomologist.* Nr. 131(6), 725–727. https://doi.org/10.4039/Ent131725-6 (20.05.2025)

Galvino-Costa, S. B. F., Figueira, A. R., Camargos, V. V., Geraldino, P. S., Hu, X.-J., Nikolaeva, O. V., Kerlan, C., Karasev, A. V. (2012a). A novel type of *Potato virus Y* recombinant genome, determined for the genetic strain PVY^{E}. – *Plant Pathology.* Nr. 61(2), 388–398. https://bsppjournals.onlinelibrary.wiley.com/doi/10.1111/j.1365-3059.2011.02495.x (25.05.2025)

Galvino-Costa, S. B. F., Figueira, A. R., Rabelo-Filho, F. A. C., Moraes, F. H. R., Nikolaeva, O. V., Karasev, A. V. (2012b). Molecular typing of *Potato virus Y* isolates from Brazil reveals a diverse set of recombinant strains. – *Plant Disease.* Nr. 96(10), 1451–1458. https://apsjournals.apsnet.org/doi/abs/10.1094/PDIS-02-12-0163-RE (25.05.2025)

Glasa, M., Hančinský, R., Šoltys, K., Predajňa, L., Tomašechová, J., Hauptvogel, P., Mrkvová, M., Mihálik, D., Candresse, T. (2021). Molecular characterization of potato Virus Y (PVY) using high-throughput sequencing: constraints on full genome reconstructions imposed by mixed infection involving recombinant PVY strains. – *Plants*. Nr. 10(4), 753. https://doi.org/10.3390/plants10040753 (27.05.2025)

Green, K. J., Brown, C. J., Gray, S. M., Karasev, A. V. (2017). Phylogenetic study of recombinant strains of *Potato virus Y*. – *Virology.* Nr. 507, 40–52. https://www.sciencedirect.com/science/article/pii/S0042682217301010?ref=pdf_download&fr=RR-2&rr=9453b5467c3a5c23 (25.05.2025)

Green, K. J., Brown, C. J., Karasev, A. V. (2018). Genetic diversity of potato virus Y (PVY): Sequence analyses reveal ten novel PVY recombinant structures. – *Archives of Virology*. Nr. 163(1), 23–32. https://doi.org/10.1007/s00705-017-3568-x (27.05.2025)

Hall, B. G. (2013). Phylogenetic Trees Made Easy: A How-To Manual by Barry G. Hall. – *The Quarterly Review of Biology*. Nr. 88(1). https://www.researchgate.net/publication/259712397_Phylogenetic_Trees_Made_Easy_A_How-To_Manual_by_Barry_G_Hall (07.05.2025)

Hawkes, J. G., Francisco-Ortega, J., (1993). The Early History of the Potato in Europe. – *Euphytica*. Nr. 70. https://link.springer.com/article/10.1007/BF00029633 (26.03.2025)

He, S., Cao, B., Yi, Y., Huang, S., Chen, X., Luo, S., Mou, X., Guo, T., Wang, Y., Wang, Y., Yang, G. (2021). DNA precipitation revisited: A quantitative analysis. – *Wiley Online Library*. Nr. 3. [e-ajakiri] https://onlinelibrary.wiley.com/doi/full/10.1002/nano.202100152 (06.05.2025)

Heather, J. M., Chain, B. (2016). The sequence of sequencers: The history of sequencing DNA. – *Genomics*. Nr. 107(1), 1–8. https://doi.org/10.1016/j.ygeno.2015.11.003 (06.05.2025)

Hema, M., Konakalla, N. C. (2021). Chapter 16 – Recent developments in detection and diagnoosis of plant viruses. – *Recent Developments in Applied Microbiology and Biochemistry*. Nr. 2. [e-ajakiri] https://www.sciencedirect.com/science/article/abs/pii/B9780128214060000163?via%3Dihub (06.05.2025)

Hon, T., Mars, K., Young, G., Tsai, Y.-C., Karalius, J. W., Landolin, J. M., Maurer, N., Kudrna, D., Hardigan, M. A., Steiner, C. C., Knapp, S. J., Ware, D., Shapiro, B., Peluso, P., Rank, D. R. (2020). Highly accurate long-read HiFi sequencing data for five complex genomes. – *Scientific Data.* Nr. 7(399). https://www.nature.com/articles/s41597-020-00743-4 (24.05.2025)

Hu, X., Karasev, A. V., Brown, C. J., Lorenzen, J. H. (2009). Sequence characteristics of Potato virus Y recombinants. – *Journal of General Virology*. Nr. 90(12), 3033–3041. https://biblio.iita.org/documents/S09ArtHuSequenceInthomNodev.pdf-c595af0b9f31a8da8f02f26bcdbbab80.pdf (13.05.2025)

Hughes, D. W., Galau, G. (1988). Preparation of RNA from cotton leaves and pollen. – *Plant Molecular Biology Reporter*. Nr. 6. [e-ajakiri] https://link.springer.com/article/10.1007/BF02670385 (06.05.2025)

Hull, R. (2004). Plant Virology. San Diego: Elsevier Academic Press. 11 lk. (25.04.2025)

Kang, H. J., Bennett, S. N., Dizney, L., Sumibcay, L., Arai, S., Ruedas, L. A., Song, J. W., Yanagihara, R. (2009). Host switch during evolution of a genetically distinct hantavirus in the American shrew mole (*Neurotrichus gibbsii*). – *Virology*, 388(1), 8–14. https://www.researchgate.net/publication/24360689_Host_switch_during_evolution_of_a_genetically_distinct_hantavirus_in_the_American_shrew_mole_Neurotrichus_gibbsii (20.05.2025)

Kaplinski, L., Lepamets, M., Remm, M. (2015). GenomeTester4: a toolkit for performing basic set operations - union, intersection and complement on k-mer lists. – *Gigascience.* Nr. 3. https://pmc.ncbi.nlm.nih.gov/articles/PMC4669650/ (07.05.2025)

Karasev, A.V., Gray, S.M. (2013). Continuous and emerging challenges of Potato virus Y in potato. – *Annual Review of Phytopathology*. Nr. 51, 571–586. https://pubmed.ncbi.nlm.nih.gov/23915135/ (13.05.2025)

Kartul. – EE kd. 4. Tallinn: Valgus, 1989, lk 353 (26.03.2025)

Kerlan C., Nikolaeva O. V., Hu X., Meacham T., Gray S. M. Karasev A. V. (2011). Identification of the molecular make-up of the Potato virus Y strain PVY^{Z}: Genetic typing of $PVY^{Z\text{-}NTN}$. – *Phytopathology*. Nr. 101: 1052 – 1060. https://www.researchgate.net/publication/51564827_Identification_of_the_molecular_make-up_of_the_Potato_virus_Y_strain_PVYZ_genetic_typing_of_PVYZ-NTN (13.05.2025)

Khanna, C., Paoloni, M. C. (2007). Molecular diagnostics. In S. J. Withrow, D. M. Vail, & R. L. Page (Eds.). – *Withrow & MacEwen's Small Animal Clinical Oncology* (4th ed., pp. 134–146). Saunders Elsevier. https://www.sciencedirect.com/science/article/abs/pii/B9780721605586500113 (26.05.2025)

Kreuze, J. F., Souza-Dias, J. A. C., Jeevalatha, A., Figueira, A. R., Valkonen, J. P. T., Jones, R. A. C. (2020). Viral diseases in potato. In H. Campos & O. Ortiz (Eds.). – *The potato crop* (lk. 389–430). Springer. https://link.springer.com/chapter/10.1007/978-3-030-28683-5_11 (26.05.2025)

Lacomme, C., Glais, L., Bellstedt, D., Dupuis, B., Karasev, A., Jacquot, E. (2017). Potato virus Y: Biodiversity, pathogenicity, epidemiology and management. – *Springer*. https://doi.org/10.1007/978-3-319-58860-5 (27.05.2025)

Levin, J. Z., Yassour, M., Adiconis, X., Nusbaum, C., Thompson, D. A., Friedman, N., Gnirke, A., Regev, A. (2010). Comprehensive comparative analysis of strand-specific RNA sequencing methods. – *Nature Methods*. Nr. 7(9), 709–715. https://pubmed.ncbi.nlm.nih.gov/20711195/ (06.05.2025)

Li, H., Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. – *Bioinformatics*. Nr. 25(14), 1754–1760. https://doi.org/10.1093/bioinformatics/btp324 (07.05.2025)

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 1000 Genome Project Data Processing Subgroup. (2009). The Sequence Alignment/Map format and SAMtools. – *Bioinformatics*. Nr. 25(16), 2078–2079. https://doi.org/10.1093/bioinformatics/btp352 (07.05.2025)

Li, J. B., Levanon, E. Y., Yoon, J. K., Aach, J., Xie, B., LeProust, E., Zhang, K., Gao, Y., Church, G. M. (2011). Genome-wide identification of human RNA editing sites by

parallel DNA capturing and sequencing. – *Science*. Nr. 324(5931), 1210–1213. https://doi.org/10.1126/science.1170995 (06.05.2025)

Li, R., Homer, N. (2010). A survey of sequence alignment algorithms for next-generation sequencing. – *Briefings in Bioinformatics*. Nr. 11(5), 473–483. https://doi.org/10.1093/bib/bbq015 (07.05.2025)

Li, R., Zhu, H., Ruan, J., Qian, W., Fang, X., Shi, Z., Li, Y., Li, S., Shan, G., Kristiansen, K., Li, S., Yang, H., Wang, H., Wang, J., Wang, J. (2010). De novo assembly of human genomes with massively parallel short read sequencing. – *Genome Research*. Nr. 20(2), 265–272. https://doi.org/10.1101/gr.097261.109 (07.05.2025)

Lorenzen, J. H., Meacham, T., Berger, P. H., Shiel, P. J., Crosslin, J. M., Hamm, P. B., Kopp, H. (2006). Whole genome characterization of *Potato virus Y* isolates collected in the western USA and their comparison to isolates from Europe and Canada. – *Archives of Virology*. Nr. 151(6), 1055–1074. https://link.springer.com/article/10.1007/s00705-005-0707-6 (25.05.2025)

Lorenzen, J., Nolte, P., Martin, D., Pasche, J. S., Gudmestad, N. C. (2008). NE-11 represents a new strain variant class of *Potato virus Y*. – *Archives of Virology*. Nr. 153(3), 517–525. https://link.springer.com/article/10.1007/s00705-007-0030-5 (25.05.2025)

MacKenzie, T. D. B., Nie, X., Singh, M. (2015). RT-PCR and real-time RT-PCR methods for the detection of potato virus Y in potato leaves and tubers. – *Methods in molecular biology*. Nr. 1236:13-26. https://www.researchgate.net/publication/266623462_RT-PCR_and_real-time_RT-PCR_methods_for_the_detection_of_potato_virus_Y_in_potato_leaves_and_tubers (24.05.2025)

Maclot, F., Candresse, T., Filloux, D., Malmstrom, C. M., Roumagnac, P., van der Vlugt, R., Massart, S. (2020). Illuminating an ecological blackbox: Using high throughput sequencing to characterize the plant virome across scales. – *Frontiers in Microbiology*. Nr. 11, 578064. https://doi.org/10.3389/fmicb.2020.578064 (06.05.2025)

Mardis, E. R. (2017). DNA sequencing technologies: 2006–2016. – Nature Protocols, 12(2), 365–368. https://doi.org/10.1038/nprot.2016.182 (06.05.2025)

Menzel, P., Ng, K. L., Krogh, A. (2016). Fast and sensitive taxonomic classification for metagenomics with Kaiju. – *Nature Communications*. Nr. 7(1), 11257. https://doi.org/10.1038/ncomms11257 (07.05.2025)

Mortazavi, A., Williams, B. A., McCue, K., Schaeffer, L., Wold, B. (2008). Mapping and quantifying mammalian transcriptomes by RNA-Seq. – *Nature Methods*. Nr. 5(7), 621-628. https://www.nature.com/articles/nmeth.1226 (06.05.2025)

Moury, B. (2010). A new lineage sheds light on the evolutionary history of Potato virus Y. – *Molecular Plant Pathology*. Nr 11. [e-ajakiri] https://bsppjournals.onlinelibrary.wiley.com/doi/full/10.1111/j.1364-3703.2009.00573.x (25.04.2025)

Nagarajan, N., Pop, M. (2013). Sequence assembly demystified. – *Nature Reviews Genetics*. Nr. 14(3), 157–167. https://doi.org/10.1038/nrg3367 (07.05.2025)

Naidu, R. A. and Hughes, J. A. (2001). Methods for the detection of plant virus diseases. In: Proceedings of a Conference Organized by IITA, Plant Virology in Sub Saharan Africa pp. 233-260. International Institute of Tropical Agriculture. Oyo State, Nigeria. https://biblio.iita.org/documents/S03ProcNaiduMethodsNothomNodev.pdf-2b2499326e69d3606f30eaf7db0844f4.pdf (06.05.2025)

Nguyen LT, Schmidt HA, von Haeseler A, Minh BQ. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. – *Mol Biol Evol*. 32(1):268–274. https://academic.oup.com/mbe/article/32/1/268/2925592 (07.05.2025)

Ogawa, T., Nakagawa, A., Hataya, T., Ohshima, K. (2012). The genetic structure of populations of Potato virus Y in Japan: Based on the analysis of 20 full genomic sequences.

– *Journal of Phytopathology*. Nr. 160(9), 661–673. https://onlinelibrary.wiley.com/doi/10.1111/j.1439-0434.2012.01959.x (25.05.2025)

Ogawa, T., Tomitaka, Y., Nakagawa, A., Ohshima, K. (2008). Genetic structure of a population of Potato virus Y inducing potato tuber necrotic ringspot disease in Japan: Comparison with North American and European populations. – *Virus Research.* Nr. 131(2), 199–212. https://www.sciencedirect.com/science/article/abs/pii/S0168170207003401 (25.05.2025)

Ohtlike taimekahjustajate ja nende peremeestaimede Eestisse toimetamise ja riigisiseselt edasitoimetamise ning sordiaretuses, katsetes ja muus teadustöös kasutamise nõuded ja eriloa väljastamise kord. (vastu võetud 30.03.2007, viimati jõustunud 01.01.2010). – Riigi Teataja. https://www.riigiteataja.ee/akt/12814297 (25.04.2025)

Onditi, J., Nyongesa, M., van der Vlugt, R. (2021). Prevalence, distribution and control of six major potato viruses in Kenya. – *Tropical Plant Pathology*. Nr. 46(3), 311–323. https://link.springer.com/article/10.1007/s40858-020-00409-x (25.04.2025)

Ortiz, O., Mares, V. (2017). The Historical, Social, and Economic Importance of the Potato Crop. – Springer International Publishing. https://link.springer.com/chapter/10.1007/978-3-319-66135-3_1 (26.03.2025)

Paez-Espino, D., Eloe-Fadrosh, E. A., Pavlopoulos, G. A., Thomas, A. D., Huntemann, M., Mikhailova, N., ... & Kyrpides, N. C. (2016). Uncovering Earth's virome. – *Nature*, 536(7617), 425-430. https://doi.org/10.1038/nature19094 (20.05.2025)

Pagan, I. (2022). Transmission through seeds: The unknown life of plant viruses. – *PLoS Pathogens*. Nr. 18. [e-ajakiri] https://journals.plos.org/plospathogens/article?id=10.1371/journal.ppat.1010707 (25.04.2025)

Palla´s, V., Sanchez, J.A., James, D. (2018). Recent advances on the mul-tiplex molecular detection of plant viruses and viroids. – *Front. Microbiol*. Nr. 9, 2087. https://doi.org/10.3389/fmicb.2018.02087 (06.05.2025)

Paten, B., Novak, A. M., Haussler, D. (2014). Mapping to a reference genome structure. – *arXiv*. Nr. 1404.5010). https://arxiv.org/abs/1404.5010 (25.05.2025)

Pearson, W. R. (2013). An introduction to sequence similarity ("Homology") searching. – *Current Protocols in Bioinformatics*. Nr. 42(1), 3.1.1–3.1.8. https://pmc.ncbi.nlm.nih.gov/articles/PMC3820096/ (07.05.2025)

Plant Viruses. (1979). /Ed. K. M. Smith. USA: Halsted Press. [on-line] ebrary (25.04.2025)

PM0281: Põllumajandusmaa ja -kultuurid maakonna järgi. – Eesti Statistika andmebaas. https://andmed.stat.ee/et/stat/majandus__pellumajandus__pellumajandussaadustetootmine__taimekasvatussaaduste-tootmine/PM0281/table/tableViewLayout2 (25.04.2025)

Pop, M. (2009). Genome assembly reborn: recent computational challenges. – *Briefings in Bioinformatics*. Nr. 10(4), 354–366. https://doi.org/10.1093/bib/bbp026 (07.05.2025)

Potato virus Y: biodiversity, pathogenicity, epidemiology and management. (2017). /Ed. C, Lacomme, L. Glais, D. U. Bellstedt, B. Dupuis, A. V. Karasev, E. Jacquot. Cham: Springer International Publishing AG. [on-line] ebrary (25.04.2025)

Potato Viruses and Their Control. (1996). /Ed. L. F. Salazar. Peru: Peru International Potato Center. [on-line] ebrary (24.05.2025)

Potato: a global history. (2011). /Ed. A. F. Smith. London: Reaktion Books Ltd. [on-line] ebrary (26.03.2025)

Potato: Nutrition and Food Security. (2020). /Ed. P. Raigond, B. Singh, S. Dutt, S. K. Chakrabarti. Singapore: Springer Nature Singapore Pte Ltd. [on-line] ebrary (24.04.2024)

Quenouille J., Vassilakos N. Moury B. (2013). Potato virus Y: a major crop pathogen that has provided major insights into the evolution of viral pathogenicity. – *Molecular Plant Pathology*. Nr. 14: 439 – 452. https://pubmed.ncbi.nlm.nih.gov/23480826/ (13.05.2025)

Rai, M. F., Tycksen, E. D., Sandell, L. J., Brophy, R. H. (2019). Advantages of RNA-seq compared to RNA microarrays for transcriptome profiling of anterior cruciate ligament

tears. – *Journal of Orthopaedic Research*. Nr. 36(1), 484–497. https://pmc.ncbi.nlm.nih.gov/articles/PMC5787041/ (06.05.2025)

Reuter, J. A., Spacek, D. V., Snyder, M. P. (2015). High-throughput sequencing technologies. – *Molecular Cell.* Nr. 58(4), 586–597. https://www.cell.com/molecular-cell/fulltext/S1097-2765(15)00340-8 (27.05.2025)

Roux, S., Hallam, S. J., Woyke, T., & Sullivan, M. B. (2015). Viral dark matter and virus–host interactions resolved from publicly available microbial genomes. – *eLife*, 4, e08490. https://doi.org/10.7554/eLife.08490 (20.05.2025)

Rubio, L., Galipienso, L., Ferriol, I. (2020). Detection of Plant Viruses and Disease Management: Relevence of Genetic Diversity and Evolution. – *Front Plant Sci*. Nr. 11. [e-ajakiri] https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7380168/ (06.05.2025)

Samarskaya, V., Kuznetsova, M., Gryzunov, N., Spechenkova, N., Bagdasarova, P., Ryabov, E., Taliansky, M., Kalinina, N. O. (2024). Identification of two novel recombinant types of potato virus Y from *Solanum tuberosum* plants in southern region of Russia. – *Plant Disease*. https://doi.org/10.1094/PDIS-10-24-2151-SC (27.05.2025)

Sanger, F., Nicklen, S., Coulson, A. R. (1977). DNA sequencing with chain-terminating inhibitors. – *Proceedings of the National Academy of Sciences*. Nr. 74(12), 5463–5467. https://doi.org/10.1073/pnas.74.12.5463 (06.05.2025)

Sastry, K. S. (2013). Plant Virus Transmission Through Vegetative Popagules (Asexual Reproduction. – *Seed-borne plant virus diseases*. [e-ajakiri] https://link.springer.com/chapter/10.1007/978-81-322-0813-6_9 (25.04.2025)

Scholthof, H. B. (2005). Plant virus transport: motions of functional equivalence. – *Trends in Plant Science*. Nr. 10. [e-ajakiri] https://www.sciencedirect.com/science/article/abs/pii/S1360138505001512 (25.04.2025)

Scholthof, K-B. G., Adkins, S., Czosnek, H., Palukaitis, P., Jaquot, E., Hohn, T., Hohn, B., Saunders, K., Candresse, T., Ahlquist, P., Hemenway, C., Foster, G. D. (2011). Top 10 plant viruses in molecular plant pathology. – *Mol Plant Pathol*. Nr. 12. [e-ajakiri] https://pubmed.ncbi.nlm.nih.gov/22017770/ (25.04.2025)

Schubert, J., Fomitcheva, V., Sztangret-Wiśniewska, J. (2007). Differentiation of *Potato virus Y* strains using improved sets of diagnostic PCR-primers. – *Journal of Virological Methods.* Nr. 140(1–2), 66–74. https://www.sciencedirect.com/science/article/abs/pii/S0166093406003843 (25.05.2025)

Seemnekartuli kategooriad ning seemnekartuli tootmise ja turustamise nõuded. (vastu võetud 03.05.2006, viimati jõustunud 01.01.2025). – Riigi Teataja. https://www.riigiteataja.ee/akt/110082012009?leiaKehtiv (25.04.2025)

Shendure, J., Balasubramanian, S., Church, G. M., Gilbert, W., Rogers, J., Schloss, J. A., Waterston, R. H. (2017). DNA sequencing at 40: Past, present and future. – *Nature*. Nr. 550(7676), 345–353. https://doi.org/10.1038/nature24286 (06.05.2025)

Simpson, J. T., Wong, K., Jackman, S. D., Schein, J. E., Jones, S. J. M., Birol, I. (2009). ABySS: a parallel assembler for short read sequence data. – *Genome Research*. Nr. 19(6), 1117–1123. https://doi.org/10.1101/gr.089532.108 (07.05.2025)

Singh, M., Singh, R. P., Fageria, M. S., Nie, X., Coffin, R., Hawkins, G. (2013). Optimization of a Real-Time RT-PCR Assay and its Comparsion with ELISA, Conventional RT-PCR and the Geow-out Test for Large Scale Diagnosis of Potato virus Y in Dormant Potato Tubers. – *American Journal of Potato Research*. Nr. 90. [e-ajakiri] https://link.springer.com/article/10.1007/s12230-012-9274-z (06.05.2025)

Singh, R. P., Valkonen, J. P., Gray, S. M., Boonham, N., Jones, R. A., Kerlan, C., Schubert, J. (2008). Discussion paper: The naming of potato virus Y strains infecting potato. – *Archives of Virology*. Nr. 153(1), 1–13. https://doi.org/10.1007/s00705-007-1059-1 (27.05.2025)

Singh, R.P., Valkonen, J.P.T., Gray, S.M., Boonham, N., Jones, R.A.C., Kerlan, C., Schubert, J. (2008). Discussion paper: The naming of Potato virus Y strains infecting

potato. – *Archives of Virology*. Nr. 153(1), 1–13. https://link.springer.com/article/10.1007/s00705-007-1059-1 (13.05.2025)

Singh, S., Awasthi, L. P., Jangre, A. (2020). Transmission of plant viruses in fields through various vectors. – *Applied Plant Virology*. [e-ajakiri] https://www.sciencedirect.com/science/article/pii/B9780128186541000244 (25.04.2025)

Stamatakis, A. (2014). RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. – *Bioinformatics.* Nr. 30(9), 1312–1313. https://pmc.ncbi.nlm.nih.gov/articles/PMC3998144/ (26.05.2025)

Takacs, A., Horvath, J., Gaborjanyi, R., Kazinczi, G., Mikulas, J. (2014). Hosts and non-hosts in plant virology and the effects of plant viruses on host plants. – *Plant Virus-Host Interaction*. [e-ajakiri] https://www.sciencedirect.com/science/article/pii/B9780124115842000056 (25.04.2025)

Tao, T. (2010). Standalone BLAST Setup for Unix. In *BLAST® Help* [Internet]. https://www.ncbi.nlm.nih.gov/books/NBK52640/ (07.05.2025)

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., Pimentel, H., Salzberg, S. L., Rinn, J. L., Pachter, L. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and Cufflinks. – *Nature Protocols*. Nr. 7, 562–578. https://doi.org/10.1038/nprot.2012.016 (06.05.2025)

Tsedaley, B. (2015). A Review Paper on Potato Virus Y (PVY) Biology, Economic Importance and its Managements. – *Journal of Biology, Agriculture and Healthcare*. Nr 9. [e-ajakiri] https://core.ac.uk/download/pdf/234660999.pdf (24.04.2025)

van Dijk, E. L., Jaszczyszyn, Y., Naquin, D., Thermes, C. (2018). The third revolution in sequencing technology. – *Trends in Genetics*. Nr. 34(9), 666–681. https://doi.org/10.1016/j.tig.2018.05.008 (06.05.2025)

Virus and Virus-like Diseases of Potatoes and Production of Seed-Potatoes. (2001). /Ed. G. Loebenstein, P. H. Berger, A. A. Brunt, R. H. Lawson. Dordrecht: Kluwer Academic Publishers. [on-line] ebrary (25.04.2025)

Virus and Virus-like Diseases of Potatoes and Production of Seed-Potatoes. (2001). /Ed. G. Loebenstein, P. H. Berger, A. A. Brunt, R. H. Lawson. Dordrecht: Kluwer Academic Publishers. [on-line] ebrary (24.05.2025)

Wang, Z., Gerstein, M., Snyder, M. (2009). RNA-Seq: A revolutionary tool for transcriptomics. – *Nature Reviews Genetics*. Nr. 10(1), 57-63. https://www.nature.com/articles/nrg2484 (06.05.2025)

Webster, C.G., Wylie, S.J., Jones, M.G.K. (2004). Diagnosis of plant viralpathogens. – *Curr. Sci*. Nr. 86 (12), 1604–1607. https://www.jstor.org/stable/24108010 (06.05.2025)

Wenger, A. M., Peluso, P., Rowell, W. J., Chang, P.-C., Hall, R. J., Concepcion, G. T., Ebler, J., Fungtammasan, A., Kolesnikov, A., Olson, N. D., Töpfer, A., Alonge, M., Mahmoud, M., Qian, Y., Chin, C.-S., Phillippy, A. M., Schatz, M. C., Myers, G., DePristo, M. A., ... Hunkapiller, M. W. (2019). Accurate circular consensus long-read sequencing improves variant detection and assembly of a human genome. – *Nature Biotechnology*. Nr. 37(10), 1155–1162. https://www.nature.com/articles/s41587-019-0217-9 (24.05.2025)

Wood, D. E., Lu, J., Langmead, B. (2019). Improved metagenomic analysis with Kraken 2. – Genome Biology. Nr. 20(1), 257. https://doi.org/10.1186/s13059-019-1891-0 (07.05.2025)

Youssef, A. (2017). Presence of recombinant Potato virus Y genotypes in Sweden. – *Master's thesis, Swedish University of Agricultural Sciences*. https://www.researchgate.net/publication/318876937_Presence_of_Recombinant_Potato_Virus_Y_Genotypes_in_Sweden (13.05.2025)

Zerbino, D. R., Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. – *Genome Research*. Nr. 18(5), 821–829. https://pmc.ncbi.nlm.nih.gov/articles/PMC2336801/ (07.05.2025)

Zhao, W., He, X., Hoadley, K. A., Parker, J. S., Hayes, D. N., Perou, C. M. (2014). Comparison of RNA-Seq by poly (A) capture, ribosomal RNA depletion, and DNA microarray for expression profiling. – *BMC Genomics*. Nr. 15(1), 419. https://doi.org/10.1186/1471-2164-15-419 (06.05.2025)

Zheng, K., Chen, T. C., Yeh, S. D., Rahman, M. S., Su, X., Wu, K., Li, T., Zhang, Z., Dong, J. (2017). Characterization of a new isolate of pepper chlorotic spot virus from Yunnan province, China. – *Archives of Virology*, 162(9), 2809–2814. https://www.researchgate.net/publication/317038658_Characterization_of_a_new_isolate_of_pepper_chlorotic_spot_virus_from_Yunnan_province_China (20.05.2025)

LISAD

Lisa 1. PVY tuvastamine ELISA meetodil

1. Kartuli taimi võib koguda analüüsimiseks, kui neil on 3-4 väljaarenenud lehte (mitte varem, kui 3 nädalat peale kasvama panekut).
2. Eemalda kastist kasvav taim koos silmatükiga.
3. Eemalda taime maapealne osa silmatükist, murdes see taime ladvast.
4. Kui mugulast on tõusnud mitu taime, võta proovi jaoks tükk ainult ühest taimest.
5. Kogu kõik ühe proovi tärganud taimede ladvad plastikust kasti, millel on kleeppaber, kuhu peale on kirjutatud proovi number ja kastide numbrid, millest taimed korjatakse.
6. Valmista siduvate antikehade lahus
7. Pipeteeri igasse plaadi auku 200 µl katvate antikehade lahust
8. Sulge plaat kilega (sealing tape)
9. Aseta plaat termostaati ja inkubeeri termostaadis/inkubaatoris +30±2˚C juures 4 tundi või jäta külmikusse ööpäevaks
10. Kui soovid kasutada kaetud plaate edasiseks analüüsimiseks, pese plaadid käsitsi. Selleks pese plaate 2 korda jooksva külma kraaniveega, lisa aukudesse PBST 1x pesupuhver, leota minimaalselt 30 sekundit pesupuhvriga ja pese veel 2 korda jooksva kraanivee all. Kuivata plaadid
11. Vastavalt kontsentratsioonile, kanna plaadile ekstraktsiooni puhver ja mahlalahjendus
12. Kanna ELISA plaatidele positiivsed ja negatiivsed kontrollid ja tavalised proovid
13. Sulge plaat kilega (sealing tape)
14. Jäta mikroplaat kantud proovidega üheks ööks külmkappi (+5 ±3 °C) inkubeerima
15. Järgmisel hommikul pese plaadid käsitsi.. Selleks pese plaate 2 korda jooksva külma kraaniveega, lisa aukudesse PBST 1x pesupuhver, leota minimaalselt 30 sekundit pesupuhvriga ja pese veel 2 korda jooksva kraanivee all. Kuivata plaadid
16. Valmista konjugeeritud antikeha lahus
17. Pipeteeri igasse plaadi auku 200 µl konjugeeritud antikehade lahust
18. Sulge plaat kilega (sealing tape)
19. Aseta plaat termostaati ja inkubeeri termostaadis/inkubaatoris +30 ± 2 ˚C juures 5 tundi
20. Pese plaate 2 korda jooksva külma kraaniveega, lisa aukudesse PBST 1x pesupuhver, leota minimaalselt 30 sekundit pesupuhvriga ja pese veel 2 korda jooksva kraanivee all. Kuivata plaadid
21. Valmista substraat puhvrit vahetult enne kasutamist
22. Pipeteeri igasse plaadi auku 200 µl substraadi lahust
23. Inkubeeri toatemperatuuril pimedas kohas 1-2 tundi

24. Mõõtmist teosta spektrofotomeetriga (Tecan Infinite F200) ja arvutiprogrammiga Magellan

Lisa 2. RNA ekstraktsioon Hughes ja Galau meetodil

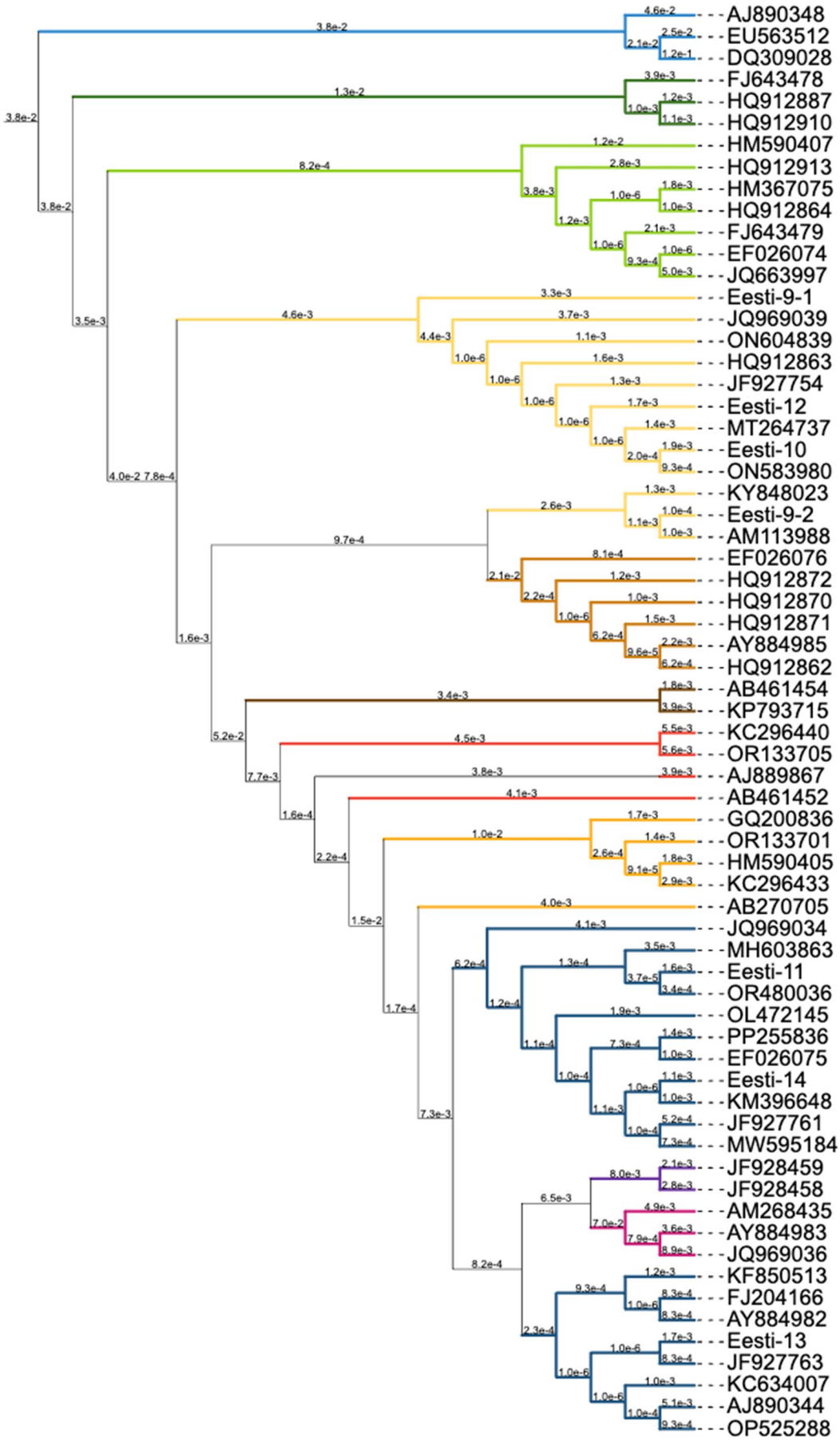
1. Kaalu 0,5 – 1 g proovi kilekotti
2. Hõõru läbi kile proov (näiteks leht) puruks, kasutades matseraatorit – rullikut või tambi haamriga mugulate puhul. Lisa 2,5 – 5 ml H&G ekstraktsiooni tööpuhvrit ja hõõru matseraatoriga
3. Vala saadud lahus nummerdatud 15 ml tuubi
4. Lisa sama kogus 6M KOAc, sega tuubi üles-alla keerates
5. Aseta jääle 15 minutiks
6. Tsentrifuugi max kiirusel 15 minutit
7. Mõõda 700 µl saadud lahust 1,5 ml tuubi
8. Lisa 700 µl 4M LiCl, sega tuubi võrekesil
9. Inkubeeri +4 ºC juures üleöö
10. Tsentrifuugi max kiirusel (13400 rpm) 25 minutit +4 ºC juures
11. Ettevaatlikult eemalda vedelikuosa, vigastamata pelletit
12. Resuspendeeri pellet 200 µl TE-puhvris, mis sisaldab 1% SDS. Lisa 100 µl 5M NaCl ja 300 µl jääkülma isopropanooli.
13. Sega tuubi üles-alla keerates ja inkubeeri -20 ºC juures 20 minutit
14. Tsentrifuugi +4 ºC juures 10 minutit max-kiirusel, et nukleiinhape põhja sadestada
15. Eemalda pealt soola/isopropanooli lahus (viska jääkide anumasse)
16. Pese pelletit lisades 400 – 500 µl 70% etanooli, nipsuta tuubi
17. Tsentrifuugi 5 minutit +4 ºC juures max-kiirusel
18. Eemalda pealt etanool, vigastamata pelletit
19. Kuivata pellet, tuubi õhu õhu juures avatuna hoides
20. Resuspendeeri pellet 50 µl ultra puhtas steriilses vees, sega vortexil
21. Jäta seisma toatemperatuuril 5 minutiks
22. Saadud lahus sisaldab RNA/DNA ja seda võib kasutada edasi PCR analüüsi läbiviimiseks
23. Lühiajaliseks säilitamiseks hoia +4 ºC juures, kauemaks säilitamiseks -20 ºC.

Lisa 3. GenomeTester4 *gmer_counter* andmebaasi faili struktuur

Failis on neli tulpa, millest esimene on infoväli, teine k-meride arv reas ning järgmised k-mer järjestused, mille sagedusi gmer_counter arvutab.

32	2	AGAAAAACAACGCAAAAACACTCA**T**	AGAAAAACAACGCAAAAACACTCA**C**
33	2	GAAAAACAACGCAAAAACACTCA**T**A	GAAAAACAACGCAAAAACACTCA**C**A
34	2	AAAAACAACGCAAAAACACTCA**T**AA	AAAAACAACGCAAAAACACTCA**C**AA
35	2	AAAACAACGCAAAAACACTCA**T**AAA	AAAACAACGCAAAAACACTCA**C**AAA
36	2	AAACAACGCAAAAACACTCA**T**AAA**C**	AAACAACGCAAAAACACTCA**C**AAAA
37	2	AACAACGCAAAAACACTCA**T**AAA**C**G	AACAACGCAAAAACACTCA**C**AAAA**G**
38	2	ACAACGCAAAAACACTCA**T**AAA**C**GC	ACAACGCAAAAACACTCA**C**AAAA**G**C
39	2	CAACGCAAAAACACTCA**T**AAA**C**GCT	CAACGCAAAAACACTCA**C**AAAA**G**CT
40	2	AACGCAAAAACACTCA**T**AAA**C**GCTT	AACGCAAAAACACTCA**C**AAAA**G**CTT

Lisa 4. Fülogeneetiline puu koos harupikkustega



Lisa 5. Lihtlitsents lõputöö reprodutseerimiseks ja üldsusele kättesaadavaks tegemiseks

Mina, Anette Stražev,

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) minu loodud teose:
   „Kõrge läbilaskevõimega sekveneerimise rakendamine viiruste tuvastamiseks Eesti seemnekartulis ja kartuli Y-viiruse tüvede identifitseerimine bioinformaatiliste meetoditega“,
   mille juhendajad on Piret van der Sman ja Tarmo Puurand,
   reprodutseerimiseks eesmärgiga seda säilitada, sealhulgas lisada Tartu Ülikooli digitaalarhiivi kuni autoriõiguse kehtivuse lõppemiseni;

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

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

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

Anette Stražev

29.05.2025