

Vantaanjoen simpukkakartoitus eDNA- menetelmällä

**Jouni Leinikki, Micaela Hellström, Panu Oulasvirta,
Elli Leinikki, Juha Syväranta ja Myyri Saarman**



Alleco

MARINE BIOLOGICAL AND LIMNOLOGICAL CONSULTANTS

Veneentekijäntie 4

FI-00210 Helsinki, Finland

Tel. +358 (0)45 679 0300

OTSIKKO: Vantaanjoen simpukkakartoitus eDNA-menetelmällä

PÄIVÄMÄÄRÄ: 9.12.2021

TEKIJÄ(T): Jouni Leinikki, Micaela Hellström, Panu Oulasvirta, Elli Leinikki, Juha Syväranta ja Myyri Saarman

JULKAISU: Alleco Oy raportti n:o 30/2021

JULKAISIJA: Alleco Oy, Veneentekijäntie 4, 00210 Helsinki, <http://www.alleco.fi>

VIITTAUSOHJE: Leinikki, J., Hellström, M., Oulasvirta, P., Leinikki, E., Syväranta, J. & Saarman, M. 2021. Vantaanjoen simpukkakartoitus eDNA-menetelmällä. Alleco Oy raportti n:o 30/2021. Alleco Oy 9.12.2021.

KANSIKUVA: Vesinäytteen suodatusta näytepisteellä VA4 (Micaela Hellström 2021)

Sisällys

Johdanto	4
Tutkimusalue ja menetelmät	5
Tulokset	8
Kalalajit	9
Tulosten tarkastelu	10
Johtopäätökset	11
Lähteet	11
LIITE 1: Single and multiple species analyses	13
LIITE 2. Laboratory methods – Metabarcoding	14
B.2.1 Extraction	14
B.2.2. PCR metabarcoding	14
B.2.2.1 Freshwater mussels (Unionidae)	14
<i>B.2.2.2 Fish</i>	14
B.2.3. Bioinformatics and species identification	15
LIITE 3: Required Quality controls- from collection to metabarcoding	16
LIITE 4: Quality control results	17
LIITE 5: Distribution of sequence reads for the different species in each sample	18

Johdanto

Vantaanjoki Natura 2000 -alueeseen kuuluu Vantaanjoen pääuoman 59 kilometrin pituinen osa. Joessa on yli neljäkymmentä koskea ja siihen laskee useita sivujokia, jotka eivät kuitenkaan sisälly Natura 2000-alueeseen. (Ympäristöhallinto 2021). Keskeisenä perusteena Natura-alueen perustamiselle oli vuollejokisimpukka (*Unio crassus*), jota esiintyy Vantaanjoen pääuoman (Valovirta 2008) lisäksi sivujoista ainakin Tuusulanjoessa (Kivimäki ym. 2021), Keravanjoessa ja Luhtaanmäen-Lepsämänjoessa (Syväranta & Leinikki 2018).

Vuollejokisimpukka kuuluu EU:n luontodirektiivin liitteiden II ja IV lajeihin (92/43/ETY). Luontodirektiivin liitteen II lajit ovat **”yhteisön tärkeänä pitämiä lajeja, joiden suojelemiseksi on osoitettava erityisen suojelutoimen alueita”** (Natura 2000). Liitteen IV lajit ovat **”yhteisön tärkeänä pitämiä lajeja, jotka edellyttävät tiukkaa suojelua, ja joiden lisääntymis- ja levähdyspaikkojen häiritseminen ja heikentäminen on kielletty”**. Kansallisella tasolla vuollejokisimpukka on rauhoitettu ja on luokiteltu uhanalaisuutensa puolesta vaarantuneeksi (VU) lajiksi (LsL 1996/1096).

Vuollejokisimpukka esiintyy Suomen etelä- ja lounaisosissa. Suurimmat kannat ovat Vantaanjoessa (Valovirta 2008) ja Kokemäenjoessa (Leinikki & Leppänen 2014, Leinikki 2017, Leppänen ym. 2020). Laji vaatii virtaavaa vettä ja kaivautumiseen soveltuvaa pohjaa (Uudenmaan ELY-keskus, 2010). Lisääntymisen kannalta tärkeää on myös joen kalasto, sillä vuollejokisimpukka tarvitsee toukkavaiheessa väli-isännäkseen tiettyjen kalalajien poikasia. Soveltuvia ovat ainakin särkikalat ja kivisimppu sekä ahven. Lohikalat eivät sovellu väli-isänniksi vuollejokisimpukoille. Myös veden laatu on tärkeä tekijä simpukoiden menestymisessä. Hapen puute vaikeuttaa pohjaan kaivautuneiden pikkusimpukoiden elämää ja liiallisen nitraattipitoisuuden (yli 10 mg/l) tiedetään rajoittavan vuollejokisimpukan hedelmällisyyttä (Nordsieck, 2000). Douda (2009) havaitsi, että verrattaessa useita jokiosuuksia, vuollejokisimpukoiden esiintyminen väheni nitraattipitoisuuksien kasvaessa. Saman tutkimuksen laboratorioskokeissa ei kuitenkaan voitu havaita nitraatin aiheuttavan akuutteja myrkytysoireita simpukoilla.

Myös myrkylliset aineet voivat uhata vuollejokisimpukoita. Harjavallassa pääsi vuonna 2014 Kokemäenjokeen suuret määrät nikkelisulfaattia (NiSO_4), mikä tappoi välittömästi arviolta 1 000 000 vuollejokisimpukkayksilöä (Leppänen ym. 2020).

Vantaanjoen Natura-alueen vuollejokisimpukkainventoinnissa (Valovirta 2008) virtaussuuntaan nähden ylin esiintymä oli Nukarinkosken alapuolella, missä lajin osuus oli 1/3 kaikista simpukoista. Seuraavilla linjoilla ylävirtaan päin Rantakulman Jokiniemessä ei lajia enää havaittu.

Ympäristö-DNA, josta käytetään yleisesti lyhennettä eDNA (englanniksi environmental DNA), on molekyylibiologinen menetelmä eliöiden tutkimiseksi. Se perustuu tietoon siitä, että kaikista eliöistä vapautuu ympäristöön perimätietoa sisältäviä soluja. Menetelmällä analysoidavaksi sopivaa DNA:ta on esimerkiksi eläinten jäljissä, ihossa, karvoissa, limassa ja muissa eritteissä tai kasvien kappaleissa ja siitepölyssä (Taberlet et al. 2012, Pedersen et al. 2015, Bruce et al. 2021). Eristämällä, monistamalla ja analysoidamalla näiden materiaalien sisältämää DNA:ta voidaan muun muassa selvittää ympäristön lajistoa, lajien suhteellisia osuuksia sekä niiden ominaisuuksia. Käytettäessä eDNA-menetelmää tutkittaviin yksilöihin ei tarvitse kajota, aineisto saadaan vesinäytteistä. Näytteen tilavuus on tavallisesti 100 millilitrasta joihinkin litroiin. Virtaavasta vedestä otetusta näytteestä voidaan tutkia, mitä eliöitä esiintyy ylävirtaan päin näytteenottoaikaan verrattuna.

eDNA säilyy vedessä vain lyhyen aikaa, minkä vuoksi sen avulla voidaan tutkia myös lajiston muutoksia ajan suhteen (Li *et al.*, 2019, Brys *et al.*, 2021). Menetelmällä havaitaan myös lajeja, joiden pyydystäminen on vaikeaa tai kiellettyä niiden suojelun vuoksi, minkä vuoksi se soveltuu erinomaisesti luonnon monimuotoisuuden tutkimiseen. (Leese *et al.* 2016, Olds *et al.* 2016, Deiner *et al.* 2017, Bruce *et al.* 2021).

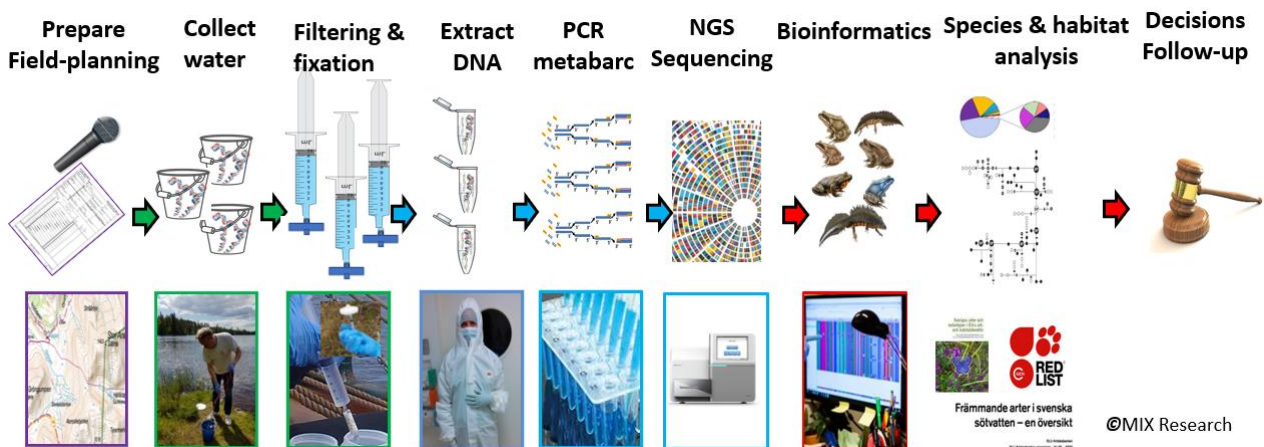
Alleco Oy selvitti yhdessä ruotsalaisen Mix Research -yhtiön kanssa eDNA-menetelmän käytettävyyttä Vantaanjoen vuollejokisimpukkapopulaation tutkimiseen. Samalla haluttiin saada tarkempaa tietoa lajin levinneisyydestä, jotta sen suojelutoimenpiteitä voidaan kohdistaa nykyistä tehokkaammin. Vantaanjoen Natura-alueen inventointien yhteydessä todettiin, että vuollejokisimpukkaa ei esiinny Nukarinkosken yläpuolella (Valovirta 2008). Sukellusmenetelmällä voidaan kattaa vain osa pitkän joen pinta-alasta, joten havaintoa ei voitu pitää täysin luotettavana rauhoitetun lajin suojelutason turvaamiseksi mm. erilaisten rakennushankkeiden aiheuttamien uhkien edessä.

Työn tilasi Uudenmaan ELY-keskus.

Tutkimusalue ja menetelmät

Työvaiheet on esitetty kuvassa 1 ja kuvailtu tarkemmin liitteissä 1–4. Vesinäytteet otettiin 14 ennalta määritellyssä sijainnissa Vantaanjoella (kuvat 2 ja 3, taulukko 1). Ylin näytteenottoaika oli joen alkulähteellä Erkylänjärven pohjoispäässä ja alin Seutulassa, missä vuollejokisimpukoita tiedettiin varmasti esiintyvän. Missään aikaisemmissa selvityksissä ei ole havaittu vuollejokisimpukoita Vantaanjoen uomassa Nukarinkosken yläpuolella. Koska tarkoituksena oli varmistaa lajien esiintymisen yläraja, aloitettiin näytteenotto yläjuoksulta alaspäin edeten.

Näytteenottoaika valittiin siten, että ne sijaitsivat kohtuullisen tasaisin välimatkoin toisistaan. Sivujokien suualueiden ylä- ja alapuolelle sijoitettiin näytepisteet, jotta niiden mahdollisesti Vantaanjoesta poikkeava lajikoostumus näkyisi tuloksissa.



Kuva 1: Ympäristö-DNA-tutkimuksen työvaiheet.

Näytteet otettiin ensimmäisen kerran toukokuussa 2021, jolloin Vantaanjoki tulvi voimakkaasti. Näytteet kuitenkin hylättiin analyysilaboratorion suosituksesta, jonka mukaan tulva-aikaan otetuissa näytteissä oli korkea hulevesien aiheuttama kontaminaatoriski. Analysoitavaksi lähetetyt näytteet otettiin heinäkuussa, kun joen vesi oli matalalla (taulukko 1).

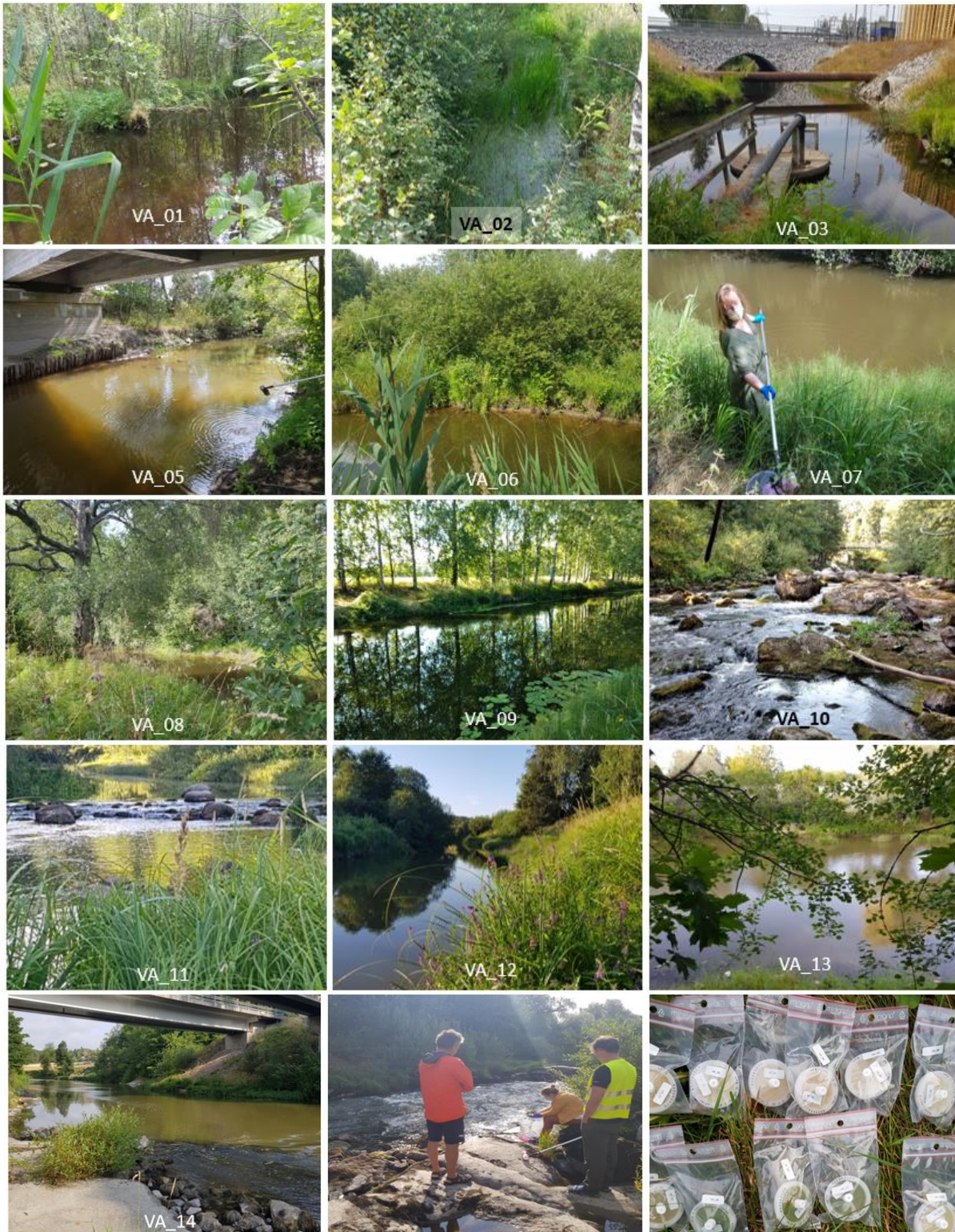
Taulukko 1: Näytteenottoaikojen koordinaatit ja kerätty mittausaineisto: Näytepaikan numero, päivämäärä, kellonaika, koordinaatit, suodatetun veden määrä, veden lämpötila, ilman lämpötila, näytteenottoisyvyys.

Sample name	Date	Time	WGS 84 N	WGS 84 E	V H2O ml	H ₂ O T °C	Air T °C	Sampling depth
VA_1	26/07/2021	12:02	60.7094	24.8787	2900	23	26	0.5
VA_2	26/07/2021	11:47	60.7368	24.8501	2700	14	28	0.5
VA_3	26/07/2021	11:18	60.7262	24.7773	3000	13	29	0.5
VA_4	26/07/2021	13:15	60.6091	24.7907	3300	19	26	0.5
VA_5	26/07/2021	14:05	60.6433	24.7983	3000	18	26	0.5
VA_6	26/07/2021	13:59	60.6426	24.7973	3000	18	26	0.5
VA_7	26/07/2021	14:30	60.6111	24.7943	2700	19	26	0.5
VA_8	26/07/2021	14:53	60.5928	24.8749	3000	19	26	0.5
VA_9	26/07/2021	17:45	60.5349	24.9131	2000	19	26	0.5
VA_10	26/07/2021	18:07	60.5264	24.9152	2000	23	26	0.5
VA_11	26/07/2021	18:30	60.4632	24.8583	2500	19	24	0.5
VA_12	26/07/2021	19:41	60.4139	24.8697	2700	23	25	0.5
VA_13	26/07/2021	20:28	60.3436	24.6556	2700	19	25	0.5
VA_14	29/07/2021	09:35	60.3436	24.8654	2700	20	23	0.5

Ennen näytteenottoa vesiastiat steriloidtiin ja puhdistettiin 10 % hypokloriittiliuoksella. Suodattamisessa käytettävät ruiskut, suodattimet, kertakäyttökäsineet ja hengityssuojat tilattiin yksittäin pakattuina DNA:sta puhdistettuihin pusseihin. Jalkineet puhdistettiin hypokloriittiliuoksella ennen kullekin näytteenottoalueelle menemistä ja jokeen astumista vältettiin.

Kultakin paikalta kerättiin yhteensä viisi litraa vettä osanäytteisiin. Vesi sekoitettiin puhdistetussa ämpärissä ja suodatettiin suljetuilla 5 µm GF / 0,8 µm PES suodattimilla (NatureMetrics, UK). Lisäksi otettiin päivittäin negatiiviset kontrollinäytteet kaupan pullovedestä mahdollisten kontaminaatioiden havaitsemiseksi. Näytteet suodatettiin ja kestäväitettiin Spens ym (2017) ohjeen mukaan.

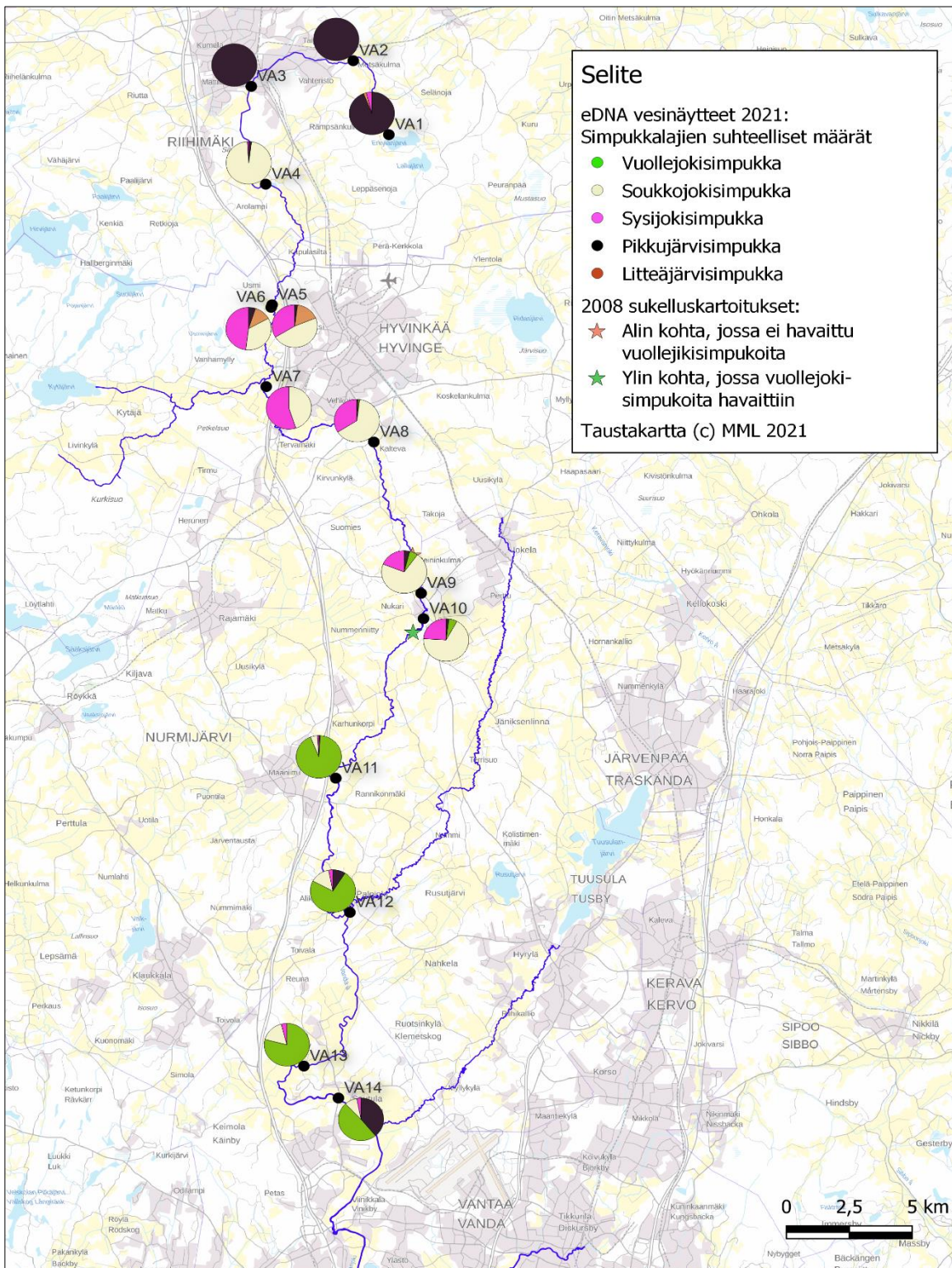
Liitteissä 1 ja 2 kuvataan laboratorioanalyysit ja luodaan yleiskatsaus molekyylibiologisiin menetelmiin. Liitteessä 3 luetellaan kentällä ja laboratoriossa kerättävät oheistiedot.



Kuva 2: Näytteenottoaikat VA1–VA14, lukuun ottamatta numeroa VA4, joka on kansikuvassa. Alarivin keskimääräinen kuva esittää näytteenottoa Vantaan Seutulassa ja alhaalla oikealla valmiiksi suodatettuja ja kestäväitä näytteitä.

Tulokset

Näytepisteiden sijainnit ja simpukkalajien osuudet kokonaisbiomassasta (luettujen lajikohtaisten DNA-sormenjälkien määrästä) on esitetty kuvassa 3.



Kuva 3: Simpukkalajien suhteelliset osuudet näytepaikoittain sekä vuoden 2008 sukelluskartoituksien (Valovirta 2008) ylin vuollejokisimpukoiden havaintopaikka Nukarinkosken alapuolella ja virtaussuuntaan nähden alin kohta, jolla laji ei havaittu. Ylin näytepaikka, jolla vuollejokisimpukkaa havaittiin vuonna 2021 oli VA8.

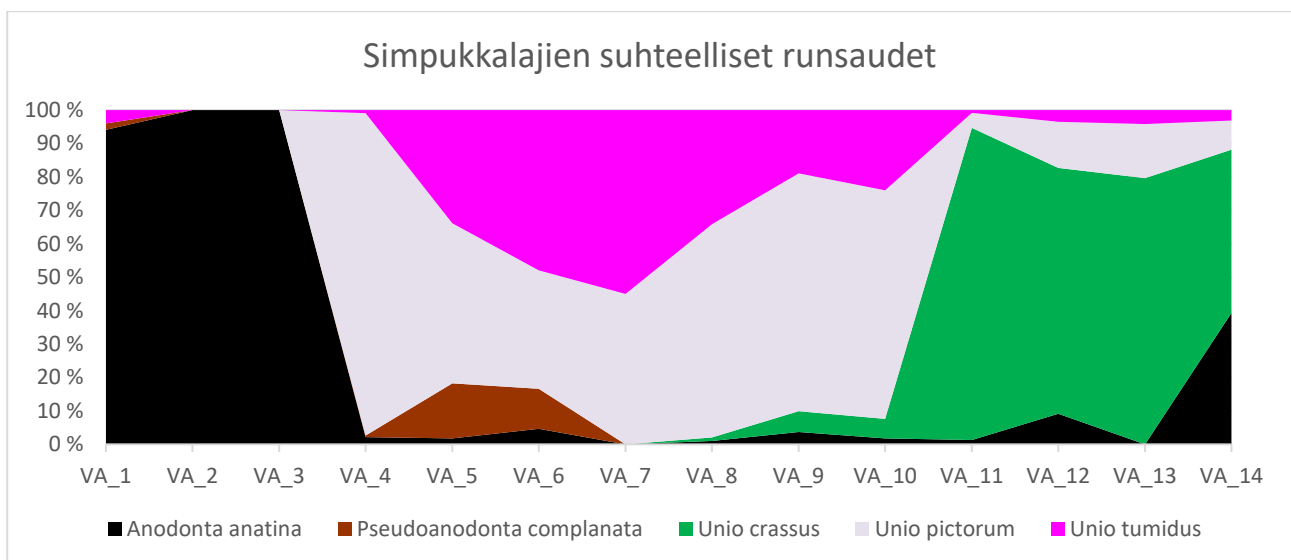
Monilajianalyseissä havaittiin kaikkiaan viisi suursimpukkalajia: litteäjärvisimpukka (*Pseudanodonta complanata*), pikkujärvisimpukka (*Anodonta anatina*), soukkojokisimpukka (*Unio pictorum*), sysijokisimpukka (*Unio tumidus*) ja vuollejokisimpukka (*Unio crassus*) (kuva 3).

Vuollejokisimpukkaa esiintyi tutkimusalueen alaosan näytteissä näytepisteelle VA8 (Kaltevantie) saakka. Lajin osuus oli verraten pieni Nukarinkosken yläpuolisilla pisteillä (1,1–6,2 %), kun sen osuus Pikkukosken ja Seutulan välisillä paikoilla oli 48,9–93,3 % (taulukko 2).

Taulukko 2: Simpukkalajien suhteelliset osuudet (%) näytepisteittäin.

Laji	VA1	VA2	VA3	VA4	VA5	VA6	VA7	VA8	VA9	VA10	VA11	VA12	VA13	VA14
<i>Anodonta anatina</i>	94	100	100	2,2	1,8	4,7	0	1	3,7	1,8	1,3	9,2	0	39,3
<i>Pseudanodonta complanata</i>	1,9	0	0	0,4	16,5	11,9	0	0	0	0	0	0	0	0
<i>Unio crassus</i>	0	0	0	0	0	0	0	1,1	6,2	5,8	93,3	73,5	79,2	48,9
<i>Unio pictorum</i>		0	0	96,5	47,8	35,5	45	63,8	71,2	68,4	4,5	13,8	16,1	8,7
<i>Unio tumidus</i>	4	0	0	0,9	33,9	47,9	55	34,1	18,9	24	0,8	3,5	4,1	3,1

Tulokset eivät kuitenkaan kerro eläinten kokonaismäärää, vaan ainoastaan niiden suhteellisen osuuden tutkituista lajeista. Litteäjärvisimpukkaa löydettiin tutkimusalueen yläosasta pisteistä VA1 sekä VA4-VA6. Lajin osuus oli suurimmillaan pisteessä VA5 (16,5 %), joka sijaitsee Usmintien sillan alla Hyvinkäällä. Pikkujärvisimpukkaa esiintyi suhteellisesti eniten Hyvinkään yläpuolisella jokiosuudella, kun sysijokisimpukkaa ja soukkojokisimpukkaa oli suhteellisesti eniten Hyvinkään alapuolisella jokiosuudella (kuvat 3 ja 4, taulukko 2).



Kuva 4: Simpukkalajien DNA-pitoisuuksien osuudet näytepisteittäin. Piste VA1 sijaitsee Erkylänjärven pohjoispäässä, piste VA4 Arolammin yläpuolella ja piste VA14 Seutulassa. Nukarinkoski sijaitsee pisteen VA10 alapuolella.

Kalalajit

Kalaston analysointi ei sisällynyt toimeksiantoomme, mutta halusimme kokeilla eDNA-menetelmän toimivuutta myös niihin. Kalalajit analysoitiin vain Pikkukosken pisteen VA11 näytteestä (paalupaikka n. 47 km). Paikalta on jokea myöten noin 10 kilometriä Nukarinkosken, jonne istutettiin mm. harjuksia vuonna 1990 (Mikkola & Saura 1994). Näytteestä havaittiin kaikkiaan 17 kalasekvenssiä, joista 15 määritettiin

lajitasolle. Toinen lajimääritystä vaille jäänyt geenikompleksi sisälsi vimpalle (*Vimba vimba*) ja pasurille (*Blicca bjoerkna*) yhteisen sekvenssin ja toinen seipin (*Leusiscus leusiscus*) ja säyneen (*Leusiscus idus*). Kalalajien DNA:n suhteelliset osuudet näytteessä VA11 on esitetty taulukossa 3.

Taulukko 3: Näytteessä VA11 havaitut kalalajit sekä niiden suhteelliset runsaudet.

Lajinimi	Tieteellinen nimi	Suhteellinen runsaus (%)
ahven	<i>Perca fluviatilis</i>	14,9
ankerias	<i>Anguilla anguilla</i>	0,1
hauki	<i>Esox lucius</i>	2,6
kiiski	<i>Gymnocephalus cernua</i>	10,5
kivisimppu	<i>Cottus gobio</i>	4,9
lahna	<i>Abramis brama</i>	3,8
taimen	<i>Salmo trutta</i>	1,5
made	<i>Lota lota</i>	<0,1
Pikkunahkiainen / nahkiainen	<i>Lampetra planeri</i> / <i>L. fluviatilis</i>	0,1
salakka	<i>Alburnus alburnus</i>	12,3
särki	<i>Rutilus rutilus</i>	43,3
seipi / säyne	<i>Leuciscus leuciscus</i> / <i>L. idus</i>	2,4
sorva	<i>Scardinius erythrophthalmus</i>	0,1
suutari	<i>Tinca tinca</i>	0,1
törö	<i>Gobio gobio</i>	2,6
toutain	<i>Aspius aspius</i>	0,1
vimpa / pasuri	<i>Vimba vimba</i> / <i>Blicca bjoerkna</i>	0,7

Runsaimmat kalalajit olivat särki, ahven ja salakka. Pikkunahkiaisen ja nahkiaisen sekvenssit ovat geneettisesti identtisiä, mutta eroavat meressä elävistä nahkiaisista. Lajit voivat risteytyä keskenään ja nahkiaisten taksonomiasta käydäänkin tieteellistä keskustelua (Beamish 1980, De Cashan *et al.* 2020). Harjuksen DNA:ta ei ollut näytteissä, vaikka laji oli istutettu Nukarinkosken vuonna 1990.

Tulosten tarkastelu

Vuollejokisimpukan esiintymisalue Vantaanjoessa ulottuu hieman pohjoisemmaksi kuin kartoituksissa vuonna 2008, jolloin laji ei löydetty lainkaan Nukarinkosken yläpuolisilta sukelluslinjoilta (Valovirta 2008). Vuonna 2021 vesinäytteessä havaittiin vuollejokisimpukan DNA:ta Kaltevantien sillan kohdalla, joka sijaitsee noin 11 jokikilometriä ylempänä kuin vuonna 2008 ylimpinä havaitut vuollejokisimpukat Nukarinkosken alapuolella. Hyvinkään Kuninkaankoskella vuollejokisimpukan DNA:ta ei enää havaittu, joten lajin pohjoinen levinneisyysraja Vantaanjoen pääuomassa lienee jossain Hyvinkään keskustan kohdalla.

Litteäjärvisimpukan ja vuollejokisimpukan DNA:ta ei havaittu samoissa näytteissä. Vaikka simpukoiden DNA:ta kulkeutuu virran mukana, ei pikkujärvisimpukan perimäainesta havaittu pisteellä VA7, joka oli 4,5 kilometriä alempana joessa kuin piste VA6. Pisteiden VA5 ja VA6 välinen etäisyys oli vain 155 metriä, joten niissä havaittua litteäjärvisimpukan DNA:n lähteenä voidaan pitää samaa esiintymää. Tuoreiden tutkimusten mukaan DNA-signaalit saattavat hiipua vesistöissä lyhyelläkin, jopa alle yhden kilometrin matkalla tai virtavesissä järven tai patoaltaan vaikutuksesta (Hellström ym. julkaisematon aineisto, Cantera ym. 2021).

Nurmijärven Pikkukoskella havaitut kalalajit ovat tyypillisiä koskialueiden ulkopuolelle. Mikkolan ja Sauran (1994) luettelemista koko Vantaanjoen vesistöalueella tavatuista kalalajeista puuttuivat mm. lohi, harjus,

turpa, kivenuoliainen sekä kolmi- ja kymmenpiikki. Vantaanjoen vesistön kalastotarkkailussa vuonna 2019 (Haikonen ym. 2020) saatiin pistettä VA11 lähinnä olevalla sähkökalastusalueella Myllykoskella seitsemän lajia: kirjolohi, kivisimppu, salakka, särki, taimen, turpa ja törö. Yhtään vesistöalueelle uutta kalalajia ei havaittu eDNA-tutkimuksessa.

Johtopäätökset

- Vuollejokisimpukan levinneisyyden yläraja Vantaanjoen pääuomassa sijaitsee Hyvinkään taajaman kohdalla. Sen alapuolella on olemassa vaara, että jokeen vaikuttavat työt voivat vaarantaa lajin suojelutason.
- eDNA-menetelmä on tehokas tapa selvittää lajien levinneisyyttä virtavesissä.

Lähteet

Beamish, R. J. 1980. Adult biology of the river lamprey (*Lampetra ayresi*) and the Pacific lamprey (*Lampetra tridentata*) from the Pacific coast of Canada. Canadian Journal of Fisheries and Aquatic Sciences, 37(11), 1906-1923.

Bruce, K. Bourlat, S. Blackmann, R., Hellström, M., m. fl. & Deiner, K. 2021. A practical guide to DNA-based methods for biodiversity assessment". PenSoft Publishers Bulgaria 2021. ISBN 978-619-248-052-3 (paperback), ISBN 978-619-248-053-0 (e-book) DOI: 10.3897/ab.e68634 i press.

Brys, R., Haegeman, A., Halfmaerten, D., Neyrinck, S., Staelens, A., Auwerx, J., & Ruttink, T. (2021). Monitoring of spatiotemporal occupancy patterns of fish and amphibian species in a lentic aquatic system using environmental DNA. Molecular ecology, 30(13), 3097-3110.

Cantera, I., Decotte, J. B., Dejean, T., Murienne, J., Vigouroux, R., Valentini, A., & Brosse, S. (2021). Characterizing the spatial signal of environmental DNA in river systems using a community ecology approach. bioRxiv, 2020–10.

Deiner, K., H. M. Bik, E. Mächler, M. Seymour, A. Lacoursière-Roussel, F. Altermatt, S. Creer, I. Bista, D. M. Lodge, N. de Vere, M. E. Pfrender, & L. Bernatchez. 2017. Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. Molecular Ecology 26:5872-5895.

De Cahsan, B., Nagel, R., Schedina, I. M., King, J. J., Bianco, P. G., Tiedemann, R., & Ketmaier, V. (2020). Phylogeography of the European brook lamprey (*Lampetra planeri*) and the European river lamprey (*Lampetra fluviatilis*) species pair based on mitochondrial data. Journal of fish biology, 96(4), 905-912.

Douda, K. 2010: Effects of nitrate nitrogen pollution on Central European unionid bivalves revealed by distributional data and acute toxicity testing. Aquatic Conserv: Mar. Freshw. Ecosyst. 20, 189–197. <https://doi.org/10.1002/aqc.1076>

Haikonen, A., Happo, L. & Hynninen, M. 2020: Vantaanjoen vesistön kalastotarkkailu 2019. Kala- ja vesijulkaisuja 284. Kala- ja vesitutkimus Oy.

Kivimäki, A-L., Clergeaud, J. & Tolanen, O. 2021: Maanläjitysalueiden vaikutukset vesiin sekä haittoja ehkäisevät kalataloudelliset kunnostukset – Loppuraportti. Vantaanjoen ja Helsingin seudun vesiensuojeluyhdistys ry. Julkaisu 88/2021. 31.8.2021

Leese, F., Altermatt, F., Hellström M. + 103 authors & Stenke, D (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes* (Rio), 2, e11321.

Leinikki, J. 2018. Suursimpukkapopulaatioiden tila Kokemäenjoessa Harjavallan padon alapuolisella osuudella kesällä 2017. Alleco Oy raportti n:o 5/2018. Alleco Oy 14.6.2018.

Leinikki, J. & Leppänen, J. 2014: Kokemäenjoen simpukkaselvitykset elokuussa 2014 – Loppuraportti. Alleco Oy raportti n:o 6/2014. Alleco Oy 27.11.2014.

Leppänen, J., Leinikki, J. & Väisänen, A. 2020: NiSO₄ spill inflicts varying mortality between four freshwater mussel species (including protected *Unio crassus* Philipsson, 1788) in a western Finnish river. *Environmental Pollution* 256 (2020) 113402

Li, J., Lawson Handley, L. J., Harper, L. R., Brys, R., Watson, H. V., Di Muri, C., & Hänfling, B. (2019). Limited dispersion and quick degradation of environmental DNA in fishponds inferred by metabarcoding. *Environmental DNA* 1 (3): 238–250.

Mikkola, J. ja Saura, A. 1994: Viemäristä lohijoeksi – Vantaanjoen vaelluskalatutkimuksia vuosilta 1987–1993. Riista- ja kalatalouden tutkimuslaitos raportti numero 84/1994.

Nordsieck, R. 2000: The living world of molluscs. Haettu 17. 11 2015 osoitteesta <http://www.molluscs.at/bivalvia/index.html?/bivalvia/naiads.html>

Olds, B. P., Jerde, C. L., Renshaw, M. A., Li, Y., Evans, N. T., Turner, C. R., ... Shirey, P. D. (2016). Estimating species richness using environmental DNA. *Ecology and Evolution*, 6, 4214–4226.

Pedersen, M. W., S. Overballe-Petersen, L. Ermini, C. D. Sarkissian, J. Haile, M. Hellstrom, J. Spens, m. fl. 2015. Ancient and modern environmental DNA. *Philosophical Transactions of the Royal Society B* 370:20130383

Spens, J., A. R. Evans, ... M. Hellström. 2017. Comparison of capture and storage methods for aqueous microbial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635-645.

Syväranta, J. & Leinikki, J. 2018. Suursimpukkaselvitys Lepsämänjoella 2018. Alleco Oy raportti n:o 4/2018. Alleco Oy 15.6.2018.

Taberlet, P., Coissac, E., Hajibabaei, M., & Rieseberg, L. H. 2012. Environmental DNA. *Molecular Ecology* 21, 1789-1793

Uudenmaan ELY-keskus. 2010: Tulkintaohje vuollejokisimpukan lisääntymis- ja levähdyspaikan määrittämiseksi ja turvaamiseksi vesistökuunnostuksissa. Uudenmaan ELY-keskus.

Valovirta, I. 2008: Vantaanjoen Natura-alueen vuollejokisimpukkainventointi 2004–2007. Luonnontieteellinen keskusmuseo, Eläimuseo ja Maailman luonnosäätiö (Suomen WWF). Helsinki 2008

Ympäristöhallinto 2021: Tietolomake

<http://paikkatieto.ymparisto.fi/natura/2018/tietolomakkeet/FI0100104.pdf>; Ladattu 22.9.2021

LIITE 1: Single and multiple species analyses

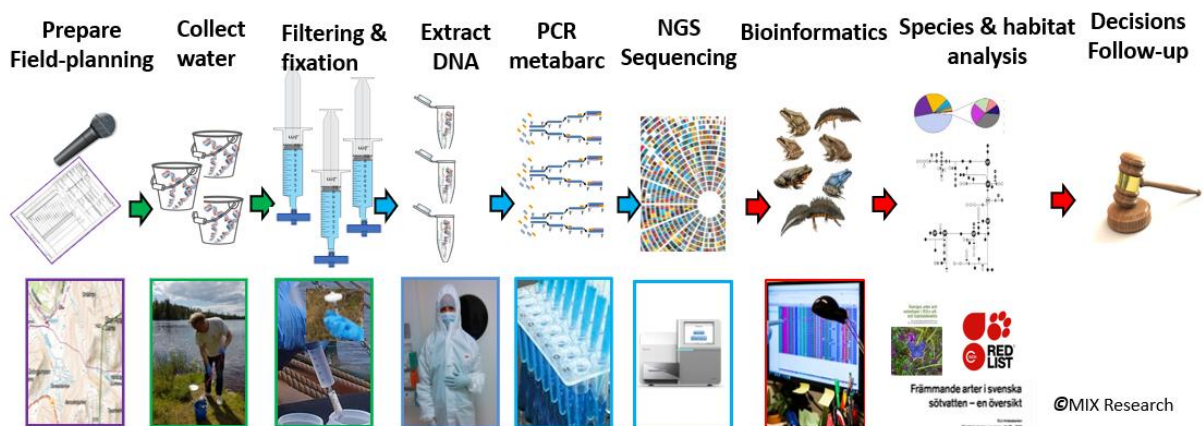
Single species assay (barcoding)- qPCR or ddPCR

Species detections when aiming for one single species is performed with qPCR or dd PCR. The question to be asked is: Can species X be found here? Each test marker is specific for a certain species of interest. Sometimes a species group can be of interest. The results indicate presence or absence and can sometime, if the assay aims for the purpose, give an indication of biomass.

Multispecies assay – Metabarcoding using NGS (Next Generation Sequencing)

The question for multispecies analyses or metabarcoding is: How many species are present here, which ones are the most dominant species? You do not need any previous knowledge of species composition at a given site. The number of species investigated within a taxon group (for example fishes, or amphibians or mussels) is unlimited. The analysis time is longer than that for single species studies, but the data generated the so called “Big Data” and the possibilities for large scale surveys on a much larger magnitude has not been possible before eDNA analyses as an inventory tool. The flowchart in figure B1_1 shows the different steps involved in an eDNA survey. Figure B1_2 summarizes the different types of data that are generated in single and multiple species survey analyses.

Figure B1-1. Flowchart showing the different steps for multispecies analyses from field planning to decisions and follow up based on eDNA data.



Single species barcoding	CGCCGCGGTTATACGAGA	Yes/No
		Species list Dominance
	CGCCGCGGTTATACGAGA	OTU 1 Match → 10 %
	CACCGCGGTTATACGAGA	OTU 2 Match → 65 %
	CGCCGCGGTTACACCACT	OTU 3 Match → 5 %
	CGCCGCGGCTACACCGTG	OTU 4 Match → 20 %

©MIX Research

Figure B1-2. Type of data which is given via single and multiple species analysis. Single species analysis give presence/absence data. Multispecies analysis results in a species list and give an indication of species dominance within a sample.

LIITE 2. Laboratory methods – Metabarcoding

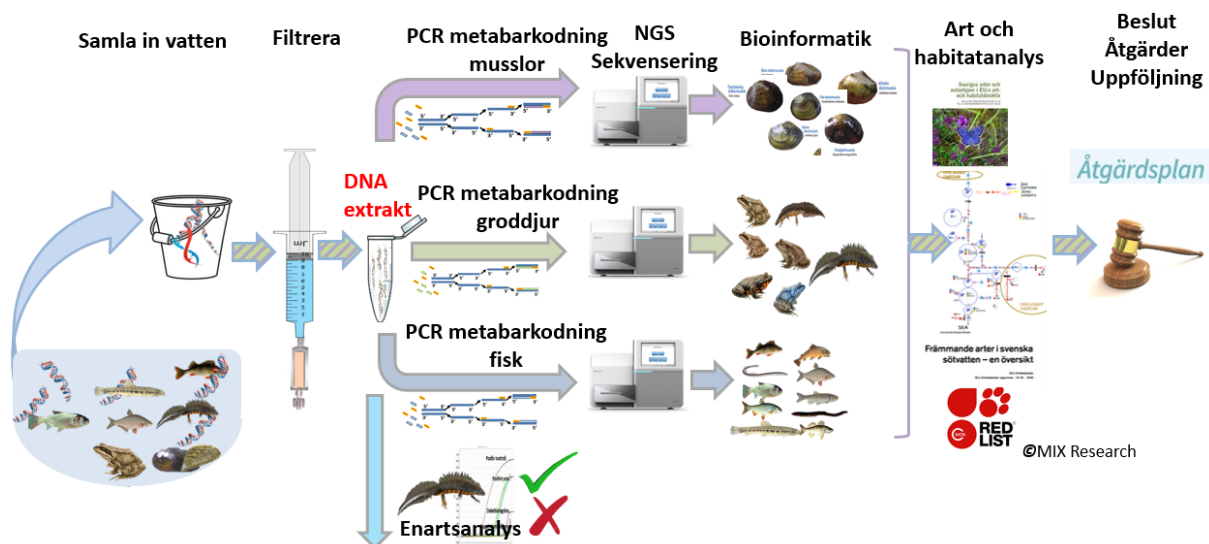
B.2.1 Extraction

The eDNAs were extracted from the filters in laboratories designated for eDNA use only. The extraction protocol followed Spens et al (2017) following the EtOH protocol. (Spens et al. 2017).

B.2.2. PCR metabarcoding

Each marker (in this case unionid mussels and for one sample unionid mussel and fish) the PCR were run in 11 replicates. The PCR aliquots were pooled for each sample during sequencing. Mock libraries with DNA from mussels or fish from tropical areas with known species composition were used as positive laboratory controls- Negative controls were added to each step of the analyses to ensure that the samples were clean from cross-contamination. Analyses for different taxonomic groups can be performed from the very same eDNA sample and different taxa are analysed separately (Figure B2_1). The samples are sequenced separately due to different lengths of the sequences. Sequencing protocols and bioinformatics are outlined in Kačergytė et al. (2021).

Figure B2-1. The very same eDNA extract can be analysed in parallel for different taxa in separate analyses. .



B.2.2.1 Freshwater mussels (Unionidae)

Freshwater mussels within family *Unionidae* were analysed with markers covering 325 base pairs on the 16S genetic region based on the markers described in Prié *et al.* (2021).

B.2.2.2 Fish

Metabarcoding for the fish samples were analysed with a marker targeting a 175 base pairs hypervariable region on the 12S mtRNA gene following Miya, et al. (2015) with modifications described in Miya et al. (2020). Additionally, an overhang was added to the 5' part of the primer to match Illumina Nextera Index markers (See principle in Kačergytė et al. 2021 and NGI - National Genomics Infrastructure, Illumina 16S).

B.2.3. Bioinformatics and species identification

Almost every species has a unique genetic identity (sometimes many unique identities) or genetic barcode in hypervariable gene regions, which gives a molecular identity. The resolutions of these identities are improving with the rapid advancement of molecular techniques in combination with growing reference databases. The different sequences were initially matched against a large and publicly available database based on GenBank and maintained by National Centre for Biotechnology Information, NCBI (<https://www.ncbi.nlm.nih.gov/>). Sequences for more than 450 000 known species are available (Sayers et al. 2020) displaying 1 billion sequences and 6,25 trillion base pairs according to NCBI's homepage. The different sequences are matched against this database and reveal the identity of fish, mussels, amphibians, or mammals. The Sequences on NCBI are not always curated. Therefore, after filtering, taxa were identified by comparing unique sequences against a curated inputs on NCBI and other publicly available reference databases. The species-level identification requires 100% identity. Where multiple reference sequences match equally to the query sequence then all those references are reported in the table. Number of sequence-reads for a particular species compared to number of reads for other species in the same sample provides an idea of species dominance in relation to other species.

References

- Hänfling, B., Handley, L. L., Read, D. S., Hahn, C., Li, J., Nichols, P., ... Winfield, I. J. (2016). Environmental DNA metabarcoding of lake fish communities reflects long-term data from established survey methods. *Molecular Ecology*, 25(13), 3101–3119.
- Kačergytė, I., Petersson, E., Arlt, D., Hellström, M., Knape, J., Spens, J., Žmihorski, M. & Pärt, T. (2021). Environmental DNA metabarcoding elucidates patterns of fish colonisation and co-occurrences with amphibians in temperate wetlands created for biodiversity. *Freshwater Biol.* 66, 1915–1929.
- Miya, M., Sato, Y., Fukunaga, T., Sado, T., Poulsen, J. Y., Sato, K., ... & Kondoh, M. (2015). MiFish, a set of universal PCR primers for metabarcoding environmental DNA from fishes: detection of more than 230 subtropical marine species. *Royal Society open science*, 2(7), 150088.
- Miya, M., Gotoh, R.O. & Sado, T. MiFish metabarcoding: a high-throughput approach for simultaneous detection of multiple fish species from environmental DNA and other samples. *Fish Sci* 86, 939–970 (2020).
- NCBI websida <https://www.ncbi.nlm.nih.gov/>
- Prié, V., Valentini, A., Lopes-Lima, M., Froufe, E., Rocle, M., Poulet, N., ... & Dejean, T. (2021). Environmental DNA metabarcoding for freshwater bivalves biodiversity assessment: methods and results for the Western Palearctic (European sub-region). *Hydrobiologia*, 848(12), 2931-2950.
- Sayers, E.W., ..., Karsch-Mizrachi, I. (2020). GenBank. *Nucleic Acids Research*, 48: D84-D86.
- Spens, J., A. R. Evans, ... M. Hellström. 2017. Comparison of capture and storage methods for aqueous microbial eDNA using an optimized extraction protocol: advantage of enclosed filter. *Methods in Ecology and Evolution*, 8(5), 635-645.
- National Genomics Infrastructure, Illumina 16S. <https://ngisweden.scilifelab.se/methods/illumina-16s-sequencing/>



LIITE 3: Required Quality controls- from collection to metabarcoding

Please record:

1. Volume of water collected/filtered
2. Total eDNA concentration for samples (including negative controls). *The concentrations vary largely between samples but give a first indication on quality of extraction.*
3. Inhibition (Yes/No). Inhibition means that there is a risk that species in the samples will not be detected because of environmental factors such as humic acids interfere with the PCR performance which increases the risk of not detecting species available. Inhibition can be removed by anti-inhibition protocols. *Results of inhibition control needs to be recorded.*
4. Band on gel confirming success of target species PCR needs to be documented (presence/absence, possibly weak bands) for all samples, including negative controls. *This shows that the PCR assay has worked properly.*
5. Negative controls divided into a) field-negatives (equipment negatives) b) extraction negatives c) PCR negatives. *This step is important to make sure that samples are not contaminated and that false positives of target species can be detected. Additionally, the controls makes the source of contamination easily detectable.*
6. DNA sample from a so called “mock community” comprised of species that are not in the area. *This sample is used as a PCR and sequencing positive control to ensure that the system works.*
7. Total reads and % of target species in the reads. Some markers give only 1% target species and very few reads. AquaBiota aims for targets over 50% with the primers used in this study. *This gives an indication on how the target species sequencing works and how suitable the markers are for the target species.*
8. Background contamination (%) from human, cow, and pig and bacteria. These sequences show some background in the reagents and in many cases actual presence of the contaminant species in the water. *Indicates how well the sequencing covers the target species.*
9. A minimum of 9 PCR replicates per sample. *Fewer replicates increase the risk of false negatives as the risk of missing out on rare species increases.*
10. Number of samples in a MiSeq run does not exceed 100, exclusive of sequencing controls. *This ensures that the number of reads in a sequencing run is optimized in order to detect elusive or very rare species. Regardless of all precautions it is important to note that many species are detected but a certain percentage can go missing.*

References:

- Goldberg, Caren S., *et al.* 2016. Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, 7.11: 1299-1307.
- Griffiths Anthony *et al.* 2016. *An Introduction to Genetic Analysis*. 11th edition. WH Freeman. New York. ISBN-13: 978-1464109485.
- Hedman, Johannes *et al.* 2017. Pre-PCR processing-projektet, P4 Stärkt beredskapskapacitet via rationell laboratoriediagnostik samt förenklad provberedning. -*Nationellt Forensiskt Centrum, NFC 2017-05-07. NFC Rapport Avdelningskansliet 2017:04*
- SFMG, Svensk Förening för Medicinsk Genetik. 2011. Riktlinjer för kvalitetssäkring i klinisk genetisk verksamhet. http://sfmg.se/download/riktlinjer/Kvalitetsriktlinjer/sfmg_riktlinjer-for-kvalitetssakring_rev101228.pdf



LIITE 4: Quality control results

The data for the samples and controls are given in table B4_1. All negative controls were negative for target species. eDNA concentrations were as expected and read numbers were high. No inhibition was detected, and the samples were sequences and PCR analysed without cleaning.

Fish MiSeq pairwise sequencing for Miya 12S worked as expected and gave 68 287 reads which is equivalent to **20 million** base pairs in the one sample analysed for fish. Read length covered 300-320 bp.

MiSeq Unio 16S parvis sequencing resulted in 753 794 sequences with an average of 53 840 reads per sample and **226 million** read base pairs. Read length covered ca 300 bp.

Table B_4.1. Quality control for the eDNA samples. Sample ID (see table 1 in the main text). eDNA concentration was measured using *Qubit Fluorometric Quantitation (Fisher Scientific)* given in ng/μl. V annotates volume of water filtered given in millilitres. Inhibition control was tested with qPCR where No means no inhibition and yes inhibition. All PCR samples were performed in 12 replicates and positive PCRs denote number of successful replicates, raw sequence is given before bioinformatics filtering and final sequences number of sequences passing the quality controls.

Sample_ ID	V	DNA c (ng/μL)	Inhibition	Marker	Positive PCRs	Raw sequence #	Final # sequences
VA_01	2700	26.0	No	Unionidae	12	64053	50992
VA_02	2700	4.2	No	Unionidae	12	61970	57586
VA_03	3000	23.4	No	Unionidae	12	63044	8728
VA_04	3300	16.8	No	Unionidae	12	88305	80336
VA_05	3000	14.6	No	Unionidae	11	70829	45913
VA_06	3000	10.8	No	Unionidae	12	72471	43529
VA_07	2700	1.19	No	Unionidae	12	79897	60670
VA_08	3000	31.4	No	Unionidae	12	107209	87062
VA_09	2000	50.4	No	Unionidae	11	80186	57481
VA_10	2000	37.6	No	Unionidae	12	56302	41633
VA_11	2500	79.4	No	Mussel/Fish	12	80629/NA	42673/68287
VA_12	2700	1.6	No	Unionidae	12	87864	49670
VA_13	2700	2.38	No	Unionidae	12	87994	66129
VA_14	2700	6.0	No	Unionidae	12	96410	61392
LabNeg	NA	< 0.01	Yes	Unionidae	0	28493	0

LIITE 5: Distribution of sequence reads for the different species in each sample

Tabell B_5.1 Mussel marker (*Unionidae*). Number of sequence-reads per species within the different sites. The numbers annotate how many times a specific sequence is read on species level and gives an estimate of species dominance or relative biomass. A high number of reads mean high biomass. Total reads per sample is given.

The sequence reads for *Unio crassus* are coloured in purple for presence detected above previous species detection and in bold black for reads detected within the known distribution area.

Species latin	VA_01	VA_02	VA_03	VA_04	VA_05	VA_06	VA_07	VA_08	VA_09	VA_010	VA_011	VA_12	VA_13	VA_14	Lab Neg
<i>Anodonta anatina</i>	47 952	57 586	8 728	1 743	815	2 029		865	2116	765	563	4 552		24 143	
<i>Pseudanodonta complanata</i>	990			337	7 577	5 178									
<i>Unio crassus</i>								964	3 536	2 408	39 813	36 506	52 398	30 013	
<i>Unio pictorum</i>				77 519	21 944	15 452	27 314	55 541	40 954	28 463	1 939	6 875	10 992	5 330	
<i>Unio tumidus</i>	2 050			737	15 577	20 870	33 356	29 692	10 875	9 997	358	1 737	2 739	1 906	
	48 942	57 586	8 728	79 599	30 336	22 659	27 314	57 370	46 606	31 636	42 315	47 933	63 390	59 486	

Tabell B_5.2 Fish marker *Miya* (*Unionidae*). Number of sequence-reads per for the one sample analyzed at site VA_11. The numbers annotate how many times a specific sequence is read on species level and gives an estimate of species dominance or relative biomass. A high number of reads mean high biomass. Total reads were 68 287.

Species latin	# reads
<i>Rutilus rutilus</i>	29547
<i>Perca fluviatilis</i>	10204
<i>Alburnus alburnus</i>	8392
<i>Gymnocephalus cernua</i>	7145
<i>Cottus gobio</i>	3360
<i>Abramis brama</i>	2608
<i>Gobio gobio</i>	1758
<i>Esox lucius</i>	1757
<i>Leuciscus leuciscus/idus</i>	1624
<i>Salmo trutta</i>	1043
<i>Vimba vimba/Blicca bjoerkna</i>	459
<i>Tinca tinca</i>	102
<i>Anguilla anguilla</i>	84
<i>Aspius aspius</i>	70
<i>Scardinius erythrophthalmus</i>	62
<i>Lampetra planeri/fluviatilis</i>	40
<i>Lota lota</i>	32