

Tartu Ülikool
Loodus- ja tehnoloogiateaduskond
Ökoloogia ja Maateaduste Instituut
Taimeökoloogia õppetool

Kätlin Lindsaar

KROHMSEENTE DNA-PÕHINE MÄÄRAMINE

Bakalaureusetöö

Juhendaja: vanemteadur Maarja Öpik

Tartu 2015

Sisukord

1. Sissejuhatus.....	3
2. DNA järjestuste kasutamine krohmseente määramisel.....	5
3. Krohmseente DNA-põhiseks määramiseks kasutatavad molekulaarsed markerid.....	8
3.1 ITS piiirkond rRNA operonis.....	9
3.2 Tuuma rRNA väikest subühikut (SSU rRNA) kodeeriv geen	9
3.3 Tuuma rRNA suurt subühikut (LSU rRNA) kodeeriv geen	10
4. Krohmseente järjestusepõhine rühmitamine.....	12
4.1 Virtuaaltakson	12
4.2 Liigihüpotees.....	12
4.3 MOTU (ingl <i>molecular operational taxonomic unit</i>)	13
5. Andmebaasid.....	15
5.1 MaarjAM andmebaas	15
5.2 UNITE andmebaas	16
5.3 PHYMYCO-DB	16
5.4 SILVA andmebaas	16
5.5 Nomenklatuuri andmebaasid.....	16
6. Molekulaarsete meetodite kasutamine krohmseente taksonoomias.....	18
Kokkuvõte	19
Summary	20
Tänuavaldused.....	21
Kasutatud kirjandus.....	22

1. Sissejuhatus

Läbi ajaloo on inimesed proovinud leida seaduspärasid maailmast, mis neid ümbritseb. Selline teadmisanu on olnud alguseks mitmetele tänapäeval tuntud teadusharudele nagu näiteks süstemaatika. Süstematiseerimise üheks alustajaks võiks pidada Aristotelest, kes oma toonaste teadmistega jagas loomariigi verega ja vereta loomadeks. Peale Aristotelest on süstemaatika vastu huvi tundnud ka teised teadlased nagu näiteks botaanik Clusius, kes jaotas 1698. aastal seeneriigi kunstliku süsteemi alusel söödavateks ja mürgisteks seenteks (Parmasto 1996). Sellises süsteemis ei arvestatud organismi omaste tunnustega, vaid lähtuti nende mõjust inimesele.

Määramise ehk identifitseerimise abil paigutatakse isend taksonoomilisse süsteemi. Paljale silmale nähtamatute organismide määramine ja kirjeldamine muutus võimalikuks tänu mikroskoobi leiutamisele. Tänapäevaks on muutunud kättesaadavaks molekulaarbioloogilised meetodid, mis pakuvad suuremaid võimalusi mikroskoopiliste organismide määramiseks.

Arbuskulaarne mükoriisa on sümbioos hõimkonda krohmseened (*Glomeromycota*) kuuluvate seente ja taimejuurte vahel (Parniske 2008). Krohmseened on mikroorganismid. Arbuskulaarset mükoriisat moodustavate seente üheks iseloomulikuks ehituslikuks komponendiks on juuresisene põõsakujuliste struktuuride ehk arbuskulite moodustumine (Smith & Read 2008). Nende kaudu toimub kahesuunaline ainete transport, kus toitained (põhiliselt fosfor ja lämmastik) transporditakse taimerakku ja süsinikuühendid taimerakkust seenehüüfidesse, mis on seenele ainsaks energiaallikaks (Parniske 2008). Lisaks arbuskulaarset mükoriisat moodustavatele seentele kuulub krohmseente hõimkonda ka tsüanobakteriga *Nostoc punctiforme* sümbioosi moodustav seen *Geosiphon pyriformis* (Smith & Read 2008).

Kuni viimase kümnendini määrati krohmseeni peamiselt morfoloogiliste tunnuste abil (Robinson-Boyer jt 2009). Määramine põhines enamasti krohmseente liikide eostel, mis isoleeriti otse keskkonnaproovidest või püüniskultuuridest ning nende vaatlemiseks kasutati mikroskoopi (Robinson-Boyer jt 2009). Antud meetod on aeganõudev ning selle kasutus on piiratud liikidele, mis ei moodusta eoseid. Lisaks vajab morfoloogiapõhine määramine spetsiifilisi ekspertoskusi (Öpik jt 2014).

Morfoloogilisele krohmseente määramismeetodile on lisandunud DNA-põhised meetodid. Molekulaarseid meetodeid kasutades on võimalik krohmseeni määrata paljudest erinevatest proovidest (Hart jt 2015). Seda meetodit on võimalik kasutada ka krohmseente liikide puhul, mis ei moodusta eoseid. Võrreldes morfoloogilise määramismeetodiga, on molekulaarsed meetodid objektiivsemad ja standardiseeritumad. Lisaks on molekulaarset määramist võimaldav tehnoloogia muutunud laialt kättesaadavaks ning võrdlemisi odavaks (Öpik jt 2014).

Käesoleva töö eesmärgiks on anda ülevaade, kuidas on võimalik DNA-põhiselt silmaga nähtamatuid krohmseeni määrata. Põhjalikumalt tutvustan krohmseente määramisel kasutatavaid markerpiirkondi, klassifitseeritud DNA järjestuste rühmitusi ja krohmseente liikide määramiseks kasutatavaid andmebaase. Lõpetuseks annan lühikese ülevaate molekulaarsete meetodite kasutamisest krohmseente taksonoomias.

2. DNA järjestuste kasutamine krohmseente määramisel

Krohmseente DNA-põhiseks määramiseks eraldatakse proovist (nt mullast) kõikide organismide DNA (Öpik 2012). Eraldatud DNA markerpiirkondadest paljundatakse PCRi abil krohmseente DNA, mis sekveneeritakse (Öpik 2012). Saadud DNA järjestusi võrreldakse andmebaasides olevatega. Andmebaasist saab vaadata, kas seal leidub sarnase DNA järjestusega varem sekveneeritud liike (Hart jt 2015). Järjestused, mida andmebaasides ei leidu, klassifitseeritakse iseseisvateks rühmadeks (Öpik 2012). Joonisel 1 on toodud välja peamised etapid, mida kasutatakse krohmseente DNA-põhiseks määramiseks. Antud peatükis annan täpsema ülevaate krohmseente DNA-põhisel määramisel kasutatavatest etappidest ja selle meetodi põhjal saadud liikide nimetamisest.

Krohmseente molekulaarseks analüüsiks vajalikku DNAd on võimalik eraldada taime juurtest, eostest, hüüfidest või mullaproovidest (Hart jt 2015). Proovi säilitamiseks kasutatakse erinevaid meetodeid. Kõige lihtsam ja odavam juurtest pärineva proovi säilitusmeetod on kuivatuskapis kuivatamine madalatel temperatuuridel (50–60°C) (Janouškova jt 2015). Kõige kiirem ja mugavam meetod, mida laboris kasutatakse, on proovi külmutamine lämmastikus (Hart jt 2015). Lisaks kasutatakse ka proovi hoiustamist DNA eralduspuhvris ja külmu kuivatusmeetodit (Hart jt 2015). Üldiselt tuleks hoida proovi säilitamise aega selle võtust võimalikult väiksena. See on oluline, et vähendada DNA degradatsiooni ja seente koosluste muutusi (Hart jt 2015).

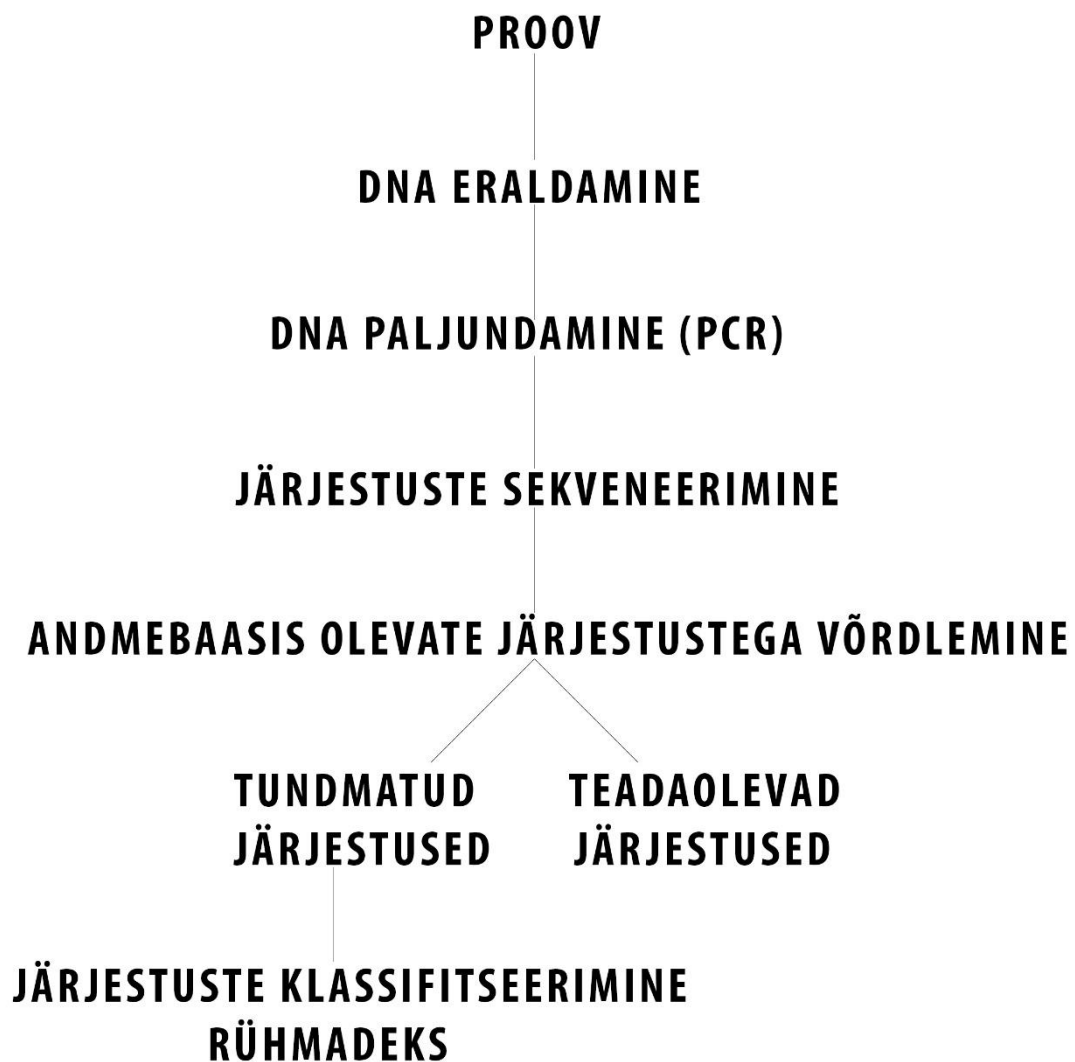
Kõige enam kasutatakse krohmseente molekulaarsel määramisel taimejuurtest pärinevaid proove (Hart jt 2015). Juureproovist eraldatakse kõigi seal olevate organismide DNA (Öpik 2012). Eraldatud DNA markerpiirkondadest on võimalik paljundada PCRi abil ainult krohmseente DNA (Öpik 2012). Selleks, et PCRi käigus oleks võimalik paljundada krohmseente DNA fragmente, on vaja kasutada spetsiifilisi praimereid. Praimer on lühike nukleotiidjärjestus, millelt toimub matriitsahela alusel DNA süntees (Heinaru 2012). PCRi analüüsimiseks on võimalik kasutada geelelektroforeesi (Heinaru 2012).

Saadud sekventse võrreldakse andmebaasides olevate referentsjärjestustega, et teada saada, kas seal leidub sarnase DNA järjestusega varem sekveneeritud liike (Hart jt 2015). Referentsjärjestused saadakse kultuurist eraldatud krohmseente liikide eostest või hüüfidest. Juhul, kui andmebaasides teada olevatele liikidele ei leidu

samasugust DNA järjestust, siis pole võimalik liike määrata. Sellised sekventsids tuleb klassifitseerida eraldi rühmadeks, nagu näiteks virtuaaltaksoniteks (Öpik 2012).

Paljudel DNA-põhistel krohmseente liikidel puuduvad nimed. Seentele nimede andmiseks tuleb mükoloogidel jälgida rahvusvahelist vetikate, seente ja taimede nomenklatuuri koodeksit (ingl *International Code of Nomenclature for Algae Fungi and Plants*) (McNeill jt 2012). Koodeks ei luba ametlikult kirjeldada liiki, mis põhineb ainult sekventsidel (Herr jt 2015). Nimetuse andmiseks on vaja eksemplari füüsilisel kujul (Hibbett & Taylor 2013). Seetõttu ainult keskkonnajärjestustel põhinevaid liike ei ole võimalik kehtivalt nimetada, sest enamikel juhtudel puudub järjestusel võrdluseks tüüpeksemplar (Herr jt 2015). Teatud olukordades võidakse joonist aktsepteerida kui tüüpeksemplari (Herr jt 2015). Koodeksit võiks modifitseerida, et oleks lubatud sekventsipõhiste liikide kirjeldamine. See parandaks kommunikatsiooni ja tõstaks teadlikust seente mitmekesisusest ja nende ökoloogilistest rollidest (Herr jt 2015).

Hibbett jt pakkusid välja, et DNA-põhiste liikidele tuleks anda ladinakeelne perekonnanimi ning liigiepiteet (Hibbett jt 2011). Soovitavalt võiks selliselt piiritletud liikidele nimetus anda siis, kui sama järjestusega uue liigi on leidnud ka teine uurimisgrupp (Hibbett & Taylor 2013). Selleks, et eristada DNA-põhiseid liike morfoloogia põhjal määratud liikidest, võiks antud liiginimedele lisada järelliiteid, nagu näiteks ENAS (ingl *environmental nucleic and sequence*) või eMOTU (ingl *environmental molecular operational taxonomic unit*) (Hibbett & Taylor 2013).



Joonis 1. Peamised etapid, mida kasutatakse krohmseente DNA-põhiseks määramiseks.

3. Krohmseente DNA-põhiseks määramiseks kasutatavad molekulaarsed markerid

Molekulaarseid markereid kasutatakse DNA-põhisel määramisel krohmseente taksonite üksteisest eristamiseks. Kõige sagedamini kasutatakse tuuma ribosomaalse RNA operonis olevaid markerpiirkondi, nagu tuuma ribosomaalse RNA väikest subühikut (SSU rRNA) kodeerivat geeni (Helgason jt 1999), ITS (ingl *internal transcribed spacer*) regiooni (Wubet jt 2004) ja tuuma ribosomaalse RNA suurt subühikut (LSU rRNA) kodeerivat geeni (Gollotte jt 2004). Joonisel 2 on kujutatud nende markerpiirkondade paiknemist rRNA operonis. rRNA operoni on võimalik lihtsasti paljundada, sest seda leidub tuumas paljudes koopiates (Hart jt 2015). Krohmseente ökoloogilistes uurimustes kasutatakse sagedamini tuuma rRNA SSU kodeerivat geeni ning taksonoomilistes uurimustes rohkem tuuma rRNA LSU kodeerivat geeni ja ITS piirkonda (Öpik jt 2014). Tabelis 1 on omavahel võrreldud kõige sagedamini kasutatavaid rRNA operonis olevaid markerpiirkondi.

Lisaks tuuma rRNA operonile on kasutusel ka teisi markereid. Neid kasutatakse, sest rRNA operoni erinevad osad ei ole iseseisvalt arenevad (Öpik jt 2014).

Mitokondriaalsetest markeritest kasutatakse krohmseente DNA-põhiseks määramiseks mtLSUd, mille abil on võimalik kirjeldada liigisisest variatsiooni (Thiéry jt 2010). Valke kodeerivatest geenidest on võimalik kasutada näiteks RNA polümeraas II suurt subühikut kodeerivat geeni (Stockinger jt 2014) ja beeta-tubuliini kodeerivat geeni (Msiska & Morton 2009). Valke kodeerivaid genee leidub genoomis tavaliselt ühe koopiana (Lindahl jt 2013).

Krohmseente DNA-põhiseks määramiseks kasutatakse erinevaid markereid (Stockinger jt 2010). Praeguseks pole veel leitud ühte kõige paremat markerpiirkonda, mida oleks võimalik kasutada krohmseente määramiseks keskkonnaproovidest (Öpik jt 2014). Ideaalne marker peaks olema piisavalt varieeruv ning sisaldama püstitatud hüpoteesi testimiseks sobivaid andmeid (Hart jt 2015). Markerite valimisel tuleb arvestada järjestuste pikkustega. Üldiselt arvatakse, et pikemad järjestused on paremad liikide eristamiseks kui lühemad (Hart jt 2015). Pikemad fragmendid sisaldavad fülogeneetiliste analüüside jaoks rohkem informatsiooni (Lindahl jt 2013). Markerpiirkondade pikkused on varieeruvad, aga tavaliselt jäävad need 500-800 bp vahele (Schoch jt 2012).

Markerpiirkond, mis sobib liigi määramiseks, ei pruugi olla kasutatav liigist kõrgema taksoni (nt seltsi) kindlakstegemiseks (Öpik jt 2014). Markeri valik sõltub järjestuste varieeruvusest. Näiteks markerpiirkond, mille varieeruvus on sobilik seltside eristamiseks, ei pruugi olla piisav liikide määramiseks. Krohmseente liikide eristamiseks peab markeri varieeruvus olema liikidevaheliselt suurem kui liikidesiseselt (Öpik jt 2014). Antud peatükis annan ülevaate põhilistest markeritest, mida kasutatakse krohmseente molekulaarsetes uurimustes.



Joonis 2. Tuuma rRNA operoni osa. Joonisel kujutatud piirkonnad pole proportsioonis.

3.1 ITS piirkond rRNA operonis

Tuuma ribosomaalne ITS regioon on ametlik seente ribakood (Schoch jt 2012). ITS piirkonda kasutatakse sekventsipõhiseks liigi määramiseks paljudes seente hõimkondades (Hibbett & Taylor 2013). ITS regiooni kuuluvad ITS1 ja ITS2 speisserid ning nende vahele jääv 5.8S geen (Schoch jt 2012). Selle piirkonna abil ei ole võimalik määrata lähedalt suguluses olevaid krohmseente liike (Stockinger jt 2009). ITS regiooni eeliseks on suur varieeruvus ning seda on PCRil võimalik lihtsasti paljundada (Schoch jt 2012). Sugukondade ja seltside eristamiseks on ITS piirkond liialt varieerunud (Lindahl jt 2013). Võrreldes tuuma SSU ja LSU rRNA geenijärjestusega, on ITS regioonis nukleotiidide variatsioon kõrgem (Schoch jt 2012).

3.2 Tuuma rRNA väikest subühikut (SSU rRNA) kodeeriv geen

Tuuma rRNA SSU kodeeriv geen on kõige sagedamini kasutatav markerpiirkond krohmseente määramiseks (Öpik jt 2014). Paljude teiste seenerühmade jaoks on markerpiirkond liiga konserveerunud (Begerow jt 2010). Tegemist on esimese markeriga, mida hakati laialdaselt kasutama krohmseente taksonoomias (Öpik jt 2014). Võrreldes ITS regiooniga, on SSU rRNA geen konserveerunud (Lindahl jt

2013). SSU rRNA geenijärjestusel põhinevast krohmseente liigisisest varieeruvusest on vähe teada (Stockinger jt 2010).

3.3 Tuuma rRNA suurt subühikut (LSU rRNA) kodeeriv geen

Krohmseente tuuma LSU rRNA geenijärjestusel on kõrge liigisisene varieeruvus (Schoch jt 2012). Seda geeni on võimalik kasutada markerpiirkonnana paljudes seente hõimkondades (Seifert 2009). LSU geenijärjestustel põhinevaid andmebaase on vähe (Krüger jt 2009). Võrreldes ITS piirkonnaga, on LSU rRNA geen ja SSU rRNA geen konserveerunud (Lindahl jt 2013). LSU rRNA kodeerivat geeni ja ITS regiooni on võimalik omavahel kombineerida (Stockinger jt 2010).

Tabel 1. DNA-põhises määramises sagedamini kasutatavate markerpiirkondade võrdlus. Tabel põhineb Begerow jt 2010, Hibbett & Taylor 2013, Krüger jt 2009, Lindahl jt 2013, Schoch jt 2012, Seifert 2009, Stockinger jt 2010, Öpik jt 2014 andmetel

	ITS piirkond	SSU rRNA geen	LSU rRNA geen
Kasutatavad seente hõimkonnad	Paljud seente hõimkonnad	Peamiselt krohmseened	Paljud seente hõimkonnad
Eelised	Suur varieeruvus	Konserveerunum kui ITS regioon	Konserveerunum kui ITS regioon
Puudused	Sugukondade ja seltside eristamiseks liialt varieerunud	Paljude seente hõimkondade jaoks liialt konserveerunud	LSU järjestustel põhinevaid andmebaase on vähe
Kasutusvaldkond krohmseente uurimustes	Sagedamini taksonoomilistes uurimustes	Sagedamini ökoloogilistes uurimustes	Sagedamini taksonoomilistes uurimustes
Markeri liigisisene varieeruvus krohmseente hõimkonnas	Kõrgem	Vähe teada	Kõrgem
Markeri nukleotiidide variatsioon krohmseente hõimkonnas	Kõrgem	Madalam	Madalam

4. Krohmseente järjestusepõhine rühmitamine

Krohmseente järjestusepõhisel rühmitamisel saadud grupe nimetatakse erinevalt, nagu fülotüübid, fülogrupid ja MOTUd (Öpik jt 2014). Nendele antakse nimetus või kood. Uus metoodika taksonite leidmiseks on tekkinud tänu molekulaarökoloogidele (Hibbett jt 2011).

Üheks põhiliseks probleemiks järjestusepõhisel määramisel on fülogruppide piiritlemine (Herr jt 2015). Erinevates uuringutes ja meeskondades kasutatakse selleks erinevaid meetodeid ja algoritme (Öpik jt 2014). Taksonite piiritlemise probleemiks on ühtsete standardite puudumine, mille tulemusena võidakse erinevates töörühmades samu andmeid kasutades saada erinevaid tulemusi (Herr jt 2015). Selles peatükis annan ülevaate peamistest järjestuspõhistest rühmadest. Tabelis 2 on välja toodud nende peamised erinevused.

4.1 Virtuaaltakson

Üheks võimaluseks krohmseente klassifitseerimiseks on looduslikest proovidest saadud DNA järjestuste rühmitamine ja joondamine fülogeneesianalüüsil virtuaaltaksoniteks (Öpik 2012). Kõige rohkem kasutatakse virtuaaltaksoni piiritlemiseks tuuma SSU rRNA geeni (Öpik jt 2014). Virtuaaltaksonid on fülogeneetiliselt määratletud järjestuste grupid, mis ligikaudu vastavad liigitasemele (Öpik jt 2014). Virtuaaltaksonite abil on võimalik kirjeldada krohmseente liigirikkuse mustreid (Öpik jt 2014). Ligikaudu liigitasemel saadud määratlused nimetatakse perekonnanime ning virtuaaltaksoni numbriga (Öpik 2012). Virtuaaltaksoneid on võimalik vajadusel mitmeks jagada või kokku liita (Öpik jt 2014). Virtuaaltaksoni piiritlemisel kasutatavad tüüpsekventsidsid valib manuaalselt kuraator ning need peavad olema nii pikad ja kvaliteetsed kui võimalik (Öpik jt 2014). Võimalusel valitakse virtuaaltaksoni tüüpjärjestused sellistest sekventsidsidest, mis pärinevad kultuuridest (Öpik jt 2014). 2013 aasta andmetel on MaarjAM andmebaasis 348 SSU rRNA geenijärjestusel põhinevat virtuaaltaksonit (Öpik jt 2014).

4.2 Liigihüpotees

Liigihüpoteesi kasutatakse paljudes erinevates seente hõimkondades, mis võimalusel vastavad ligikaudu liigitasemele (Kõljalg jt 2013). Liigihüpoteesi taksonite

klasterdamine põhineb tuuma ribosomaalsel ITS regioonil (Kõljalg jt 2013). Liigihüpoteesi nimetus koosneb organismi nimest, sekventsikoodist ja liigihüpoteesi koodist (Kõljalg jt 2013). Liigihüpoteesi piiritlemiseks kasutatakse referentsjärjestusi ja esindusjärjestusi (Kõljalg jt 2013). Tuleb jälgida, et automaatselt valitud esindusjärjestused oleksid taksonoomiliselt kõige sobivamad (Kõljalg jt 2013). Kuraatorid võivad esindusjärjestust ignoreerida ja selle asemel määrata referentsjärjestused, mis võivad pärineda erinevast bioloogilisest materjalist (herbaareksemplar, elusorganism, vesi, õhk, koeproovid jne) (Kõljalg jt 2013).

4.3 MOTU (ingl *molecular operational taxonomic unit*)

DNA järjestusi on võimalik rühmitada MOTUdeks (Schloss & Handelsman 2005). Molekulaarökoloogilised uurimused kasutavad MOTUsid eeldusena, et tegemist on ligikaudu liigitasemel vastavate rühmitustega (Hibbett jt 2011). MOTUde puhul on tegemist ükskõik mis järjestuse klastriga (Schloss & Handelsman 2005). Kivlin jt leidsid oma metauuringus, et SSU rRNA geenil põhinevaid MOTUsid on 967 ja LSU rRNA geenil põhinevaid 1159 (Kivlin jt 2011). Nende tulemuste saamiseks kasutati geenijärjestuste piiritlemiseks 99% sarnasusläve. MOTUde piiritlemiseks puuduvad üldiselt kokkulepitud standardkriteeriumid (Hibbett jt 2011). MOTUde nimetamisel kasutatakse pigem numbrilisi koode kui ladinakeelset nimetust (Hibbett jt 2011).

Tabel 2. Järjestusepõhisel rühmitamisel saadud gruppide võrdlus. Tabel põhineb Hibbett jt 2011, Kõljalg jt 2013, Schloss & Handelsman 2005, Öpik 2012 ja Öpik jt 2014 andmetel

	Virtuaaltakson	Liigihüpotees	MOTU
Nimetamine	Perekonnanime ning virtuaaltaksoni numbriga	Organismi nimi, sekventsikood ja liigihüpoteesi kood	Numbriliselt
Piiritlemine	Tüüpjärjestustega	Tüüpjärjestusega (referentsjärjestus ja esindusjärjestus)	Puuduvad üldiselt kokkulepitud kriteeriumid.
Peamine kasutatav markerpiirkond	Tuuma SSU rRNA geen	ITS piirkond	Ükskõik mis järjestuste klaster

5. Andmebaasid

Organismide DNA-põhisel määramisel on oluline andmebaaside olemasolu, kus saab võrrelda referentsjärjestusi uuritava järjestusega (Hart jt 2015). Andmebaasides olevaid sekventse täiendavad metaandmed, mis sisaldavad informatsiooni krohmseente fülogeneetilisest asetusest, levikust, arvukusest, ökoloogiast ja biokeemiast (Herr jt 2015).

Järjestusepõhise määramise üheks väljakutseks on usaldusväärsete nukleiinhappe andmebaaside arendamine, mis sisaldaks korrektseid referentsjärjestusi metaandmetega (Herr jt 2015). Referentsjärjestusi peaks hoiustama iseseisvas andmebaasis, mis on usaldusväärsed ja disainitud määramise eesmärgil (Hart jt 2015). Peamises suures järjestuste andmebaasis INSDC (ingl *International Nucleotide Sequence Database Collaboration*) leidub suures osas sissekandeid, mis ei ole täielikult tuvastatud liigitasemeni (Nilsson jt 2006). Lisaks võivad selles puududa ka olulised metaandmed (nt proovi päritolu) (Tedersoo jt 2011). Seetõttu on loodud väiksemaid andmebaase, mis keskenduvad DNA-põhiste liikide määramisele.

5.1 MaarjAM andmebaas

MaarjAM andmebaas koondab keskkonnaproovidest ja seenekultuuridest saadud krohmseente DNA järjestused (Öpik 2012). MaarjAM on hetkel ainus andmebaas, mis sisaldab ainult krohmseente järjestusi (Öpik jt 2010). Andmebaasi on kogutud referentsjärjestused, mida saab kasutada uute järjestuste määramiseks. Algselt põhines MaarjAM SSU rRNA kodeerival geenijärjestusel, sest tegemist on kõige rohkem kasutatava markeriga krohmseente uuringutes (Öpik jt 2014). Nüüdseks sisaldab andmebaas lisaks veel ITS regiooni, tuuma LSU rRNA kodeerivat geenijärjestust, valke kodeerivaid geene ja mitokondriaalseid markerpiirkondi (mt LSU) (Öpik jt 2010). Andmebaasis olevad järjestused, millel puudub sarnase DNA järjestusega varem sekveneeritud liik, rühmitatakse fülogeneesianalüüsil virtuaaltaksoniteks (Öpik jt 2014). Sekventsidega on seotud metaandmed, milles on kirjas proovi päritolu, peremeestaim, bioom jne. (Öpik jt 2010). MaarjAM andmebaasi uuendatakse pidevalt (Öpik jt 2010).

5.2 UNITE andmebaas

UNITE andmebaas hõlmab seente ITS järjestusi, mille peamiseks eesmärgiks on pakkuda molekulaarse määramise jaoks referentsjärjestusi, mille sekvents ja nimetus sobiks kindlasti kokku (Kõljalg jt 2013). Esimene versioon andmebaasist keskendus ainult ektomükoriisa seente ITS järjestustele. UNITE andmebaas jälgib Index Fungorum-i nomenklatuuri (Kõljalg 2013). Nüüdseks on andmebaas globaalne ja sisaldab ka arbuskulaarset mükoriisat moodustavate seente ITS järjestusi. UNITE andmebaasi uuendatakse väljalasetena kaks korda aastas (Kõljalg jt 2013). 67% andmebaasis olevatest krohmseente ITS järjestustest on piiritletud ainult perekonna tasemeni (Hart jt 2015). UNITE andmebaasis on kasutusel kaheosaline klasterdamisprotsess, mille lõpptulemusena moodustatakse väiksematest klasterdustest liigihüpoteesid (Kõljalg jt 2013).

5.3 PHYMYCO-DB

PHYMYCO-DB sisaldab seente SSU rRNA kodeeriva geeni ja EF1-alfa markereid. (Mahé jt 2012). 2012 aasta andmetel sisaldab PHYMYCO-DB 9120 SSU rRNA geenijärjestust ja 672 EF1-alfa markereid (Mahé jt 2012). PHYMYCO-DB üheks eesmärgiks on arendada lihtsasti kasutatav andmebaas, mis sisaldab kõrge-kvaliteedilisi seente SSU rRNA kodeeriva geeni ja EF1-alfa markereid (Mahé jt 2012). Andmebaasi uuendatakse automaatselt 4 korda aastas ning peale automaatset järjestamist kontrollib kuraator andmed üle (Mahé jt 2012).

5.4 SILVA andmebaas

SILVA andmebaas sisaldab bakterite, arhede ja eukarüootide (k.a seened) rRNA geenijärjestusi (Quast jt 2012). Seda andmebaasi ei uuendata pidevalt, vaid väljalasetena (Quast jt 2012). Juuli 2012 seisuga sisaldas andmebaas 3 194 778 SSU rRNA ja 288 717 LSU rRNA geenijärjestust (Quast jt 2012).

5.5 Nomenklatuuri andmebaasid

Mükoloogid peavad uute seeneliikide liiginimed registreerima avalikult kättesaadavatesse andmebaasidesse, nagu MycoBank või Index Fungorum (Hibbett & Taylor 2013). See on oluline, et ei tekiks olukorda, kus sama liiginimi on kasutusel erinevatel organismidel. Ekspertid kontrollivad pakutud nimede kehtivust, õigsust ja

lingvistilist korrektsust (Crous jt 2004). Pakutud nimed jäävad konfidentsiaalseks kuni publitseerimiseni (Crous jt 2014). Suur osa krohmseente taksonitest, mis pärinevad ainult metagenoomika uurimustest, ei ole nimetatud. Seetõttu pole võimalik neid sisestada ka nimedepõhisesse andmebaasi (Herr jt 2015).

6. Molekulaarsete meetodite kasutamine krohmseente taksonoomias

Viimastel aastatel on krohmseente taksonoomia pidevas muutumises, sisaldades ka ümberkorraldusi kõrgemates taksonites (Schüßler & Walker 2010, Oehl jt 2011). 2001. aastal lisasid Schüßler ja Kluge krohmseente hulka seene *Geosiphon pyriforme*, mille eripäraks on endosümbioos tsüanobakteriga *Nostoc punctiforme*. Saadud tulemus põhines rRNA SSU geenijärjestuse uuringul (Schüßler & Kluge 2001).

Taksonoomide tegevuse käigus saadud tulemusi kasutavad näiteks ökoloogid. Ökoloogidel on oluline teada, milline on organismide taksonoomiline kuuluvus. Taksonoomid peaksid kasutama taksonoomia koostamisel nii morfoloogilisi kui ka molekulaarseid tunnuseid (Herr jt 2015). See on oluline, et ökoloogid, kellel on isendi kohta ainult DNA sekvents, saaksid oma andmeid võrrelda olemasolevatega.

Viimastel aastatel on ilmunud kaks taksonoomilist tööd, mis sisaldavad molekulaarseid andmeid. 2010. aastal koostasid Schüßler ja Walker taksonoomilise tööd, mis põhines ainult molekulaarsel analüüsil (Schüßler & Walker 2010). Sellele järgnenud Oehl jt poolt koostatud töös kasutati nii molekulaarseid kui ka morfoloogilisi meetodeid (Oehl jt 2011). Saadud tulemused on omavahel erinevad. Näiteks Schüßler ja Walker jaotasid sugukonna *Glomeraceae* perekondadeks *Glomus*, *Funneliformis*, *Sclerocystis* ja *Rhizophagus* (Schüßler & Walker 2010). Oehl jt jaotasid selle sugukonna perekondadeks *Glomus*, *Funneliformis*, *Septoglomus* ja *Simiglomus* (Oehl jt 2011).

Kokkuvõte

Kuni viimase kümnendini kasutati krohmseente määramiseks peamiselt morfoloogilisi tunnuseid. Tänapäevaks on lisandunud juurde DNA järjestustel põhinevad meetodid. Käesolev töö annab kirjandusel põhineva lühiülevaate krohmseente DNA-põhisest määramisest.

Põhjalikumalt tutvustan krohmseente määramisel sagedamini kasutatavaid rRNA operoni markerpiirkondi. Markerpiirkondadest paljundatakse PCRi abil krohmseente DNA järjestused. Saadud järjestused sisestatakse nukleiinhapete andmebaasidesse. Andmebaaside arendamine on järjestusepõhise määramise üheks väljakutseks. Töös antakse lühiülevaate MaarjAM, UNITE, PHYMYCO, SILVA ja nomenklatuuri andmebaasidest.

Peamiseks probleemiks krohmseente DNA-põhisest määramises on molekulaarsete taksonite piiritlemine. Hetkel puuduvad ühtsed kriteeriumid DNA-põhiste liikide piiritlemiseks. DNA Järjestusi on võimalik klassifitseerida rühmadeks, mis võimalusel vastavad liigitasemele. Käesolev töö tutvustab põhjalikumalt järjestusepõhistest rühmadest MOTU, virtuaaltaksoni ja liigihüpoteesi kontseptsioone.

Lõpetuseks annab töö ülevaate molekulaarsete meetodite kasutamisest krohmseente taksonoomias, mis on viimastel aastatel pidevas muutumises. Selle juures on oluline, et taksonoomid koostaksid võrreldavaid andmeid, mida oleks võimalik kasutada ka DNA-põhiselt määratud isendite puhul.

Summary

DNA-based identification of Glomeromycota

Until recent decades the main means of identification of Glomeromycota was to use morphological methods. Nowadays, DNA- based methods have been implemented. The current thesis aims to give a brief overview of the DNA-based identification methods of Glomeromycota.

The more frequent methods, which involve nuclear ribosomal operon are introduced in more detail. Markers are used with PCR to replicate Glomeromycota DNA sequences. The resulting sequences are then entered to nucleic acid databases. The development of the aforementioned databases is one of the main challenges of sequence based identification. The thesis gives an overview of the MaarjAM, UNITE, PHYMYCO, SILVA and nomenclature databases.

The biggest problem in regards of DNA-based identification is the delimitation of the molecular taxa. Currently there are no common criteria to delimitate DNA-based species. DNA sequences can be however classified into sequence groups that correspond roughly to the species level. From the sequence based methods MOTU, virtual taxonomy and species hypothesis concepts are discussed.

Lastly, the thesis gave an overview of the molecular methods in regards to taxonomic use. Aforementioned methods have been in constant change during the last years. It is also important, that taxonomists would create comparable data so it could be used for DNA-based identification of individual organisms.

Tänuavaldused

Käesoleva töö valmimise eest tahaksin tänada oma juhendajat Maarja Öpikut.

Kasutatud kirjandus

Begerow, D., Nilsson, H., Unterseher, M., & Maier, W. 2010. Current state and perspectives of fungal DNA barcoding and rapid identification procedures. *Applied Microbiology and Biotechnology* 87: 99-108.

Crous, P., Gams, W., Stalpers, J., Robert, V., & Stegehuis, G. 2004. MycoBank: an online initiative to launch mycology into the 21st century. *Studies in Mycology* 50: 19–22.

Gollotte, A., van Tuinen, D., & Atkinson, D. 2004. Diversity of arbuscular mycorrhizal fungi colonising roots of the grass species *Agrostis capillaris* and *Lolium perenne* in a field experiment. *Mycorrhiza* 14: 111-117.

Heinaru, A. 2012. *Geneetika*. Tartu Ülikooli Kirjastus, Tartu.

Herr, J., Öpik, M., & Hibbett, D. 2015. Towards the unification of sequence-based classification and sequence-based identification of host-associated microorganisms. *New Phytologist* 205: 27-31.

Hart, M., Aleklett, K., Chagnon, P., Egan, C., Ghignone, S., Helgason, T., Lekberg, Y., Öpik, M., Pickles, B., & Waller, L. 2015. Navigating the labyrinth: a guide to sequence-based, community ecology of arbuscular mycorrhizal fungi. *New Phytologist*. doi: 10.1111/nph.13340

Helgason, T., Fitter, A., & Young, J. 1999. Molecular diversity of arbuscular mycorrhizal fungi colonising *Hyacinthoides non-scripta* (bluebell) in a seminatural woodland. *Molecular Ecology* 8: 659-666.

Hibbett, D., Ohman, A., Glotzer, D., Nuhn, M., Kirk, P., & Nilsson, R. 2011. Progress in molecular and morphological taxon discovery in Fungi and options for formal classification of environmental sequences. *Fungal Biology Reviews* 25: 38-47.

Hibbett, D., & Taylor, J. 2013. Fungal systematics: is a new age of enlightenment at hand?. *Nature Reviews Microbiology* 11: 129-133.

Janoušková, M., Püschel, D., Hujslová, M., Slavíková, R., & Jansa, J. 2015. Quantification of arbuscular mycorrhizal fungal DNA in roots: how important is material preservation?. *Mycorrhiza* 25: 205-214.

Kivlin, S., Hawkes, C., & Treseder, K. 2011. Global diversity and distribution of arbuscular mycorrhizal fungi. *Soil Biology and Biochemistry* 43: 2294-2303.

Krüger, M., Stockinger, H., Krüger, C., & Schüßler, A. 2009. DNA-based species level detection of Glomeromycota : one PCR primer set for all arbuscular mycorrhizal fungi. *New Phytologist* 183: 212-223.

Kõljalg, U., Nilsson, R., Abarenkov, K., Tedersoo, L., Taylor, A., Bahram, M., Bates, S., Bruns, T., Bengtsson-Palme, J., Callaghan, T., Douglas, B., Drenkhan, T., Eberhardt, U., Dueñas, M., Grebenc, T., Griffith, G., Hartmann, M., Kirk, P., Kohout, P., Larsson, E., Lindahl, B., Lücking, R., Martín, M., Matheny, P., Nguyen, N., Niskanen, T., Oja, J., Peay, K., Peintner, U., Peterson, M., Põldmaa, K., Saag, L., Saar, I., Schüßler, A., Scott, J., Senés, C., Smith, M., Suija, A., Taylor, D., Telleria, M., Weiss, M., & Larsson, K. 2013. Towards a unified paradigm for sequence-based identification of fungi. *Molecular Ecology* 22: 5271-5277.

Lee, J., Lee, S., & Young, J. 2008. Improved PCR primers for the detection and identification of arbuscular mycorrhizal fungi. *FEMS Microbiology Ecology* 65: 339-349.

Lindahl, B., Nilsson, R., Tedersoo, L., Abarenkov, K., Carlsen, T., Kjølner, R., Kõljalg, U., Pennanen, T., Rosendahl, S., Stenlid, J., & Kauserud, H. 2013. Fungal community analysis by high-throughput sequencing of amplified markers - a user's guide. *New Phytologist* 199: 288-299.

Mahé, S., Duhamel, M., Le Calvez, T., Guillot, L., Sarbu, L., Bretaudeau, A., Collin, O., Dufresne, A., Kiers, E., & Vandenkoornhuyse, P. 2012. PHYMYCO-DB: a curated database for analyses of fungal diversity and evolution. *PLoS ONE* 7: e43117.

McNeill, J., Barrie, F.R., Buck, W.R., Demoulin, V., Greuter, W., Hawksworth, D.L., Herendeen, P.S., Knapp, S. Marhold, K., Prado, J., Prudhomme van Reine, W.F., Smith, G.F., Wiersema, J.H., & Turland, N.J. 2012. International Code of Botanical Nomenclature (Melbourne Code) adopted by the Eighteenth International Botanical Congress Melbourne, Australia, July 2011.

Msiska, Z., & Morton, J. 2009. Phylogenetic analysis of the Glomeromycota by partial β -tubulin gene sequences. *Mycorrhiza* 19: 247-254.

Nilsson, R., Ryberg, M., Kristiansson, E., Abarenkov, K., Larsson, K., & Kõljalg, U. 2006. Taxonomic Reliability of DNA Sequences in Public Sequence Databases: A Fungal Perspective. *PLoS ONE* 1: e59.

Oehl, F., Silva, G.A., Goto, B.T., & Ewald, S. 2011. Glomeromycota: three new genera and glomoid species reorganized. *Mycotaxon* 116: 75-120.

Parmasto, E. 1996. *Biosüsteematika teooria ja meetodid*. Tartu Ülikooli Kirjastuse Trükikoda, Tartu.

Parniske, M. 2008. Arbuscular mycorrhiza: the mother of plant root endosymbioses. *Nature Reviews Microbiology* 6: 763-775.

Quast, C., Pruesse, E., Yilmaz, P., Gerken, J., Schweer, T., Yarza, P., Peplies, J., & Glockner, F. 2012. The SILVA ribosomal RNA gene database project: improved data processing and web-based tools. *Nucleic Acids Research* 41: D590-D596.

Robinson-Boyer, L., Grzyb, I., & Jeffries, P. 2009. Shifting the balance from qualitative to quantitative analysis of arbuscular mycorrhizal communities in field soils. *Fungal Ecology* 2: 1-9.

Schoch, C.L., Seifert, K.A., Huhndorf, S., Robert, V., Spouge, J.L., Levesque, C.A., Chen, W., & Fungal Barcoding Consortium, . 2012. Nuclear ribosomal internal transcribed spacer (ITS) region as a universal DNA barcode marker for Fungi. *Proceedings of the National Academy of Sciences of the United States of America* 109: 6241–6246.

Seifert, K. 2009. Progress towards DNA barcoding of fungi. *Molecular Ecology Resources* 9: 83-89.

Schüßler, A., & Kluge, M. 2001. Geosiphon pyriforme, an endocytosymbiosis between fungus and cyanobacteria, and its meaning as a model system for arbuscular mycorrhizal research. In Hock, B. (ed.), *The Mycota IX*, pp. 151-161. Springer, Berlin.

Schüssler, A., and Walker, C. 2010 *The Glomeromycota: a species list with new families and new genera*. Gloucester, UK: Arthur Schüssler and Christopher Walker. ISBN- 13: 978-1466388048, ISBN-10: 1466388048. URL [http:// www.amf-phylogeny.com](http://www.amf-phylogeny.com). Published in December 2010 in libraries at The Royal Botanic Garden Edinburgh, The Royal Botanic Garden Kew, Botanische Staatssammlung Munich, and Oregon State University. Electronic copy freely available online at URL <http://www.amf-phylogeny.com>.

Schloss, P., & Handelsman, J. 2005. Introducing DOTUR, a Computer Program for Defining Operational Taxonomic Units and Estimating Species Richness. *Applied and Environmental Microbiology* 71: 1501-1506.

Smith, S., & Read, D. 2008. *Mycorrhizal symbiosis*. Academic Press, Amsterdam.

Stockinger, H., Walker, C., & Schüßler, A. 2009. ' Glomus intraradices DAOM197198', a model fungus in arbuscular mycorrhiza research, is not Glomus intraradices. *New Phytologist* 183: 1176-1187.

Stockinger, H., Krüger, M., & Schüßler, A. 2010. DNA barcoding of arbuscular mycorrhizal fungi. *New Phytologist* 187: 461-474.

Stockinger, H., Peyret-Guzzon, M., Koegel, S., Bouffaud, M., & Redecker, D. 2014. The largest subunit of RNA polymerase II as a new marker gene to study assemblages of arbuscular mycorrhizal fungi in the field. *PLoS ONE* 9: e107783.

Tedersoo, L., Abarenkov, K., Nilsson, R., Schüssler, A., Grelet, G., Kohout, P., Oja, J., Bonito, G., Veldre, V., Jairus, T., Ryberg, M., Larsson, K., & Kõljalg, U. 2011. Tidying up international nucleotide sequence databases: ecological, geographical and sequence quality annotation of ITS sequences of mycorrhizal fungi. *PLoS ONE* 6: e24940.

Thiéry, O., Börstler, B., Ineichen, K., & Redecker, D. 2010. Evolutionary dynamics of introns and homing endonuclease ORFs in a region of the large subunit of the mitochondrial rRNA in Glomus species (arbuscular mycorrhizal fungi, Glomeromycota). *Molecular Phylogenetics and Evolution* 55: 599-610.

Wubet, T., Weiß, M., Kottke, I., Teketay, D., & Oberwinkler, F. 2004. Molecular diversity of arbuscular mycorrhizal fungi in Prunus africana, an endangered medicinal tree species in dry Afromontane forests of Ethiopia. *New Phytologist* 161: 517-528.

Öpik, M., Vanatoa, A., Vanatoa, E., Moora, M., Davison, J., Kalwij, J., Reier, Ü., & Zobel, M. 2010. The online database MaarjAM reveals global and ecosystemic distribution patterns in arbuscular mycorrhizal fungi (Glomeromycota). *New Phytologist* 188: 223-241.

Öpik, M. 2012. DNA-liigid ehk kuidas uurida organisme, keda oleme näinud vaid DNA järjestustena. In Öpik, M., Laanisto, L., Vanatoa, A. & Kull, K. (ed.), *Schola Biotheoretica XXXVIII Elurikkuse mõte ja mõõt.*, pp. 94-100. Eesti Loodusuurijate Selts, Tartu.

Öpik, M., Davison, J., Moora, M., & Zobel, M. 2014. DNA-based detection and identification of Glomeromycota: the virtual taxonomy of environmental sequences. *Botany* 92: 135-147.

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, Kätlin Lindsaar,

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

KROHMSEENTE DNA-PÕHINE MÄÄRAMINE,

mille juhendaja on Maarja Öpik,

- 1.1. reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;
 - 1.2. üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu kuni autoriõiguse kehtivuse tähtaja lõppemiseni.
2. olen teadlik, et punktis 1 nimetatud õigused jäävad alles ka autorile.
3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, 25.05.2015