UNIVERSITY
OF OSLO

Margret Veltman

Hitting a moving target

Novel genomic resources to identify and trace traded orchids

Thesis submitted for the degree of Philosophiae Doctor

Natural History Museum
Faculty of Mathematics and Natural Sciences

University of Oslo

2023

Series of dissertations submitted to the
Faculty of Mathematics and Natural Sciences, University of Oslo
No. 2647

ISSN 1501-7710

Botanical illustrations adapted from:

Jan Moninckx, Maria Moninckx, and Alida Withoos (1699-1706). Moninckx atlas: a catalog in the form of a painted herbarium of the plants in the Hortus Medicus of the City of Amsterdam (Hortus Medicus Amstelodamensis). Allard Pierson Museum - The Collections of the University of Amsterdam. Amsterdam, the Netherlands. https://hdl.handle.net/11245/3.27224

Cover: UiO.
Print production: Graphic center, University of Oslo.

Preface

Doing doctoral research as part of a training network called 'Plant.ID', it is perhaps not surprising that the research presented here is about plant identification: the identification of plants with relevance for society, on which we depend as humans and that we would like to protect for future use. In my case, specifically, orchids. A broad group of species from which I selected a narrow subset (namely edible orchids from the Mediterranean), to focus my efforts on. The working title of this PhD was therefore 'orchid targets', which sounded relatively straightforward to me as it implied a well-defined goal that could be met through a well-planned, if difficult process. Not realising exactly how difficult this process would be at the start, I did know that targets could mean two things: the species that are targets for human consumption and trade, and the genetic targets that can help us say something about these species identities.

While finalising this research - which was far from straightforward - it dawned on me that the targets I had been working on for so long, appeared to be constantly in flux; and that what I was trying to 'hit' was actually a moving target, in multiple ways. Within one year after starting this project, a substantial number of species that were on my 'target list' had already changed name or were no longer accepted, while others were introduced; a sign that orchid research and taxonomy is very much alive and that new insights about their species relationships are being generated every day. While keeping track of the nomenclature was a frustrating and time consuming part of my PhD, it also generated an appreciation for the tremendous amount of work that goes into describing and delineating recalcitrant species: those that resist classification.

Name changes notwithstanding, I was committed to dealing with these recalcitrant targets and characterising their trade patterns. Plants naturally have a propensity to move, with or without the assistance of humans. In addition to trade and migration, which introduces plants and their uses to new territories, climate change nowadays affects the distributions of orchids and their availability for human use; as do a whole range of other environmental problems that impact the habitat and living conditions of orchids in one way or another. But it is not just the plants that are (for better or worse), moving or being moved around: one of the main conclusions of this dissertation is that the targets themselves are shifting, as changing conditions lead people to search for and use new species.

The first question I therefore asked myself was: how do you identify and trace something that is traded, sometimes across large distances, without knowledge of its provenance? This question gradually transformed into: how do you develop a resource that can accommodate changes in the species that are targeted, and that is adaptable to a changing understanding of the boundaries between them and what they should be called? The genetic targets that I defined to achieve

these goals build on and complement existing resources, and together form a toolkit that will hopefully continue to evolve as well - just like orchids do - to meet the changing needs of plant identification.

Acknowledgements

This PhD itself has been a moving target of some sort, with constantly changing outlines (and deadlines), that would never have been brought to completion without the help of a substantial number of people.

First and foremost, I would like to thank my supervisors, **Hugo de Boer**, **Barbara Gravendeel**, **Mike Fay** and **Alex Antonelli**, for their guidance and support. Thank you, **Hugo** for always encouraging me to pursue my own interests and enabling my (sometimes off-topic) conference and workshop visits. Your strategic and diplomatic skills helped this research stay on track, and I appreciated your patience and availability during times of crisis, whether real or imagined. **Barbara**, thank you for always being the critical voice in a discussion and for your infectious enthusiasm about everything orchid-related. You have connected me to many useful people, ideas and resources, for which I am grateful! Thanks **Mike** and **Alex** for providing advice and encouragement from a distance, and being there when I needed it. I hope to make up for my 'virtual secondment' by visiting Kew in person some time.

I would also like to thank my PhD committee at the Natural History Museum. **Lutz**, **Charlotte**, **Jørn**, and **Geir**, you helped me to clarify my priorities and straighten out my timeline when I needed it most. Thanks to your impartial assistance and recommendations, I was able to finish my PhD with the resources I needed and without compromising on quality.

While I was sitting in front of a screen all day, there were other people sitting at the bench doing the essential lab work, without which this thesis would not have had any data to build on. **Bastien** and **Audun**, I don't know how many hundreds of samples you processed (it must have been nearly a thousand?) and how many gels, Qubits and Nanodrops you ran. Your diligence and accuracy in the lab are something I could never strive to achieve myself. Thanks **Bastien**, for coming to visit Oslo and being such a fun and reliable collaborator throughout our Plant.ID trials and tribulations. Thanks, **Audun** for going the extra mile when time was running out and I really, really needed that data. I am eternally grateful to both of you for all your hard work!

There are numerous others who have contributed to this research with sampling and knowledge on orchids and their uses. I would specifically like to thank fellow and former PhD students working on salep, **Susanne Masters** and **Abdolbaset Ghorbani**, who have helped with data collection, interpretation and other resources that were necessary to complete this research. Thanks also to my other collaborators who provided essential expertise that I didn't have. Thanks **Jesus** (aka Adrian) for your knowledge on species distribution modelling and spatial phylogenetics, it gave a new dimension to the project. And a big thank you to my former students, **Théo** and **Audun**, with whom I was able to

foray into non-genetic territory and explore the various trade and policy aspects of this research.

A PhD is more than the research that was done. Like it takes a village to raise a child, it takes a research environment to raise a scientist. I would like to thank my former supervisors at Wageningen University, **Tinde** and **Eric**, for planting the seed for my later research interests and inspiring me to pursue a PhD. Many thanks to the **EDGE group** where I did my PhD, for providing such an open and laid-back space to study everything and anything related (or unrelated!) to plants. I have enjoyed learning about all your various research interests, be they ethnobotany, speciation, plant-insect interactions, fish eDNA or protists. I don't think anyone could've hoped for a more eclectic group of people to derive inspiration from and exchange thoughts with. Thanks **Rebecca**, **Lise**, **María** and **Feli** for sharing an office during the various stages of my PhD. Doing a PhD can be lonely, but your presence made it less so. **Quentin**: thank you for always being so cheerful and optimistic. You're a guiding light for those of use who hope to finish our PhD some day and continue doing research! **Sondre**, **Rita** and all the other master students that have found their way to the EDGE group: thanks for making the museum a more fun and vibrant place to work. **Megan**: you joined the EDGE group with a drive and curiosity that reminds me why I got into science. Thanks for sharing your enthusiasm! Finally, thanks to **Micah** and former EDGE group leaders, **Anneleen** and **Hugo**, and everyone who performed duty work for the EDGE group, for facilitating all the Christmas lunches, ShaSup trips and writing retreats. They provided some much-needed diversions and changes of scenery throughout the years.

While largely virtual, the **Plant.ID** network has been a stimulating environment to do my research in, providing many opportunities for networking and training. I would like to thank all my fellow **ESRs** for sharing in the fun (and sometimes not so fun) times, and **Marcella** and **Brecht**, for making everything run smoothly. I would also like to extend a special thanks to all **European** and **Norwegian tax payers**, without whom there would be no funding to do science in the first place, and no one to do it for. I hope you will continue to value the outputs of science, and enjoy reading or hearing about it at events such as 'Vin og Vitenskap', with your botanical beverage of choice. And thank you, **Norway**, for participating in European funding schemes such as **Horizon 2020** and being such a great country to do a PhD in. I consider it a privilege to have enjoyed your (generous) research funding, social security system, language, culture, ridiculous beer prices and extraordinary beauty.

A PhD would not be complete without having fellow PhDs to complain to and share in the misery. Thank you to everyone at the museum and beyond who has made the journey more bearable and enjoyable. **Helene** (who should get an honorary PhD in emotional support) and **Jesus**, you guys make Oslo feel like home. I hope we can continue to visit each other despite the long distance. **Lasse**: the pandemic unfortunately interrupted our festival and concert visits, but I am hopeful we will find an excuse for you to visit Amsterdam soon! **Annie**: thank you for your relentless efforts to improve things for your co-workers. I am happy the museum has you to keep everyone on their toes! **Solveig**, **Marianne**

and **Lise**: thanks for being so sweet and caring, and occasionally helping us internationals to navigate the intricacies of Norwegian working life. **Siri** and **Nils**: you have transcended the PhD stage already (and how!), but I am so happy we overlapped and stayed friends. I hope there will be more adventuring after the Subaru!

When the pandemic was at its peak and my research at its nadir, I found a new calling in the form of PhD and postdoc rights advocacy. Thank you **UiODoc** for being there to support temporary researchers at UiO, and thank you **Michael**, **Olga**, **Christina**, **Kelly**, **Lasse**, **Erling**, **Vipin** and **Jesus** for being such a great team to work with. Thank you **SiN** for doing the same at the national level, and for giving me the chance to learn everything I ever wanted to know about Norwegian higher education, and more. **Camilla** and **Ingvild**: thank you for your hard work and support during our year on the board. I have learned so much from both of you and am confident that the spirit of Nordic democratic standards and procedures will never quite leave me!

Despite the pandemic, my PhD has carried me to other places outside Norway. I would like to acknowledge the doctoral school in the humanities at **Stockholm University** for allowing me to participate in a series of workshops in the Environmental Humanities, which has stimulated my cross-disciplinary thinking and given me new perspectives on human-plant interactions. I would also like to thank **Barbara** for inviting me to work at Naturalis Biodiversity Center, the **Evolutionary Ecology research group** for hosting me and making me feel welcome and part of their team, and the **PhD and postdoc community at Naturalis** for providing such a fun and dynamic research environment while I was there. Naturalis has been a second academic home to me during the last year and a half, for which I am very grateful.

Finally, thank you to my Dutch friends and family who came to visit during my time in Norway. **Jill** en **Jorrit**: ik ben blij dat jullie net zo enthousiast zijn over Noorwegen als ik, en hoop dat we nog vaak samen mogen wandelen en kamperen! **Fenna**: het is altijd lachen om met jou spontaan dingen te doen, in Zweden, Noorwegen of Nederland. **Eliza**: volgens mij is er geen plek op aarde waar je niet heen zou gaan. Wat heb ik geluk met zo'n reislustig vriendinnetje! **Anneke**, **Memo**, **Anna** en **Anno**: ik hoop op een spoedige reünie, en het bier zal hopelijk een stuk goedkoper zijn als jullie me elders komen opzoeken! **Papa** en **mama**: bedankt dat jullie me altijd hebben gesteund als ik weer eens wilde vertrekken. Maar vooral bedankt dat ik steeds mijn eigen interesses heb kunnen volgen en dat jullie de studies van mij, Anneke en Anno mogelijk hebben gemaakt. Dit is hopelijk de laatste diploma-uitreiking van ons waar jullie bij hoeven te zijn!

And to my favourite biologist of all time, **Peter**: thank you for being my long-distance partner, lockdown buddy, and cohabitant all at once. You inspire me to be a better scientist and person, and make me see new possibilities that I didn't see before. I hope it only gets better from here!

Margret Veltman
Amsterdam, June 2023

Contents

Preface **i**

Contents **v**

Sammendrag **vii**

Summary **xi**

List of Papers **xv**

Introduction **1**
Background 1
Approach 8
Synthesis 16
Concluding remarks 28
References 32

Papers **40**

I Wildlife trade **41**

II Orchidinae-205: a new genome-wide custom bait set for studying the evolution, systematics, and trade of terrestrial orchids **59**

III Geographic shifts and increasing species diversity of wild orchid harvesting threatens survival of natural populations **87**

Appendices **115**

A Supplementary information for Paper II **117**

B Supplementary information for Paper III **159**

Sammendrag

Ulovlig handel med dyr og planter er et av de mest presserende problemene som truer biodiversitet i dag. Med over 28 000 arter er orkidéfamilien (Orchidaceae) den største gruppen av plantearter som er oppført av konvensjonen om internasjonal handel med truede arter av vill flora og fauna (forkortet CITES), som strengt kontrollerer deres import og eksport. Likevel er handelen med spiselige knoller fra terrestriske orkidéer, som brukes i tørket og pulverisert form til en populær middelhavsdelikatesse kalt "salep", svært utbredt til tross for reguleringer. Denne avhandlingen undersøker tilgjengeligheten av molekylære verktøy for overvåking av vill plantehandel og bygger på den økende betydningen av målrettet fangstsekvensering (fra engelsk: «target capture sequencing») for artsidentifikasjon ved å utvikle et tilpasset agnsett (fra engelsk: «bait set»), Orchidinae-205, for beriking av 205 nye markører som er tilpasset Orchidinae s.l.: undergruppen som vanligvis blir målrettet for salep.

Orchidinae s.l. omfatter omtrent 1800 arter globalt, hvorav et underutvalg på omtrent 100 forekommer i den sørøstlige middelhavsregionen. Fordi salepknoller er morfologisk like, er artens identitet og opprinnelse ikke påvisbar uten molekylære metoder, og direkte observasjoner av høsting som kan bekrefte artens identitet når de blomstrer, er begrenset. Kunnskap om de artene og kildepopulasjonene som blir mest utnyttet, kan bidra til bevaringstiltak for å beskytte lokal orkidédiversitet. Men den nåværende etablerte metoden for artsidentifikasjon - tradisjonell strekkodeavlesning med et lite antall universelle plantemarkører - har ofte ikke tilstrekkelig fylogenetisk kraft til å skille mellom nært beslektede orkidéarter, og standard sekvenseringsmetoder for populasjonsgenetikk for å oppdage finmasket artsmangfold og populasjonsstruktur er for det meste uegnet for nedbrutt materiale som kokte og tørkede orkidéknoller.

Målrettet fangstsekvensering er en fremvoksende metode innen systematikk og evolusjon av ikke-modellarter uten referansegenomer, men Orchidinae s.l. er underrepresentert i målsekvensene (fra engelsk: «target sequences») til universelle eller nåværende agnsett spesifikt for orkidéer, noe som begrenser deres sekvensgjenoppretting og robustheten i fylogenetiske analyser. Denne avhandlingen prøver å fylle dette gapet ved å utvikle nye genetiske ressurser og verktøy for målrettet fangstsekvensering av Orchidinae s.l. Agnsettene ble designet med transkriptomsammenstillinger fra 14 arter som representerer åtte slekter innenfor stammen. For å kombinere fylogenetisk kraft med funksjonell relevans av mållokiene, ble 174 enkeltkopierte lokus valgt og supplert med 31 kandidatgener som antas å være involvert i biosyntesen av glucomannan, det vannløselige polysakkaridet som gir saleppulver sin karakteristiske tekstur. Agnene for å berike disse lokusene ble testet på et utvalg av 77 taxa som finnes i regionen som omfatter Hellas, Tyrkia og Iran, som er nærliggende land der salep

ofte høstes og selges.

Sekvensassembleringen viser universelt høy utvinning av eksoner på tvers av fokusartene uavhengig av artstilhørighet, og gjenoppretter over 80% av det totale målsekvensrommet (fra engelsk: «target space») på 306 kb. Sammenlignet med ytelsen til andre agnsett på identiske eller nært beslektede taksa, viser de tilpassede agnene som er designet her, at de kunne gjenopprette lengre sekvenser med mindre manglende data, høyere taksonbesetning og flere variable og parsimonisk informative områder. En grundig fylogenetisk sammenligning med det universelle agnsettet «Angiosperms-353» viser høyere støtte for gentrær og mer fylogenetisk informasjonsverdi for «Orchidinae-205»-lokusene for ni fokusslekter. Mens maksimum likelihood (ML) og multispecies coalescent (MSC) tilnærminger til rekonstruksjon av artstrær for alle prøver viser høy støtte for de fleste kladene, forblir flere arters forhold i nylig divergerte slekter eller artskomplekser omdiskuterte og krever ytterligere sekvensinformasjon for å bli løst. Dette kan oppnås ved å legge til intronsekvenser som også ble assemblert i denne studien og fanget som bifangst under målberikelse (fra engelsk: «target enrichment»). Disse kan gi ekstra sekvensjusteringer som ofte er lengre og inneholder mer fylogenetisk informasjon enn eksonene som ble analysert her, spesielt for mindre taksonomiske grupper av interesse.

Referansedatabasen som beskrives her, kan brukes til å identifisere salep-knoller ved å plassere dem i et felles fylogenetisk rammeverk med passende referanseprøver. Vi brukte agnene på et utvalg av 99 historiske salep-knoller samlet fra museer over hele Europa, datert fra midten av 1800-tallet til slutten av 1900-tallet, samt 97 moderne salep-knoller fra markeder i ulike byer i Tyrkia og Iran. Til tross for høy fragmentering og lav konsentrasjon av DNA hadde 90% av knollene tilstrekkelig utvinning av målsekvenser til å muliggjøre fylogenetisk analyse. Ved å bruke en multikriterie metode for å avgjøre artsbestemmelse basert på genetisk avstand og monofyli i ML- og MSC-trær, ble 80-85% av prøvene identifisert med tillit på artsnivå. Gjenværende bestemmelser på slektsnivå kan potensielt avgrenses til artsnivå ved å analysere intronsekvenser for sett av knoller og arter med motstridende fylogenetiske plasseringer.

Identifikasjonen av salep-knoller avslører endringer i artssammensetningen av salep over tid og sted. Mens noen få *Orchis*-arter pleide å dominere markedet, blir nå mange flere slekter høstet, og det virker som om de er geografisk strukturert. Nylig høstede arter har et bredere spekter av blomstringstider, noe som antyder at høstingen, som er nært assosiert med blomstringssesongen, nå skjer flere ganger i løpet av sesongen. Dette kan forklare delvis økningen i artsmangfoldet. Som en mulig konsekvens av overhøsting, finner vi at knollene også minker i størrelse, noe som kan skyldes høsting av yngre individer enten på grunn av ukritisk høsting eller forsvinningen av eldre individer i ville orkidépopulasjoner.

Analyse av fylogenetisk mangfold viser at diversiteten av solgt salep i Tyrkia, (som anslås å ha en rikere orkidéflora enn Iran), er lavere enn i Iran. Kombinert med den enkeltstående observasjonen av historisk populære arter i nordøst-Iran, der salep ikke tradisjonelt konsumeres, antyder dette at lokal utarming kan presse høstingen av salep-arter østover til områder der de fremdeles er relativt vanlige. Analyse av «hot nodes» i artsfylogenien som er beriket med salep

i ulike tidsperioder, fremhever kladene som ser ut til å ha høyere risiko for høsting nå enn tidligere. Disse kladene inkluderer alle *Anacamptis*- og *Serapias*-arter, som på grunn av kombinasjonen av deres høstingsrisiko og den relativt lave artsoppløsningen, bør være en topprioritet for fremtidig analyse og økt overvåking.

Denne avhandlingen kaster lys over problemet med overutnyttelse av knollformede orkidéer ved å kombinere målrettet fangst med økologiske og morfologiske data for å avdekke skiftende høstingspress, samtidige endringer i tilstanden til populært høstede arter og deres mulige innvirkning på mangfoldet og overlevelsen til lokale populasjoner. Avhandlingen presenterer også en ressurs som kan brukes av orkidéspesialister innen systematikk, bevaring og rettsgenetikk. Fremtidig forskning kan benytte denne ressursen for å bidra til vår forståelse av orkidéers evolusjon og slektskap, for å informere om bedre artsforvaltning og muliggjøre mer effektiv overvåking av orkidéhandelen i fremtiden.

Summary

Wildlife trade is one of the most pressing problems threatening biodiversity today. With over 28,000 species, the orchid family (Orchidaceae) is the largest group of plant species listed by the Convention on International Trade in Endangered Species of Wild Fauna and Flora (CITES), which tightly controls their import and export. Yet, trade in the edible tubers of terrestrial orchids, used in dried and powdered form for a popular Mediterranean delicacy called 'salep', is rampant in spite of regulations. This thesis examines the current availability of molecular tools for monitoring wildlife trade, and builds on the growing importance of target capture sequencing for species identification by developing a custom bait set (Orchidinae-205) for enrichment of 205 novel markers that are tailored to Orchidinae s.l.: the subtribe most commonly targeted for salep.

Orchidinae s.l. comprises about 1800 species globally, of which a subset of about 100 occur in the southeastern Mediterranean region. Because salep tubers are morphologically indistinguishable, their species identity and provenance are undetectable without molecular methods, and direct observations of harvesting that can corroborate species identity when they are in flower are limited. Knowledge of the species and source populations targeted the most can aid conservation actions to protect local orchid diversity. But the currently established method for species identification - traditional barcoding with a small number of universal plant markers - often does not have enough phylogenetic power to discriminate between closely related orchid species, and standard population genetic sequencing techniques for detecting fine-grained species diversity and population structure are mostly unsuitable for degraded material such as boiled and dried orchid tubers.

Target capture is emerging as a leading method in the systematics and evolution of non-model species without reference genomes, but Orchidinae s.l. are underrepresented in the target sequences of universal or orchid-specific bait sets currently available, limiting their sequence recovery and the robustness of phylogenetic inference. This thesis tries to address this gap by developing new genetic resources and tools for target capture sequencing of Orchidinae s.l. Custom baits were designed with transcriptome assemblies of 14 species representing eight genera sampled across the tribe. To combine phylogenetic power with functional relevance of the target loci, 174 single-copy loci were selected and supplemented with 31 candidate genes putatively involved in the biosynthesis of glucomannan, the water-soluble polysaccharide that gives salep powder its distinctive textural properties. Baits for enriching these loci were tested on a selection of 77 taxa occurring in the region comprising Greece, Turkey and Iran, adjacent countries where salep is commonly harvested and sold.

Sequence assembly shows universally high exon recovery across target species,

irrespective of species identity, recovering >80% of the total target space of 306 kb. A comparison with the performance of other bait sets on identical or closely related taxa reveals that the custom baits designed here were able to recover longer sequences with less missing data, higher taxon-occupancy and more variable and parsimony informative sites. An in-depth phylogenetic comparison with the universal baits of Angiosperms-353 shows higher gene tree support and phylogenetic informativeness of the Orchidinae-205 loci for nine target genera. While maximum likelihood (ML) and multispecies coalescent (MSC) approaches to species tree reconstruction of all samples show high support for most clades, several species relationships in recently radiated genera or species complexes remain contentious, and will require additional sequence information to be resolved. This could be achieved by analysing the intron sequences also assembled in this study for smaller taxonomic groups of interest, which are captured as by-catch during target enrichment and can provide additional sequence alignments that are often longer and contain more phylogenetic information than the exons analysed here.

The reference database described here can serve to identify salep tubers by placing them in a joint phylogenetic framework with suitable reference samples. We applied the baits to a selection of 99 historical salep tubers, dating from the mid-19th century to the late 20th century and sampled from museums across Europe, and to 97 modern salep tubers sampled from markets in various cities across Turkey and Iran. Despite high fragmentation and low concentrations of DNA, 90% of tubers had sufficient target recovery to enable phylogenetic analysis. Using a multi-criteria approach for arbitrating between different species assignments based on genetic distance and monophyly in the ML and MSC trees, 80-85% of these were confidently identified at the species level. The remaining genus level identifications could potentially be narrowed down to the species level as well, by analysing intron sequences for subsets of tubers and species with conflicting phylogenetic placements.

The identification of salep tubers reveals changes in the species composition of salep over time and space. Whereas a few *Orchis* species used to dominate the supply, nowadays many more genera are harvested, that appear to be structured geographically. Newly collected species have a wider range in flowering times, suggesting that harvesting, which is closely associated with the flowering season, now occurs at multiple times in the season. This may partially explain the increase in species diversity. As a possible consequence of overharvesting, we find that tubers are also decreasing in size, which could be caused by the harvesting of younger individuals either due to indiscriminate harvesting or the disappearance of the older demographic of wild orchid populations.

Phylogenetic diversity analysis shows that the diversity of sold salep in Turkey (which is estimated to have a richer orchid flora than Iran), is lower than in Iran. Coupled with the exclusive observation of historically popular species in Northeast Iran where salep is not traditionally consumed, this hints that local depletion might be pushing the harvest of salep species eastward to areas where they are still relatively abundant. Analysis of "hot nodes" in the species phylogeny that are enriched for salep in different time periods, highlights

clades that appear to be at higher risk for being harvested now than in the past. These clades include all *Anacamptis* and *Serapias* species, which, due to the combination of their harvesting risk and the relatively low resolution of their species relationships, should be a top priority for future analysis and increased monitoring.

This thesis sheds light on the problem of overexploitation of tuberous orchids by combining target capture with ecological and morphological data to elucidate shifting harvesting pressures, concomitant changes in the state of popularly harvested species, and their possible impact on the diversity and survival of local populations. This thesis also presents a resource that can be used by communities of orchid specialists in systematics, conservation and forensics. Future research can utilise this resource to aid our understanding of orchid evolution and relationships, to inform better species management and enable more effective monitoring of orchid trade in the future.

List of Papers

Paper I

Jahanbanifard, M., Veltman, M. A., Veldman, S., Hartvig, I., Cowell, C., Lens, F., Janssens, S. & Smets, E. (2022). "Wildlife trade". In: H. de Boer, M. O. Rydmark, B. Verstraete, & B. Gravendeel (Eds.), *Molecular identification of plants: from sequence to species*, pp. 372–386. Advanced Books. DOI: 10.3897/ab.e98875.

Paper II

Veltman, M. A., Anthoons, B, Schrøder-Nielsen, A., Gravendeel, B. & de Boer, H. "Orchidinae-205: a new genome-wide custom bait set for studying the evolution, systematics, and trade of terrestrial orchids". *Submitted for publication.*

Paper III

Veltman, M. A., Anthoons, B., Schrøder-Nielsen, A., Chimal Ballesteros, J. A., Ghorbani, A., Karahan, A., Öztürk, E., Terzioglu, S., Akan-Küçükaladağ, S., Masters, S., Nesbitt, M., Fay, M. F., Antonelli, A., Gravendeel, B. & de Boer, H. "Geographic shifts and increasing species diversity of wild orchid harvesting threatens survival of natural populations". *Manuscript.*

Introduction

Background

Wild plant use and trade

Wildlife trade is a billion dollar industry (CITES Secretariat, 2022). The trade of wild animals and plants is of increasing conservation concern, as traded volumes of wild species and their derived products are growing across nearly all sectors, and could lead to overexploitation (Hughes, 2021) While much research has focused on the overexploitation and trade of animal (mostly vertebrate) species (Morton et al., 2021; Scheffers et al., 2019), plants have remained underhighlighted (Margulies et al., 2019). Nonetheless, plants constitute over 80% of all wild species used by humans in one way or another, with the majority used for food and medicine, and a large part of the world population relies on them to meet their basic needs (IPBES, 2022). The increasing demand for authentic and natural products, however, has transformed the market for traditional foods and medicinal plants. This means that for certain very popular wild plant species (e.g. those used in Traditional Chinese Medicine or in Ayurveda), harvesting is no longer a local practice, but a global industry (Booker, 2014). Unfortunately, the supply of wild plants is not easily scalable unless they can be artificially propagated or cultivated. For many species, this is either cumbersome (or impossible); financially unrewarding; or perceived to yield products of inferior quality (Schippmann, Leaman, and Cunningham, 2002). This means that, alongside climate change and land use change, unsustainable wild collecting remains one of the main risks for the survival of wild populations, and should hence be a top priority for the protection of natural plant resources (Maxwell et al., 2016).

To protect wild plants from the devastating effects of overexploitation, the Convention on the International Trade of Endangered Species of Flora and Fauna (CITES) blacklists species that are deemed vulnerable to overexploitation, only allowing international trade in cases where it is not detrimental to the survival of species in the wild and under strict permitting requirements. Despite these legal provisions, the reporting system is deficient both in terms of its completeness and its accuracy in keeping track of which plants are traded, for what purposes and in what quantities (Lavorgna et al., 2018). This opens the door to various forms of non-compliance, such as smuggling and laundering of wild plant material (Hinsley, Nuno, et al., 2017). Enforcement of CITES regulations relies on knowledge of the species being traded, which sometimes requires significant taxonomical expertise. In addition, botanical specimens may be processed before use or sometimes consist only of plant parts with insufficient diagnostic characters for reliable species identification. The lack of morphological evidence and of specialists who

can assess this evidence to confirm species identities and origins, presents a clear need for high-throughput and automated assessment of the species identity of wildlife products that are encountered on the illegal market or confiscated at border control (Bashyal and Roberts, 2023).

For this reason, molecular identification is seen as one of the most promising approaches to wildlife forensic investation (Gouda et al., 2020). Traditionally, this has been done with the use of diagnostic genetic markers called 'barcodes' (Staats, Arulandhu, et al., 2016), but the application of traditional universal barcodes to plant species has been hampered by lack of markers that are both universal as well as informative (Hollingsworth, Graham, and Little, 2011). For many plant species, broader coverage is therefore needed to tell them apart (Hollingsworth, Li, et al., 2016). Fortunately, recent years have seen tremendous progress in the quality and availability of sequencing technologies, with costs dropping below $0.01 per raw megabase and $500 genomes available as of 2021 (Wetterstrand, 2023), enabling the cost-efficient generation of genome-wide data even for non-model organisms without an assembled reference genome (Ellegren, 2014). Methods are now emerging to retrieve DNA from even the most degraded plant remains (Latorre et al., 2020), theoretically providing wildlife forensics with all the tools that it needs for monitoring and enforcement of trade regulation.

Protecting biocultural diversity

While much of the literature on wildlife trade has focused on its monitoring and enforcement, which are both necessary elements of biodiversity protection, I would like to digress a bit and mention that the international response to overexploitation is sometimes also criticised. Some of the main criticisms are its lack of consideration for historical contingencies and biological complexity (Lavorgna et al., 2018) and the inherent conflict it poses with traditonal uses and livelihoods (Cooney et al., 2018). Wild plant uses often go back centuries and can be an important part of the cultural identity of local communities. Cultural and biological diversity are often inextricably linked and the intimate connections between people and plants are not always a danger, but can also provide an incentive to conserve species and habitats. For this reason, conservation efforts are shifting away from human-exclusive to human-inclusive models such as biocultural conservation, which recognises that the progressive loss of cultural diversity is a crisis in itself that leads not just to the erosion of traditional ecological knowledge but also of biodiversity itself (Gavin et al., 2015). This trend mirrors the increasing involvement of local communities in community-based resource management, which values indigenous and local knowledge (ILK) or traditional ecological knowledge (TEK) of their environment as important sources of environmental stewardship (Fariss et al., 2023). These approaches are increasingly recognised and implemented in the arena of international wildlife trade as well (Roe and Booker, 2019).

While the current biodiversity crisis is pressing and demands urgent action, taking a step back to allow for a historical perspective on plant use and trade is necessary if we are to take local and global cultural factors that impact

their conservation seriously. Humans have mediated the dispersal of plants and their uses since prehistory (Boivin, Crassard, and Petraglia, 2017). Whereas long-distance trade is becoming more predominant in the globalised economy of today, trade connections have always been an important catalyst for cultural change and have led to the horizontal diffusion of ethnobotanical knowledge between cultures, in addition to their vertical transmission from one generation to the next (Teixidor-Toneu, Jordan, and Hawkins, 2018). Thus, to understand the geographic spread of plants and their uses today, and come to sensible conclusions about their management, we need to consider not only their biological distributions and diversity, but also the historical events that have shaped contemporary interest in these plants among different groups of people, and the socio-economic forces that drive changes in their patterns of harvest and trade.

Orchids and salep

This dissertation will examine the historical and contemporary trade of wild edible orchids in the southeastern Mediterranean region (Figure 1. The orchid family (Orchidaceae) numbers roughly 28,000 species (Fay 2018) and has a global distribution. Orchids have a variety of uses all over the world, ranging from medicine (Chinsamy, Finnie, and Van Staden, 2011) and food (Veldman et al., 2018) to horticulture (Phelps and Webb, 2015), and are even used as adhesive (Berdan et al., 2009). Most orchid species (about 80%) are epiphytes that are native to the tropics (Givnish et al., 2015), but a smaller number grow in temperate environments and most of these are terrestrial. Some of these have underground tubers that are harvested for human consumption, a use that has been documented on different continents (Bulpitt, 2005). One of the most common examples is a product "salep", a traditional delicacy in Mediterranean countries that can refer both to a drink or to ice cream, that is made with ground orchid tubers (Ece Tamer, Karaman, and Utku Copur, 2006). The orchids that provide the raw material for this product all belong to the subtribe Orchidinae (tribe Orchideae), which belongs to the Orchoideae subfamily and contains about 1800 species worldwide (Jin et al., 2017), with roughly a hundred occurring in the region (Pridgeon et al., 2001; Pridgeon et al., 2003). The tubers used for salep are prized not just for their distinctive flavour, but also for their unique polysaccharide composition, which contains a high concentration of glucomannan that is used as a thickening agent (Kurt, 2021) and is thought to promote satiety and offer other health benefits to its consumers (Ece Tamer, Karaman, and Utku Copur, 2006).

Salep has a long history of use in countries of the former Ottoman empire and is especially popular in the Balkans and Anatolia. The growing popularity of salep has caused concerns about unsustainable harvesting and overexploitation (Ghorbani, Gravendeel, Naghibi, et al., 2014; Kreziou, de Boer, and Gravendeel, 2016). Both tubers and powdered salep are now being traded internationally, and it is estimated that the harvest in Turkey and its neighboring country Iran amounts to millions of plants harvested per year (Ghorbani, Gravendeel, Naghibi, et al., 2014; Kasparek and Grimm, 1999). Reports of population decline and

Figure 1: Some of the orchid species that are the focus of this dissertation. Top: *Orchis mascula* (L.) L., *Neotinea ustulata* (L.) R.M.Bateman, Pridgeon & M.W.Chase, *Ophrys insectifera* L. Middle: *Platanthera bifolia* (L.) Rich., *Dactylorhiza majalis* (Rchb.) P.F.Hunt & Summerh., *Gymnadenia conopsea* (L.) R.Br. Bottom: *Himantoglossum robertianum* (Loisel.) P.Delforge, *Anacamptis pyramidalis* (L.) Rich., *Serapias bergonii* E.G.Camus. Photos: Rogier van Vugt.

indiscriminate harvesting suggest that different species may be targeted now than in the past, putting more species at risk of being threatened. Recent barcoding studies have revealed that as many as 35 species are being harvested in Greece, Turkey and Iran alone (Ghorbani, Gravendeel, Selliah, et al., 2017; Kasparek and Grimm, 1999; Kreziou, de Boer, and Gravendeel, 2016). Recent barcoding studies have tried to shed light on the species composition of salep (de Boer et al., 2017; Ghorbani, Gravendeel, Selliah, et al., 2017), but limited sequencing success and phylogenetic resolution has precluded firm conclusions about which species are harvested more commonly, and lack of historical sampling has prevented analysis of how this compares to past preferences.

Conceptual framework

This thesis will explore some of the historical dynamics of salep trade through the lens of genetic and ecological data, in order to shed light on the changes in, drivers of and consequences of wild orchid harvesting and trade, with a specific focus on contemporary Turkey and Iran. While detailed socio-ecological analyses are beyond the scope of this research, the results of this thesis can be placed in a conceptual framework that inspires thinking about its socio-ecological context.

The problem of overexploitation can be conceptualised through the DPSIR framework for modelling society-environment interactions (Bradley and Yee, 2015). DPSIR stands for Driver, Pressure, State, Impact, and Response, where drivers are external forces (social, demographic and economic) that propel human actions; pressures are the processes that cause negative externalities for the environment as a direct consequence of human actions; state refers to the change in state of the environment as a result of human actions; impact concerns the consequences of this state change for environmental quality, potentially influencing human and non-human well-being; and responses are the options available or the management actions taken to solve the problem (Figure 2). The DPSIR framework is a common tool for understanding causal chains in environmental systems analysis, developed to further the sustainable use of natural resources, and is therefore highly applicable to biodiversity management (Maxim, Spangenberg, and O'Connor, 2009).

Seen in this framework, overharvesting of orchids is a pressure exerted on the environment by human actions, which may lead to a change in the state of wild plant populations (e.g. lower abundance, reduced genetic diversity). This can impact environmental quality in several ways, for example by resulting in the decline and loss of species that provide ecosystem services (such as salep supply), which are subsequently less or no longer available for human enjoyment, and of biodiversity itself. Typical responses to overexploitation are the designation of protected areas, listing plants as endangered or restricting their trade. Unfortunately, if the drivers that promote overharvesting (such as high demand, large profits, and established trade and sales routes) persist and enforcement is inadequate, the problem is unlikely to disappear completely.

While there is a general understanding of the causes for the overexploitation of orchids (drivers, such as profits made with trade), and the measures available

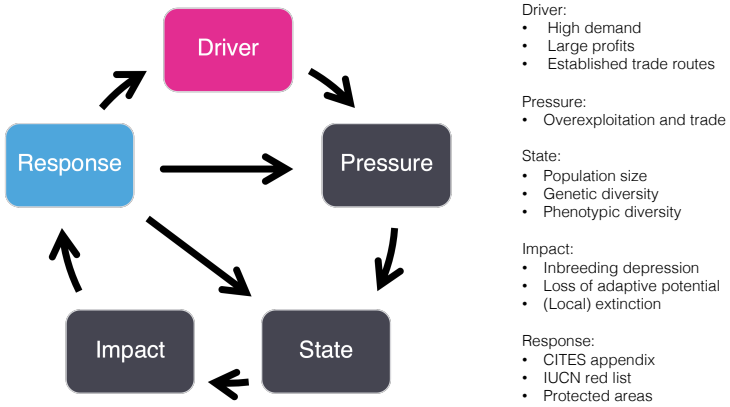

Figure 2: The DPSIR framework conceptualises the causal chain of environmental problems such as overexploitation of natural resources. Some examples of possible drivers, states, impacts and responses as they relate to wild orchid harvesting for salep are listed on the right. The boxes in black highlight parts of the causal chain that are hard to measure, but can be illuminated with molecular methods.

to control it (responses, such as restricting trade), quantitative knowledge of the harvesting pressure on different species, their ecological state and the ensuing conservation impacts is currently limited (Hinsley, de Boer, et al., 2017), and hence constitutes somewhat of a black box. Shedding light on this part of the causal chain therefore requires more advanced tools for tracking species identity, geographic provenance and genetic diversity of traded orchids. This thesis aims to contribute to filling this knowledge gap by evaluating and developing such tools and demonstrating their utility in the empirical case of salep trade in Turkey and Iran.

Research questions and aims

This thesis specifically aims to answer the following research questions (RQ):

1. How is wildlife trade, and in particular orchid trade, currently regulated and monitored?

2. What is the potential of target capture for monitoring the species identity of (severely degraded) illegally traded salep tubers?

3. What is the temporal and spatial variation in traded salep species, and what are its drivers and (potential) consequences?

In order to answer these research questions, the thesis includes three papers that address different parts of the DPSIR framework. Paper I reviews the current state of wild plant trade, its regulation and the challenges associated

with it. It touches on the head and tail of the DPSIR causal chain, namely the driving forces behind wildlife trade and the responses that are in place to control it (RQ1). The main contribution of Paper I is to list the tools available to enforce this response, by monitoring certain (legal) aspects of traded wildlife products that we would otherwise be unable to tell, one of which is species identity (RQ1). Paper II takes the recommendation of Paper I in developing and testing novel genetic markers for the identification of salep species (RQ2). Paper II also provides (the beginning of) a reference database against which traded salep tubers can be compared in order to assess their species identity. The tools developed in this paper theoretically enable monitoring of the pressures that are exerted on the environment (e.g., which species are being harvested and where?) and the resulting changes in environmental state (e.g., how does this influence their relative abundance and genetic diversity?). Paper III implements this new tool in practice in order to elucidate the changing pressures on salep species over time, their implications for the state of wild orchid populations, and ultimate impact on orchid availability and salep supply (RQ3). In doing so, it also provides an evaluation of the effectiveness of the tool developed in Paper II in reaching its primary aim, namely species identification (RQ2). Paper III further examines connections between salep species composition and several ecological and morphological variables, such as elevation and flowering time and tuber size and weight, in order to understand different harvesting practices and their possible consequences for orchid populations. Both Paper II and III touch on the utility of the developed genomic resources for enabling more effective responses to (illegal) harvesting and wildlife trade (RQ1), referring back to the recommendations made in Paper I. An overview of the scope of the different papers and how they relate to the research questions is given in Figure 3.

The specific objectives and hypotheses for each paper are as follows:

- Paper I aims to find the most suitable methods for detecting various types of illegal trade, by conducting a thorough literature review. We hypothesise that target capture is the most promising method for species identification of severely degraded plant materials without sufficient morphological characteristics. These findings lead to the choice of method developed in Paper II.

- Paper II aims to develop and test a custom bait set for targeted capture of novel markers that are suited for species identification of tuberous orchids (Orchidinae s.l.). We hypothesise that the custom baits designed here outperform universal loci and off the shelf kits in terms of target recovery and phylogenetic support and informativeness.

- Paper III aims to apply the method developed in Paper III to an empirical case, namely salep harvest and trade, in order to identify the species that are most at risk. We hypothesise that salep harvest is expanding taxonomically and geographically, and that this is driven by population declines (Figure 4).

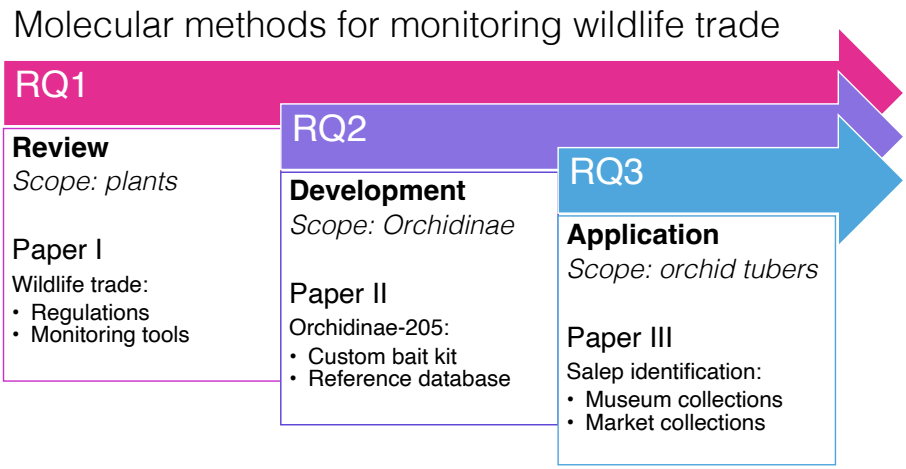

Figure 3: Graphical overview depicting the relationships among the three papers and their connection to the three research questions. Each paper focuses on methodological advancements in wildlife trade monitoring, but centers on a different aspect (review of existing methods, development of a new method and application of the method in practice). These aspects follow each other in chronological order, but findings from Paper II and Paper III hark back to earlier questions, yielding progressive insights.

Approach

We employed a five-step approach to answer our research questions, with the first step consisting of a thorough methodological literature review and orientation on the taxonomic scope of our study system (Papers I and II); the second and third steps consisting of custom target capture bait set development; and testing (Paper II); the fourth step consisting of processing and identifying salep samples (Paper III); and the fifth step consisting of the analysis of their community composition over time and space, as well as potential drivers and consequences of variation observed therein (Paper III). All the bioinformatic steps required for analysing the target capture data generated in steps 3 (Paper II) and 4 (Paper III) are summarised in a pipeline flow-chart in Figure 5.

Literature review

Methods for monitoring wildlife trade To summarise the current state of wildlife trade monitoring, we conducted a (non-systematic) review of the literature describing case studies of various species, focusing on the last five years (2016-2021). We then clustered the papers into different types of applications, determined by the type of listing of the taxon studied and the specific aspects of the traded goods determining their legality. We sorted the methods by application, focusing only on the most common methods and further defined

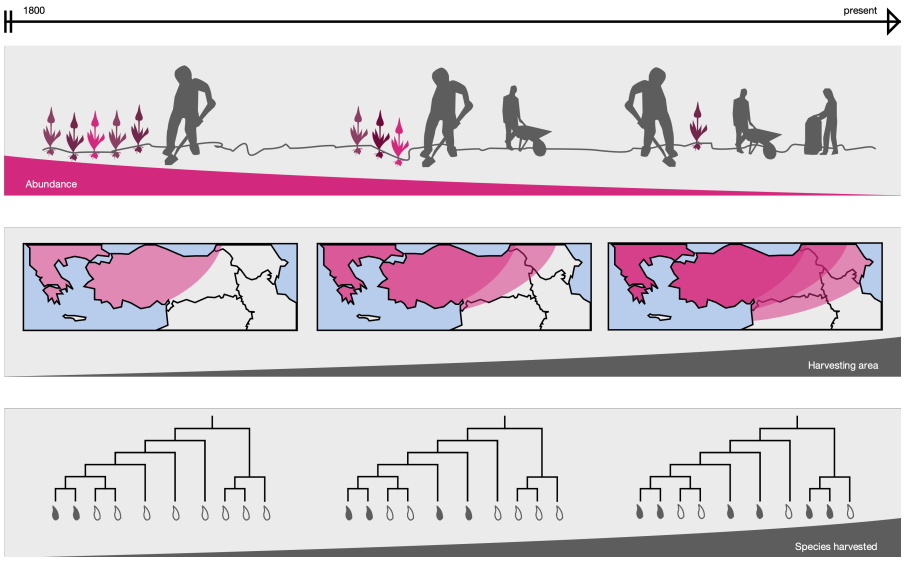

Figure 4: Graphical overview of the hypotheses explored in Paper III of this thesis. The abundance of salep species is expected to go down, leading people to shift their harvesting to new areas and new species.

these by their input requirements and available reference data. This resulted in a table summarising the strengths but also limitations of the different methods that we encountered, and their potential uses.

Potential target species for salep harvesting We also conducted a review to determine the potential target species of salep. Because one of the aims of the thesis was to identify tubers sold in present-day Turkey and Iran, and because Turkey's neighboring country Greece is also known for its high consumption of salep, an initial list of target species was drafted containing all tuberous orchids occurring in these countries. To avoid using multiple taxonomies and unstable nomenclature, we chose to rely on the Field guide to the Orchids of Europe and the Mediterranean (Kühn, Pedersen, and Cribb, 2019) for the final classification and prioritisation of target species. This field guide used the most up to date species names accepted at the time and contains range maps for each species, which we cross-referenced with the World Checklist of Selected Plant Families (Govaerts, 2019) where possible. This resulted in a list of 82 target species that could potentially be harvested for salep in these countries.

Custom bait development

Transcriptome assembly and target selection Having determined that monitoring of salep would ideally enable both species identification of degraded samples as well as the detection of (geographic) source population, we settled

on target capture as the method of choice. Target capture uses RNA baits or probes to bind to fragments of target DNA, which are subsequently amplified to increase their total share in the nuclear DNA pool, enabling the recovery of high coverage sequencing data for hundreds or even up to thousands of loci (Andermann et al., 2019). Our goal in Paper II was to design baits that uniquely target loci that are strictly single-copy within our target taxonomic group (i.e., species belonging to Orchidinae s.l.). Since target capture relies on prior availability of genomic resources to select the target loci, we searched for publicly available transcriptomes and genomes of the Orchidoideae subfamily on NCBI (https://www.ncbi.nlm.nih.gov/). At the time of bait design, no published reference genome was available for the Orchidoideae subfamily, but RNA-seq data was available for eight genera in the Orchideae tribe (subtribe Orchidinae), as well as three genera in the Cranichideae tribe (subtribe Goodyerinae) and three genera in the Diurideae tribe (subtribes Caladeniinae and Drakaeinae).

Twenty-three representative species were chosen and their reads were assembled with Trinity v2.10.0 (Grabherr et al., 2011) to generate *de novo* transcriptome assemblies for each, that were filtered and evaluated based on the recommendations by Yang and Smith (2014). The resulting contigs were used to identify orthogroups with OrthoFinder v2.5.1 (Emms and Kelly, 2019), which clusters transcripts based on shared ancestry. This was done twice: once for all Orchidoideae transcriptomes, and once for the Orchidinae transcriptomes. Due to gene duplication and missing data, not all orthogroups have exactly one gene copy per species; some have fewer and some have many more. To ensure universal enrichment across all our target taxa and avoid problems with paralogous genes, we therefore selected only orthogroups that had exactly one copy per genus. We made an exception for genes that were likely to play a role in the biosynthesis of glucomannan, which are interesting for functional analysis of the genetic variation underlying glucomannan concentration, a detailed list of which is presented in Paper II. A subset of 25 orthogroups was therefore allowed to have multiple copies, which were subsequently split into clusters with high genetic similarity that were treated as separate target loci.

Target filtering and bait development We mapped all candidate targets (which consisted of two sets: a specific set of Orchidinae-level orthogroups, and a more general set of Orchidoideae-level orthogroups) against a draft genome of *Ophrys sphegodes* Mill. (Osph-v1.1, unpublished) with GMAP (Wu, Reeder, et al., 2016) in order to eliminate genes with multiple hits, low sequence identity or short alignments. The best loci were kept, and since the top 500+ were heavily dominated by targets from the first (specific) set, we continued bait design with Orchidinae orthologs. In the end, 308 single copy targets were kept and supplemented with 31 glucomannan targets. Target sequences for all 14 Orchidinae transcriptomes were submitted to Daicel Arbor Biosciences (Ann Arbor, MI, USA), who developed a custom MyBaits kit using strict filtering for the single copy targets (to maximise sensitivity) and relaxed filtering for the glucomannan targets (to maximise coverage). To reduce the size of the kit,

single copy targets where <90% of the baits survived filtering and where less than 10 out of 14 taxa remained were removed. The baits were designed with a length of 70 bp and 3x tiling density (where each bait overlaps the previous bait by two thirds of its length) to optimise enrichment of degraded DNA, and collapsed with a minimum of 83% overlap and >95% sequence identity, yielding a final set of 60,000 baits targeting 205 genes (hereafter Orchidinae-205).

Custom bait testing

Library preparation and sequencing To test the baits (Paper II), we sourced DNA from 88 samples, representing 75 distinct taxa (including one putative hybrid and four accepted subspecies) belonging to our target species. These samples were fragmented to a size of approximately 400 base pairs and DNA libraries were prepared with unique dual indexing with one of two kits offered by Swift (Swift Biosciences, MI, USA), namely Accel-NGS 2S Hyb (Cat. No. 23023, 2021) or Turbo v2 (Cat. No. 44096, 2021). Target enrichment with the custom baits was carried out after pooling the DNA into pools of 2-8 samples, depending on their DNA quantities, and concentrating the DNA with Ampure XP (Beckman Coulter, CA, USA). RNA probes were hybridized at 62 °C for 24 hours, and 10 amplification cycles were carried out after enrichment. These samples were sequenced with a total sequencing output of 890M 150PE reads.

Gene recovery, alignment and tree reconstruction Raw reads were trimmed with Trimmomatic v0.39 (Bolger, Lohse, and Usadel, 2014) prior to assembling them into target sequences with Hybpiper v14 (Johnson, Gardner, et al., 2016). Hybpiper filters reads into on-target and off-target reads by mapping to the target reference file (consisting of transcriptome sequences used for bait development). This sorts the reads into their respective target loci, allowing for assembly of introns as well as exons due to the retention of read pairs on overhanging fragments. Due to the wide evolutionary distance spanning our target species, for the purpose of Papers II and III only exon sequences were used, which were aligned by codon with MACSE v2.06 (Ranwez et al., 2018) and trimmed with HmmCleaner (Di Franco et al., 2019), which is designed to remove poorly aligning segments that might be caused by sequence errors. This method was chosen because codon-aware alignment is often more reliable for coding sequences, and because block-trimming methods such as trimAl (Capella-Gutiérrez, Silla-Martínez, and Gabaldón, 2009), by trimming entire columns, have a tendency to remove phylogenetic information without necessarily improving phylogenetic inference.

The trimmed alignments were used to construct a maximum likelihood (ML) species tree as well as ML gene trees using IQ-TREE v2.1.2 (Minh, Schmidt, et al., 2020), which were edited to remove low support nodes and implausibly long branches before serving as input for multispecies coalescent (MSC) species tree reconstruction with ASTRAL-III (Zhang et al., 2018). Phylogenetic congruence and uncertainty was further explored by calculating gene and site concordance factors with IQ-TREE v2.1.2, and conducting a polytomy test with ASTRAL-III.

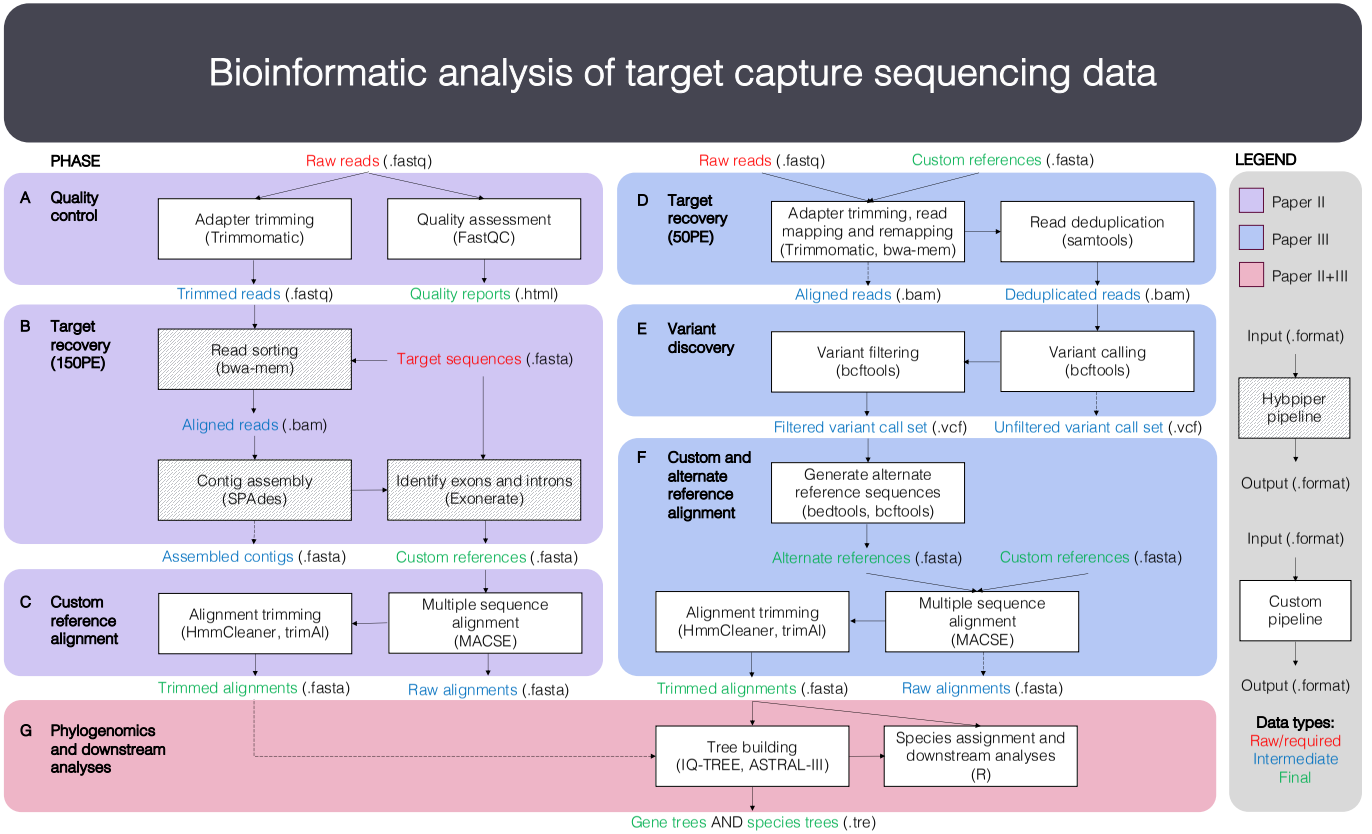

Figure 5: Bioinformatics pipeline for analysing target capture sequencing data. The figure illustrates pipelines used in steps 3 (Paper II) and 4 (Paper III) of the approach, and how they are related.

Performance assessment and comparison Locus overlap was assessed between the Orchidinae-205 kit and two alternative kits, one for enrichment of low copy nuclear genes in orchids (Orchidaceae-963) developed by Eserman et al. (2021), and one for flowering plants (Angiosperms-353) developed by Johnson, Pokorny, et al. (2019). This was done by performing a BLAST search of the target files used for probe design with BLAST+ v2.9.0 (Camacho et al., 2009). Paper II also assessed the relative performance of Orchidinae-205 compared to these kits in terms of capture and enrichment success. Target recovery information was obtained for species from the same tribe (Angiosperms-353 baits) or subfamily (Orchidaceae-963 baits) and visualised in R. A reduced nine-taxon tree was generated to enable a direct comparison of phylogenetic informativeness of the Orchidinae-205 markers with the Angiosperms-353 markers, with which the same species were enriched and sequenced by Baker et al. (2022). The phylogenetic informativeness (PI) of individual gene alignments of both sets of markers was inferred using PhyDesign (López-Giráldez and Townsend, 2011) against the calibrated Orchidinae-205 and Angiosperms-353 ML species trees. These trees were first calibrated with a relaxed clock model and a root age of 22 Mya (Inda, Pimentel, and Chase, 2012). The area under the curve (AUC) for each PI profile was subsequently calculated and used to evaluate the cumulative informativeness of each locus along the entire tree. These AUCs were ranked and their distributions were compared statistically to assess the differences between loci and impact of tree choice.

Detecting positive selection in glucomannan target genes To test the potential of the kit for answering questions about the evolution of glucomannan biosynthesis genes, we conducted a branch site test of positive selection with aBSREL (Smith et al., 2015) on 30 glucomannan target gene alignments that had (near) complete taxon coverage. This was done to ascertain whether any of these loci may have undergone episodic diversifying selection on any of the branches in the species tree. To account for the possible effect of gap-rich columns on inferences of positive selection, we applied two different gap thresholds and removed all columns which consisted of more than 25% or more than 50% gaps, respectively. aBSREL was run on both versions of each gene alignment with HyPhy v2.5 (Kosakovsky Pond et al., 2020), and p-values for each branch were corrected for multiple testing with the Holm-Bonferroni correction.

Salep identification

After testing the baits in Paper II on a selection of high quality reference material, we applied them to the identification of salep with the goal to characterise its species composition over time and space. For this purpose we sampled 99 historical salep tubers from various natural history and pharmacy museums in different countries across Europe, dating from the 1840s to the late 20th century. We also sampled 97 modern salep tubers sold on markets in different cities in Turkey and Iran, all collected in 2013-2014. Seven additional samples were added

to the reference database generated in Paper II before species identification of the tubers.

Library preparation and sequencing DNA was obtained from 186 severely degraded samples (consisting of 179 tubers and seven additional reference samples) in Paper III. DNA of these samples did not have to be fragmented because the median fragment length was already quite low, frequently falling below 80 bp. Because of the fragmentation as well as the low and often uncertain availability of DNA (caused by undetectable concentrations and high levels of contaminants), libraries for these samples were prepared with the Accel-NGS 1S kit (Cat. No. 10096, 2021) offered by Swift (Swift Biosciences, MI, USA), which was chosen for its compatibility with low input and single-stranded or otherwise damaged DNA, also using unique dual indexing. Due to the low and variable input quantities, DNA was pooled in groups of 2-4 samples instead of the standard eight, and concentrated with Ampure XP (Beckman Coulter, CA, USA) before target enrichment enrichment as described above, with the difference that more pre-capture (indexing) and post-capture PCR cycles were carried out in order to ensure sufficient DNA input for enrichment and sequencing. Given the short insert size of these samples, the libraries were sequenced with a shorter read length, generating 1.7 billion 50PE reads. Libraries for seventeen other tuber samples with exceptionally high DNA concentrations were prepared and sequenced using the methods described above for Paper II, generating 92 million 150PE reads.

Gene recovery, alignment and tree reconstruction The 150PE reads of seventeen high quality tuber samples were trimmed and assembled into target sequences with Hybpiper v14 (Johnson, Gardner, et al., 2016). The 50PE reads of all other samples were also trimmed with Trimmomatic v0.39 (Bolger, Lohse, and Usadel, 2014), but with tweaked settings to prevent unnecessary data loss and optimise short read retention. Shorter reads may be difficult to assemble, and hence we opted for a different strategy for exon recovery of the degraded samples. Instead of mapping to the original target reference file (consisting of the transcriptome sequences used for probe development), the exon sequences of reference samples generated in Paper II were used as a custom reference file for mapping the 50PE reads. The reference sample with the highest mapping rate was selected for reference-guided assembly of target loci with the bwa-mem algorithm of the BWA short read aligner v0.7.17 (Li and Durbin, 2009), followed by read deduplication with SAMtools v1.12 (Danecek et al., 2021). Consensus sequences were created after variant calling with BCFftools v1.12 (Danecek et al., 2021), masking zero coverage regions with BEDtools v2.30.0 (Quinlan and Hall, 2010). Tuber sequences with less than 60% breadth of coverage were excluded from the analysis. The untrimmed alignments generated in Paper II were enriched with all tuber sequences and sequences of the remaining seven reference samples using MACSE (Ranwez et al., 2018), and trimmed with HmmCleaner (Di Franco et al., 2019), before ML and MSC tree reconstruction.

Species assignment Species assignment based on DNA sequence information can be approached in multiple ways (Ross, Murugan, and Li, 2008). It can either rely on the estimation of genetic (or evolutionary) distances between samples and identify the most similar available reference sequence, or it can utilise a phylogenetic framework and identify the reference sample with which the query sample is most closely related. To make matters more confusing, there are distinct camps advocating different approaches to phylogenetic inference, one based on the concatenation of gene alignments into a supermatrix, and the other based on a coalescent framework that takes into account incomplete lineage sorting by considering the distinct evolutionary histories of individual gene trees (Gatesy and Springer, 2013; Springer and Gatesy, 2016; Wu, Song, et al., 2013). To handle these competing paradigms, we developed a decision-tree to arbitrate and find a consensus between alternative species assignments proposed by different methods. The alignments and species trees were used to assign species identities to salep tubers of unknown identity by using three distance-based methods (based on Kimura's 2 parameter (K2P) model (Kimura, 1980) and the ML and MSC tree branch lengths), selecting for each tuber the reference sample with the shortest distance; as well as two clade-based identifications (using the ML and MSC tree), selecting for each tuber the reference sample(s) with which it formed the smallest monophyletic clade. A species-level consensus identification was only made if at least two out of three distance-based methods (K2P, ML and MSC) agreed with at least one clade-based identification, or if both clade-based identifications were in agreement with at least one distance-based method. In all other cases a genus-level consensus identification was assigned.

Salep community composition

For Paper III, the input data for all other downstream analyses were the species assignments. These were used to quantify the variation in species composition between collections, over time and through space.

Temporal and spatial variation We divided all salep samples into five discrete age groups ranging from the 1840s to the present and calculated the distances between them using the Kulczynski dissimilarity index. We conducted multi-dimensional scaling of these distances to visualise their similarities, and assessed the variances within their respective centuries and age groups with the R package 'vegan' (Oksanen et al., 2022). Species co-occurrences within collections and age groups were analysed with the R package 'cooccur' (Griffith, Veech, and Marsh, 2016) and phylogenetic diversity metrics were calculated with the R package 'picante' (Kembel et al., 2010). Spatial variation was analysed for 21st century salep only, because these were the only collections for which we know their geographic origins (defined as point of sales). Our study area consisting of the border region of southeastern Turkey and north Iran was divided into five zones from west to east, ensuring roughly equal sample sizes between them. The relative frequencies of genera were calculated for each zone and plotted along a longitudinal gradient using loess regression.

Ecological and morphological variables To understand the relationship between (hypothesised) orchid availability and what was found on the markets, Paper III compared diversity of sold salep with the native orchid diversity for each market location. To achieve this, we built species distribution models (based on elevation and climate variables) with MaxEnt (Phillips, Anderson, and Schapire, 2006) for each of our target species and used these to generate presence-absence maps with custom habitat suitability thresholds for each species, chosen to optimise resemblance with approximate range maps based on expert opinion (Kühn, Pedersen, and Cribb, 2019). The individual presence-absence maps were then used to estimate species richness across our study area, as well as phylogenetic diversity (based on the ML species tree generated in Paper III). We also assessed different ecological variables that might explain different harvesting patterns, by taking elevational and flowering time data for each identified species and evaluating trends in the distribution of these variables over time. These variables gave us insight into the factors that might affect different species compositions now and in the past. The consequences of overharvesting were explored by measuring and weighing all 1200 tubers that were present in the historical and contemporary collections that we sampled (from which we sequenced only a subset) and plotting their weight and size distributions over time.

High-risk clades To account for sampling deficiencies, we wanted to know whether there might be species that could potentially be harvested for salep, but that did not make it into our dataset. We took the concept of "hot nodes", which are described as "nodes on the phylogeny that include significantly more plants traditionally used in medicine" (Saslis-Lagoudakis et al., 2012), to designate clades with a substantially higher chance of being used for salep in different time frames. To identify these high risk clades, we conducted a randomisation of all identified species across the MSC tree and marked nodes that had significantly more descendants that were identified as salep than expected by chance. To account for differences in frequency of observation, we did the same test based on the abundance of each species, conducting sampling of tips with replacement instead of without replacement, and using a stricter significance threshold to correct for the inflated statistical power associated with larger sample size.

Synthesis

How is wildlife trade regulated and monitored?

Wildlife trade regulations CITES is the main treaty that regulates international trade in wildlife products. Species can be listed on one of three Appendices with varying degrees of restrictions associated with them. Appendix I lists the most endangered species and places a total ban on international trade, except under some circumstances when trade does not have a commercial purpose. Some orchids are listed in Appendix I, but all of them belonging to different

subfamilies than the species analysed in this thesis. Appendix II lists species that are already threatened or may become threatened if trade is not tightly controlled, including look-alikes that are not necessarily overexploited (yet), but resemble species that are. Appendix II lists all species belonging to the orchid family, except those listed on Appendix I, and hence also applies to the species analysed in this thesis. Under the regulations of Appendix II, trade is only possible when in the possession of export permits and re-export certificates, which are issued only when accompanied by a non-detriment finding showing that trade will not be harmful to the survival of the species. Appendix III lists species that are not listed on either Appendix I or II, but whose trade is regulated by an individual Party to the Treaty, who relies on the cooperation of other Parties to help curb overexploitation. Since Appendix II places more strict requirements on the trade of species than Appendix III, no orchids are listed on Appendix III.

Methods for monitoring wildlife trade What matters for the validity of these regulations is whether a specimen was acquired before or after the date the species was listed. In cases where this can be documented, a pre-Convention certificate may be obtained which exempts the holder from standard permitting requirements. In the case of orchids, artificially propagated hybrids are also exempt from the permitting requirements, if easily recognisable. These regulations combined distinguish four types of characteristics of traded specimens that determine their legality and may require molecular and computational tools for their detection, namely: 1) the taxonomic identity (species, genus or family listed), 2) whether it was collected from the wild (source population), and if so, 3) in which country or (protected) area it was collected (geographic origin) and 4) whether it was collected pre- or post-Convention (age of the specimen). A literature search revealed that different methods are appropriate for each characteristic, which are summarised in Figure 6.

In brief, the age of a sample can only be established through isotope analysis (including radiocarbon dating), while taxon identity can be ascertained through a range of methods depending on the substrate. For hardwood and other tissues that contain detailed morphological information, computational image recognition can have high discriminatory power, and mass spectrometry may be able to identify species when their chemotypes are sufficiently different from each other. In other cases, genetic methods are preferable, with DNA barcoding as the go-to option, since it is relatively cheap and has easily accessible reference databases. Where DNA barcoding falls short is in the detection of more fine-grained genetic variation at or below the species level: for this, population genetic markers (such as single nucleotide polymorphisms (SNPs) generated by whole genome sequencing or reduced representation sequencing) are more suitable, but this requires the generation of custom reference databases for each group of species. Stable isotope analysis can also point out the geographic origins of a sample, but does not give any information on taxon identity, making population genetic markers the more sensible choice in cases where both are required.

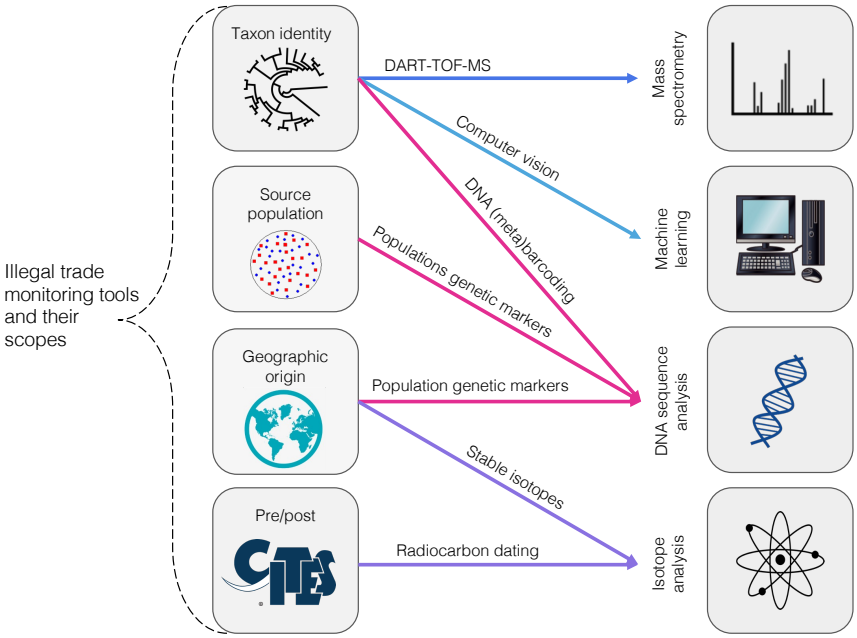

Figure 6: Some tools available for monitoring wildlife trade and their possible applications. The methods are listed over coloured arrows, connecting their respective analytical domains (on the right) to the characteristic determining whether trade is legal or not. DART-TOF-MS = Direct Analysis in Real Time (DART) coupled with time-of-flight (TOF) mass spectrometry.

Target capture as a novel approach Whereas traditional barcoding amplifies only one or a few markers, target capture generally amplifies hundreds or even thousands, and can therefore be considered an extension of the barcoding concept, but generating much more data. Very few examples exist to date that use target capture for monitoring trade (e.g. Manzanilla et al., 2022), and therefore it is not listed as a separate analytical technique in Paper I. However, one of the conclusions of Paper I is that conventional population genetic tools usually require high quality DNA, which is often unavailable for severely degraded specimens. In these cases, target capture may be a more suitable method for tracing geographic origins and source populations. In addition, even on higher taxonomic levels (i.e. species or genus level identification), DNA barcoding sometimes does not offer enough resolution to accurately identify species. Previous studies have shown that the percentage of species-level identifications for salep using traditional barcodes hovers around 40% (Ghorbani, Gravendeel, Selliah, et al., 2017), suggesting that more sequence information is needed to discriminate these orchids. Due to the degraded nature of salep tubers, sequencing success in itself is also variable, meaning that data might not even be obtained for a large percentage of tubers (Ghorbani, Gravendeel, Selliah, et al., 2017). This problem is exacerbated by

the fact that primers are not always tailored to the taxonomic group in question, resulting in amplification bias.

What sets target capture apart is that - when the baits are designed well - it is simultaneously able to effectively enrich DNA from even the most degraded DNA, but it also recovers large amounts of coding and non-coding DNA that can be used to differentiate specimens across different taxonomic scales, from the population level (Manzanilla et al., 2022) to the family level (Mandel et al., 2014) and even all flowering plants (Johnson, Pokorny, et al., 2019), thus enabling high-resolution but also broad-scale identification of traded specimens with severely fragmented or little DNA. As such, target capture falls somewhere between and *de facto* encompasses both DNA barcoding and population genetic markers in the methods overview presented in Paper I (Table 1), as it is able to meet the purposes that both of these methods are designed to achieve. A benefit of target capture is that it requires less material input than conventional population genetic sequencing approaches, although custom reference data will also need to be generated from scratch (except in the case of some pre-existing kits for which resources have already been developed). The latter necessitates substantial financial and bioinformatic investments before being able to deploy the method. Nonetheless, once the resources are available, target capture can outcompete traditional barcoding (Paper III) and yield a pool of tailored, taxon-specific loci from which a subset of markers can be mined that have higher enrichment and sequencing success and offer more phylogenetic resolution than existing universal markers (Paper II).

What is the potential of target capture for monitoring salep?

Orchidinae-205: strengths and limitations This thesis provides a foundation for the deployment of target capture for identification of salep tubers, by developing a custom bait set for enrichment of 205 markers tailored to the Orchidinae s.l., as well as generating the beginning of a reference database for potential target species (currently from Greece, Turkey and Iran), that can easily be expanded to new taxa from other countries. A custom bait set that is able to obtain large amounts of sequencing information from severely degraded specimens will be useful not just for identifying salep species, but can provide a stepping stone for more detailed genus or species specific analyses regarding country of origin or even source populations. Even though all orchid species are CITES listed and taxon identification at the family level would therefore theoretically suffice to determine the legality of trade, knowledge of the exact species and geographic origins of traded tubers will aid our understanding of what species and populations are most at risk and most in need of protection. This will help to prioritise conservation actions, making any fine-grained discriminatory power above and beyond the species level a valuable attribute of our kit.

In addition, Orchidinae-205 was designed to be more broadly useful to evolutionary and molecular biologists interested in the genetic basis of glucomannan production, by incorporating 31 candidate loci that code for enzymes involved in polysaccharide biosynthesis. A test case of this is

Table 1: A comparison of the methods used for identifying plants in trade with an indication of their applications and limitations. The methods for which target capture can be a substitute are highlighted in purple.

Tool	DNA (meta) barcoding	Population genetic markers	Computer vision and pattern recognition	DART-TOF MS	AMS/ ^{14}C dating	Stable isotope
Material input	Whole plants, organs, tissues, powder	Whole plants, organs, tissues, powder	Timber, leaves, flowers, pollen	All	Anything containing organic matter	Anything containing organic matter
Purpose of application	Determine taxonomic identity from genus to species level	Determine population or region of origin	Determine taxonomic identity, from genus to (sometimes) species level	Determine taxonomic identity at species level	Determine age of material	Determine the region of origin
Availability of reference data	Well-developed for temperate species, less for tropical species and regions	Needs to be developed and referenced for each species separately	Being developed for CITES protected timber and plants	Being developed for CITES protected timber	Calibration might be required depending on the sample	Needs to be developed for each region separate

demonstrated in Paper II, which shows the utility of pre-selecting candidate loci for assessing rates of evolution and the ratio of non-synonymous to synonymous substitutions (dN/dS ratio) in genes of interest, detecting positive selection on a number of terminal branches and some internal branches in the species tree. Positive selection events may help explain differences in glucomannan concentration and, knowing which species are harvested for salep, also explain human preferences for different species. Since this is only a proof of concept (demonstrating that it is possible to do selection scans on these loci), future studies into the evolutionary constraints and pressures shaping glucomannan target sequences will have to corroborate candidate sequence variation through functional validation of putative causal SNPs and experimental verification of glucomannan concentrations.

What target capture in general and this bait set in particular is not able to do is ascertain whether a tuber was harvested pre- or post-Convention. Even though the level of DNA damage is associated with the age of a sample (Orlando et al., 2021; Staats, Cuenca, et al., 2011), these molecular signatures cannot be used to authenticate age on short time scales such as the existence of CITES (which entered into force in 1975), and hence other methods such as radiocarbon dating are preferable. However, salep is so popular and traded so widely that, excepting museum specimens, most tubers on the market have probably been collected recently and almost certainly after 1975, rendering the assessment of tuber age a moot point.

Target enrichment and recovery success The final bait set presented in Paper II contains 60K probes covering 205 separate loci and a total target space of 306 kb. Of the 88 samples that were sent for sequencing in Paper II, seven did not yield sufficient sequencing reads for target recovery and were excluded from further analysis. Five of these belonged to the same DNA pool, so it is likely that either enrichment of this pool failed, or that the available DNA in this pool was too low. One other sample was sequenced twice; once with (theoretically) enough reads, but poor target recovery, and once with insufficient reads. Given the low enrichment of this sample, we think it may have been mislabelled and not actually belong to our target species. One other sample failed once, but was resequenced successfully. For the remaining samples, more than half (and frequently around two thirds) of the reads mapped on target, indicating that enrichment was successful. Target recovery was universally high (>80% of the target space), ranging from 254-330 kb.

This target recovery pertains to high quality DNA from leaf tissue. Target recovery for degraded leaf tissue and tubers was therefore expected to be (somewhat) lower. The target recovery of seventeen tubers that were sequenced with 150PE reads (Paper III) was on par with the reference samples from Paper II, ranging from 300-329 kb. For the remaining tubers and degraded reference samples, target enrichment was variable with 6-67% (averaging 42%) of reads mapping on target. Nineteen samples had a breadth of coverage of less than 60%, and 24 samples of less than 80%. The remaining 162 samples had a coverage of 262-328 kb.

The lower sequence recovery of the degraded samples may just be that - degradation of DNA and consequently less DNA to enrich and less to sequence. But part of it may also be caused by the lower sequencing depth, which, although the total number of reads per sample was approximately the same, was about three times lower due to the shorter length of the reads (50 versus 150 bases), and possibly even lower due to the overlap of forward and reverse reads of ultra-short fragments. Even though 11-13% of degraded samples were discarded from the analysis due to insufficient coverage (Paper III), our threshold for sequence retention was quite strict, meaning that more sample might have been retained if we are able to accept more data loss. The sequencing success with this bait kit (close to 90%) is still much higher than reported for traditional barcoding, which ranged from 19-69% (Ghorbani, Gravendeel, Selliah, et al., 2017).

Our kit also outperforms other bait sets designed for the entire orchid family (Orchidaceae-963) and all flowering plants (Angiosperms-353), both in terms of percentage locus recovery and absolute target recovery (Figure 7), demonstrating the benefits of designing custom baits tailored to the taxonomic groups of interest (Paper II). The overlap of Orchidinae-205 with these two kits was minimal, and for the loci that did overlap, the recovered length of loci targeted by our kit tended to be longer, which is probably caused by the lower enrichment efficiency of universal probes (Angiosperms-353) or those designed exclusively with a distantly related taxon such as *Phalaenopsis equestris* (Schauer) Rchb.f. (Orchidaceae-963).

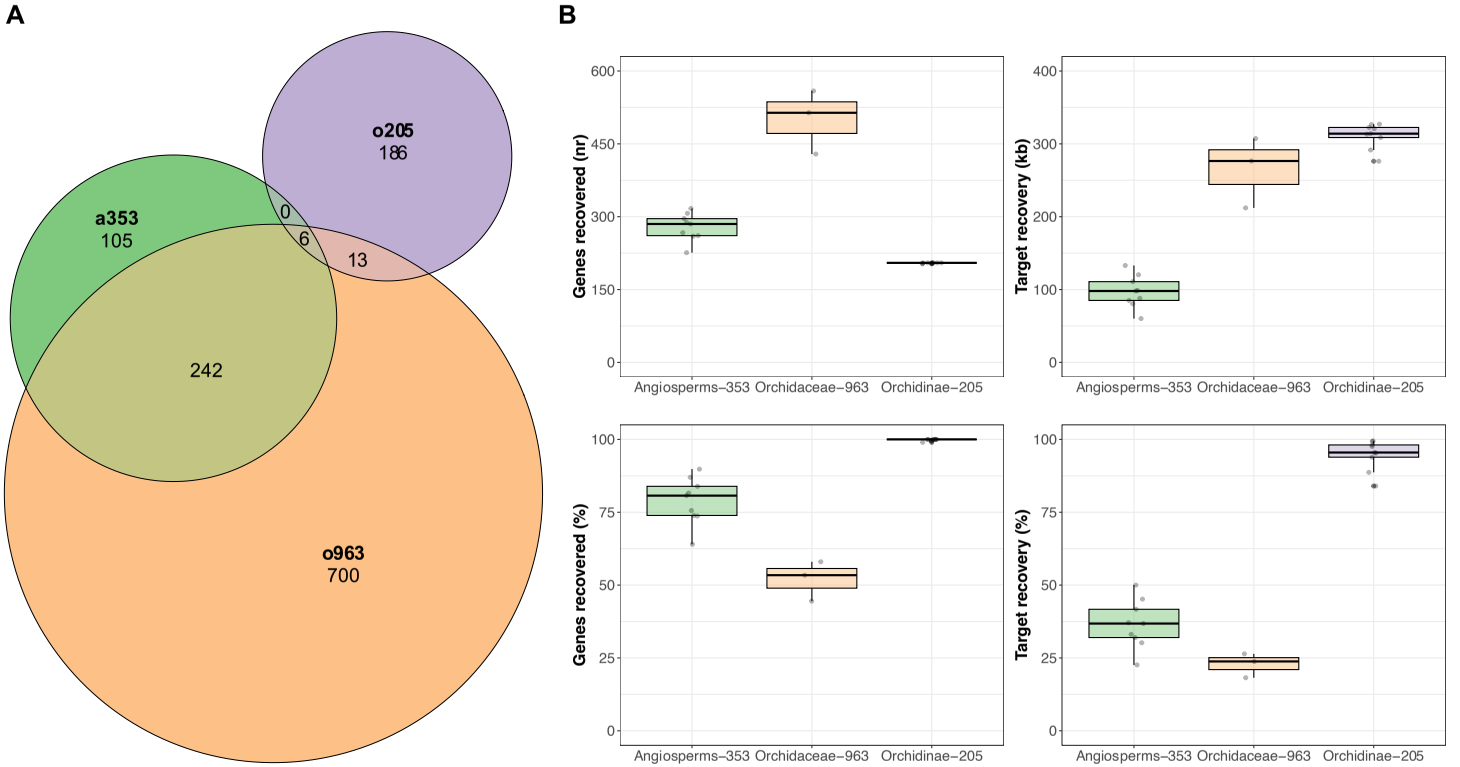

Figure 7: Comparison of Orchidinae-205 markers with two alternative bait sets. A) Overlap in target loci between Orchidinae-205, Orchidaceae-963 and Angiosperms-353. B) Target recovery in number (upper panels) and percentage (lower panels) of genes recovered (left) and base pairs recovered (right), for nine identical target species enriched with Angiosperms-353 (Baker et al., 2022) and Orchidinae-205 (this thesis) and 3 Orchidoideae species enriched with Orchidaceae-963 (Eserman et al., 2021).

Phylogenetic informativeness, support and concordance Consistent with the observation that custom baits are more effective at enrichment and target recovery, we find that when comparing alignments generated with nine identical species for Orchidinae-205 loci and Angiosperms-353 loci, the former are longer in length, have less missing data and contain more variable and parsimony informative sites than the latter. In addition, species ML and MSC trees generated with the Orchidinae-205 loci are more consistent with one another and have higher average node support, indicating that these trees are less impacted by uninformative genes and missing data (Paper II). Phylogenetic informativeness profiling further confirmed that Orchidinae-205 loci were more informative regardless of topological differences in the estimated species tree. Despite the better performance of Orchidinae-205 loci, in the full phylogeny generated in Paper II we do still observe weak support for a subset of both deep and shallow nodes characterised by rapid successive divergence events, reasons for which I will list below.

Because the MSC model accounts for incomplete lineage sorting (ILS) in its reconstruction of the species tree, discrepancies between the MSC and ML can be adequately explained by ILS if gene trees are informative and there is no gene flow. However, the difference between gene and site concordance factors (which should be equal if ILS is the only process causing conflict between gene trees) shows substantially lower gene concordance factors for some nodes in our tree, indicating that these nodes are poorly resolved and that this is probably caused by lack of phylogenetic information in the gene alignments and corresponding gene trees rather than by ILS (Minh, Hahn, and Lanfear, 2020). A polytomy test (Sayyari and Mirarab, 2018) shows for many of these same (often shallow) nodes that they are statistically better represented by a polytomy, which likely reflects a soft polytomy caused by lack of data rather than a hard polytomy (Paper II). Some of the species relationships affected by this phenomenon are recently radiated clades within *Ophrys* and *Serapias* (Breitkopf et al., 2015; Inda, Pimentel, and Chase, 2012), or rapidly hybridising lineages within *Dactylorhiza* and *Gymnadenia* (Brandrud et al., 2020; Hedrén, Lorenz, and Ståhlberg, 2018). For the former, including intron sequences might just give that extra bit of phylogenetic information required to generate reliable gene trees and allow robust species tree inference. In case of the latter, topological conflict between trees is partially explained by gene flow, which could be solved for some known or putative hybrids by separating haplotypes and constructing alignments and trees with multiple alleles corresponding to the different putative parent species (Nauheimer et al., 2021). It is noteworthy that in some cases, nodes characterised by high discordance still have high support, which indicates that relying on bootstrapping or posterior probabilities alone for interpreting phylogenetic uncertainty can be misleading, since gene trees can have disparate topologies that still consistently lead to the same species tree.

Identification of salep tubers The extent of phylogenetic discordance was limited to a few clades (as described above) and hence does not necessarily

affect all species that are harvested for salep. Paper III applied the custom bait set to 196 individual tubers in order to establish their species identity and ultimately assess variations in community composition over time and space. The identification of specimens of unknown identity requires a phylogenetic framework with comprehensive taxon sampling in order to place the unknown tubers with high confidence and match them to their closest relative. This phylogenetic framework was developed in Paper II, but was not 100% complete. Seven samples (including six species that were previously missing) were therefore added in Paper III, including some likely candidate species for salep harvesting in the past and in the present.

The species identification method explained in the Approach has the advantage that it exposes conflicting species assignments, and hence methodological choices that might inadvertently impact conclusions about species identity if not compared with their alternatives. However, it also implies that more specimens will be identified at the genus level than if a single method were adopted. In practice, this difference was small, amounting to only 10-15% of tubers, depending on the threshold of node support adopted for accepting clades. These tubers mostly belonged to two genera, namely *Anacamptis* and *Serapias*, the latter of which all clustered with the same species (*Serapias vomeracea* (Burm.f.) Briq.) in the ML tree but was inconclusive in the MSC species tree. *Anacamptis* tubers had a conclusive species identification in 75% of cases, but were sometimes inconclusive in both trees, and sometimes monophyletic with a single species in the ML tree (*Anacamptis palustris* (Jacq.) R.M.Bateman, Pridgeon & M.W.Chase) but inconclusive in the MSC tree. The same pattern was observed for *Ophrys* tubers, all of which clustered with one species in the ML tree (*Ophrys umbilicata* Desf.), but with a different one (*Ophrys kojurensis* Gölz) or inconclusively in the MSC tree. A handful of *Orchis* tubers (7%) was also identified at the genus level, but this was due to lack of monophyly in both species trees, indicating either a lack of or erroneous data, potential hybrid lineages or missing reference taxa.

The results described above demonstrate the potential of target capture to identify the majority of tubers at the species level, reaching more conclusive identifications than previous barcoding studies (Ghorbani, Gravendeel, Selliah, et al., 2017). The inability to discriminate between closely related species is very localised in only a few lineages where the ML and MSC approaches to phylogenetic inference tend to disagree. As described above, this is mostly likely caused by either ILS, or insufficient phylogenetic information in the gene alignments; and in case of the latter could be remedied by the inclusion of the longer intron sequences recovered in Paper II.

Difficulties in species identification might remain in *Serapias* and *Ophrys*, which radiated very recently, around 1 and 5 Mya, respectively (Inda, Pimentel, and Chase, 2012), and still contain many species whose names and taxonomic level and status are disputed. The bait kit presented in this thesis will therefore hopefully also be used for fundamental systematics research, which is necessary to build on existing taxonomic insights and update species classifications. This will benefit identification of salep tubers in the long run, since a stable nomenclature and coherent species concepts are the basis of any attempt at identifying species.

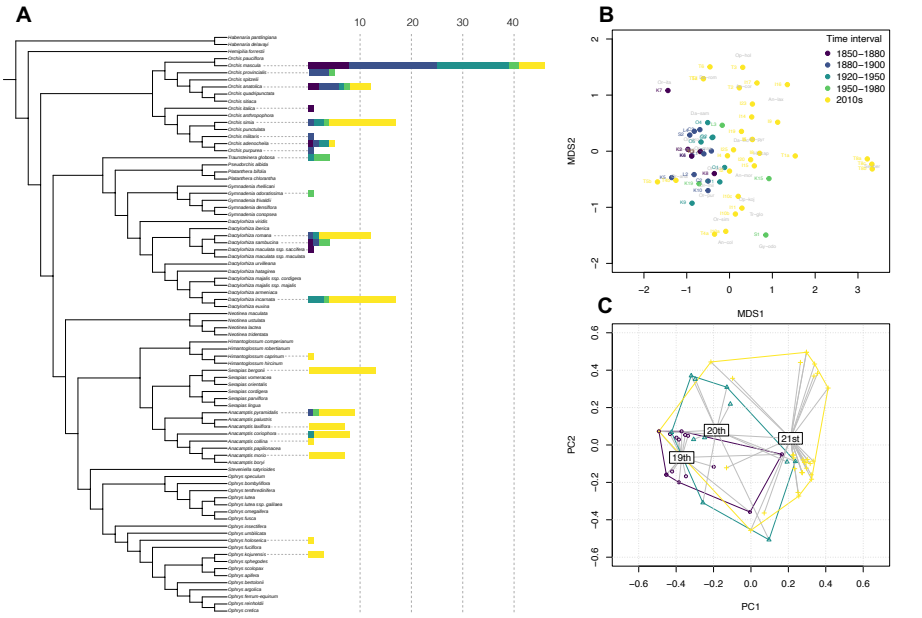

Figure 8: Heterogeneity in species composition of salep collections. A) Number of tubers identified per species based on their closest reference under the multispecies coalescent. Colours of the stacked bar chart correspond to the legend in panel B. B) Non-metric multidimensional scaling of collection diversity, coloured by age. C) Within group variance of collections, grouped by century. Coloured lines indicate the circumference of the collection dispersion within each group. Grey lines indicate the distance of each collection to the centroid of the group. Distances between collections in B and C are based on the Kulczynski dissimilarity index using the species identifications made in A.

How - and why - does traded salep vary across time and space?

Temporal variation The positive salep identifications obtained in Paper III were further analysed by examining the variation in species composition of salep across individual collections, time periods and geographic regions. One of the main findings from this analysis is that salep used to be phylogenetically clustered, with its core diversity centered on a few *Orchis* species (most notably *Orchis mascula*). Over time this pattern changes and the dissimilarity between collections, as well as the variance in distance of collections to the centroid of their respective age groups (e.g. a century) becomes larger, indicating increasing phylogenetic dispersion (Figure 8). Recent collections contain salep tubers from many more genera, and there is no clear species that dominates. *Orchis* spp. are most commonly substituted for *Dactylorhiza*, followed by *Anacamptis* and to a lesser extent *Serapias*. The latter two are observed almost exclusively among recent collections from Turkey and Iran, whereas *Dactylorhiza* was already a common alternative to the more traditional *Orchis* species in the 19th century.

Geographic variation Among the modern samples from 2013-2014, the distribution of salep species across space was highly heterogeneous. *Orchis* was found in the Western and Eastern range of our study area, but not in the middle, suggesting a local reduction in availability. Instead, *Serapias* dominates in southeastern Turkey, *Anacamptis* in northwestern Iran, and both *Anacamptis* and *Dactylorhiza* are observed in equal proportions in (central) North Iran. The presence-absence maps generated in this study suggest that while Turkey has a richer native orchid flora, fewer species are traded, whereas Iran has a lower estimated richness of native orchid flora, but sells more species. This may be partially due to data-deficiency of Iran in terms of available species records, but also because Iran is becoming a hotspot for trade, providing a new source for species (especially in the east) that are traditionally preferred but no longer available in the border region of Turkey and Iran where salep is popularly consumed by the local Kurdish population (Youssef et al., 2019).

Causes and consequences To explore the possible causes for the changes in species composition over time, we compared the median elevation and flowering times of species harvested in five discrete time intervals. As species occur at different altitudes, changes in species composition might (in part) be explained by, for example, choosing to collect orchids from higher up in the mountains versus the lowland. Similarly, since species have different flowering times and tuber collection is closely associated with flowering (Molnár V. et al., 2017; Sezik, 2002), changes in species composition could also be (partially) explained by choosing to collect salep earlier or later in the season, when different species are in flower. Whereas differences in elevation displayed no clear trend over time, we did see that the onset of flowering shifted forward and the end of flowering was on average pushed back in recent decades (Figure 9A-B). The longer average duration of flowering and higher spread around the mean of onset, midpoint and end of flowering might therefore indicate that salep is collected more widely throughout the season than before.

At the same time, we observe that individual tubers are decreasing in size and weight (Figure 9C-D). While this could be an artefact of species composition, since species vary in their tuber morphology (Molnár et al., 2017), a close up of size and weight trends in the commonly harvested genera *Orchis*, *Dactylorhiza*, and *Anacamptis*, shows that the downward trend persists at the level of individual genera. This suggests that the effect is not caused by taxonomic differences at the level of genera, and might be a universal effect among commonly harvested species, which leaves two other possible explanations for the observed trend: either tubers are harvested increasingly early in the season, or the tuber size at the end of the growing season is smaller than before. The first would go against the self-interest of salep harvesters who sell salep by weight and can make a larger profit with fully grown tubers. The second would imply that tubers are harvested from younger individuals, as terrestrial orchids are perennial species that produce larger tubers each year, and could be caused by the depletion of mature individuals from the population due to overharvesting. Anecdotal

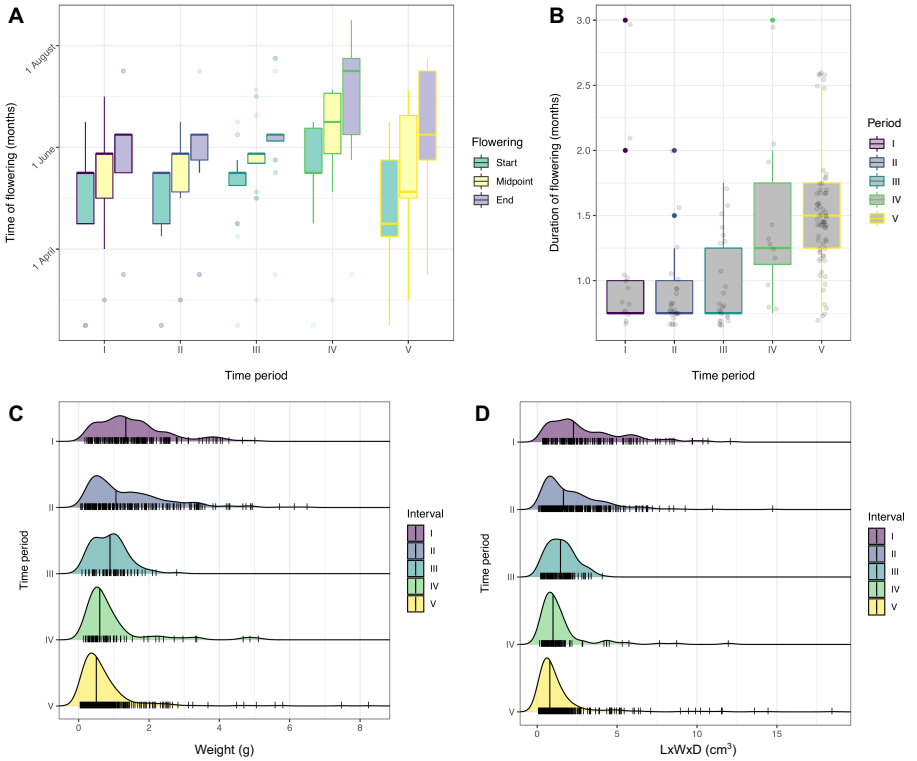

Figure 9: Variations in flowering time and tuber size and weight over time. A) Start, midpoint and end of flowering of orchid species identified as salep per time period. B) Duration of flowering of the species identified as salep per time period in number of months. Data points are individual tubers, with opaque colours indicating outliers. C) Shifts in the distribution of tuber weight over time. D) Shifts in the distribution of tuber size (measured in three dimensions) over time. In C) and D), tick marks indicate individual tubers, and the large vertical line indicates the median. Time interval I = 1840-1879; II = 1880-1919 ; III = 1920-1949; IV = 1950-1979; and V = 2013-2014.

evidence from the literature seems to suggest the second explanation is a likely scenario (Ghorbani, Gravendeel, Naghibi, et al., 2014; Kreziou, de Boer, and Gravendeel, 2016).

Future target species? Lastly, in Paper III we identified clades that were significantly enriched for salep species presence and abundance. Consistent with our observations of the temporal changes in salep species, we find that *Orchis* as an entire genus is identified as a risk clade in every single historical time interval and century (from the mid-19th century to the late 20th century), but that this risk clade contracts to a smaller portion of the genus (presence

data) or even disappears completely (abundance data) in the 21st century. In contrast, the sister genera *Anacamptis*, *Serapias* and *Himantoglossum* are considered at risk in the present, but not at all before the turn of the century. The identification of risk clades in *Dactylorhiza* is variable, as almost this entire genus is identified as significantly enriched when using abundance data, but not at all when using presence data. This highlights the impact that a concentration of many individuals in the same (few) species has for the analysis. Among historical collections, a small clade within *Dactylorhiza* is persistently marked as enriched, demonstrating the consistent popularity of a few species at least throughout the 19th century.

Considering the breakdown in phylogenetic clustering of salep over time, combined with knowledge that the harvesting season is now potentially longer and popular species might be locally depleted (lowering the average age and abundance of local populations), we consider it likely that salep harvesting will continue to push into new territory and towards new species. Thus, assuming that indiscriminate harvesting is only going to increase, species belonging to the risk clades identified here, if not already being under threat now, might become so in the near future.

Concluding remarks

The vicious cycle of salep overharvesting

New insights into pressures, state and impact The core objective of this thesis has been to shed light on the "black box" of wildlife trade, namely the pressures, state and impact of overexploitation of salep orchids (Figure 2), and to explore to what extent target capture is able to fulfill this role. While Paper I provided a methodological exploration and Paper II developed a custom bait set and reference panel for orchid identification, Paper III applied these resources to an empirical case and elucidated some of the spatio-temporal patterns in orchid trade. Linking these results back to the DPSIR model, Paper III was able to discover changing pressures on wild orchid populations as humans shift their harvesting efforts to new species, new areas, younger individuals and different seasonal timing. It was also able to detect a likely impact on commonly harvested orchid species by observing a downward trend in tuber size and weight over time, which is indicative of a younger population that cannot sustain itself. While this suggests a declining state of orchid populations in terms of their abundance and age composition, other aspects of population state such as genetic diversity and effective population size, as well as impacts like inbreeding depression, reduced adaptive potential and possible extinction risk remain underexplored. However, the resources presented here can theoretically be used to shed light on these aspects too when population level sampling of species is available, using common population genetic techniques.

The drivers of overexploitation Knowledge about the pressures, state and impact of overharvesting as they relate to different orchid species can reveal

changes in the drivers of human action and inform more effective responses to this environmental problem. The long stability and recently accelerating turnover of salep composition over time suggests that what used to be an ecological equilibrium is now increasingly under stress from outside forces. The consumption of salep is a century (or even millennia) old tradition (Ece Tamer, Karaman, and Utku Copur, 2006) that has arguably long been, and could conceivably in the future also be, enjoyed sustainably when conducted by communities holding traditional ecological knowledge of sustainable harvesting practices and associated with home use (Molnár V. et al., 2017). However, shifting market forces have rendered orchid harvesting a lucrative business and turned it into a primarily profit driven enterprise, especially in areas where salep is not traditionally consumed (Ghorbani, Gravendeel, Naghibi, et al., 2014), while on the side of consumers living in urban areas (where demand is growing most rapidly), increasingly complex supply chains over large distances disrupt the connections with the source of their food and hinders transparency of its provenance and sustainability (Trienekens et al., 2012). For people who harvest and trade wild plants to make a living and not for their own consumption, these market forces in turn diminish the interest in proper maintenance of and stewardship over orchid populations by prioritising short-term revenues over long-term availability (especially under economic conditions of poverty), which incentivises the exploitation of more species over larger areas, with less regard for sustainable harvesting practices (Leão, Lobo, and Scotson, 2017).

Monitoring and other responses Currently, the main response to this problem is the permitting system introduced by CITES; but this is a stick rather than a carrot for more sustainable harvesting, and by its very nature is limited in scope because it only addresses international trade and does not concern the domestic market, where demand may still be substantial. In addition, the enforcement of compliance to CITES regulations is cumbersome and heavily relies on monitoring techniques that may not be easy to implement at scale, sometimes preventing species from being listed to begin with (Hughes, 2021). The bait set and resources presented in this thesis here open the door to more effective monitoring of traded orchid tubers. Yet, without ways to prevent laundering and other forms of illegal harvesting and trade, the current institutional response against overexploitation is likely to remain ineffective as long as it neglects the drivers that promote the collection and trade of wild orchids, which are primarily economic in nature (Ghorbani, Gravendeel, Naghibi, et al., 2014; Kreziou, de Boer, and Gravendeel, 2016).

In this respect, initiatives to promote and increase awareness of sustainable harvesting or cultivation of salep are promising avenues to break the vicious cycle of species overexploitation. Reconnecting consumers to their food through knowledge of its origins and production could help drive commercial interests towards more sustainable harvesting or cultivation practices (Dulić et al., 2020). Increasing awareness of the history of salep and what species are ‘traditional’ as opposed to recent substitutes, might spark an interest in authentic salep that is

superior in quality and taste (Charitonidou et al., 2019) and, when accompanied by certification schemes such as FairWild that ensure sustainable harvesting (FairWild Foundation, 2010), incentivise the protection of these species rather than their depletion. Lastly, the introduction of artificially propagated salep could be one strategy to increase supply, while releasing pressure on existing wild populations, although a prerequisite for the success of such endeavours is that the cultivated tubers are perceived by consumers as a valid and high quality substitute, on par with wild collected tubers (Schippmann, Leaman, and Cunningham, 2002).

While a detailed socio-ecological analysis is beyond the scope of this thesis, future research could further investigate the society-environment interactions that characterise orchid trade, in order to better understand its dynamics and promote change towards a more sustainable system. In such a system, monitoring can hopefully not only be a tool for detecting and enforcing compliance, but also a resource to shed light on and improve changing human-plant relations.

Future directions

Priorities for development and application The papers presented in this thesis provide a first step towards accurate and reliable species identification of traded orchids, but is by no means the final destination. There are two ways to expand the resources developed here: by adding individuals to the reference database, or by adding sequence information to existing reference samples. A first priority should be to add missing species and subspecies from countries where salep might also be harvested (such the Balkans), because these would otherwise run the risk of going undetected. Subsequent investments can be prioritised according to the conceptualisation below (Figure 10).

Species relationships with low phylogenetic resolution or high discordance due to rapid divergence, will benefit from increasing genomic breadth of coverage, either by adding intron sequences to the exons (increasing locus length), or by adding new targets (increasing the number of loci). As Paper II has argued, it pays off to improve locus length and information and information before adding more loci, because reliable gene trees are the rate limiting step for accurate species tree inference under the two-step multispecies coalescence (Xu and Yang, 2016).

Species that have a high risk of being harvested will benefit from more dense sampling, allowing more fine-grained identification at the population level. Species that have undergone rapid divergence *and* suffer a high risk of being harvested should therefore be a first priority for continued resource development, both by adding reference material and sequence information. On the basis of results from Paper III, we can conclude that species in this category mostly belong to *Anacamptis* and *Serapias*. *Ophrys* spp. are also in need of more sequence information, but are less commonly harvested and therefore a lower priority for forensics (although they might be a priority for fundamental research aiming to understand the complex radiation of *Ophrys* species).

Less pressing in terms of sequence information, but still at high risk of being traded are *Orchis* and *Dactylorhiza* species. These will mostly benefit from added population level sampling to trace their origins to and manage specific source populations.

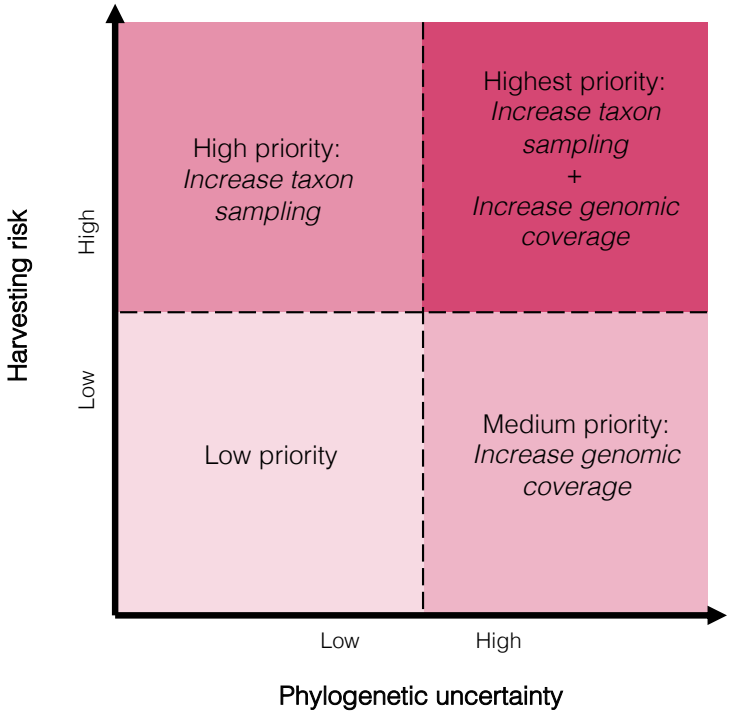

Figure 10: Schematic overview of the prioritisation of resource allocation based on the harvesting risk and the level of phylogenetic uncertainty impacting species identification.

Costs and scalability A parallel (perhaps counter-intuitive) development to enable more effective monitoring could be to downsize the bait set. As mentioned, target capture is financially and computationally intensive, and even though the baits have been developed and a reference database exists for three countries, the application to new (unknown) samples can still be prohibitively expensive. While it is possible in many cases to lowering the per-sample costs by using different reagents and buying in bulk (Hale et al., 2020), the most cost-effective way to scale up salep identification would be by identifying a subset of markers that are be able to reconstruct most species relationships, without having to enrich and sequence all 205. Such a reduced set of baits for custom barcodes that are tailored to Orchidinae s.l. are probably more effective at enriching DNA and

producing sufficient phylogenetic information than existing universal barcodes, and could be implemented on a larger scale at significantly lower costs.

Beyond identification As a relatively novel and costly method, target capture is an underexplored tool for monitoring wildlife, but a promising one that can tackle three out of four major questions concerning illegal trade, namely: species identity, geographic origin and source population. While this thesis has only demonstrated the utility of Orchidinae-205 in answering the first of these questions, a logical next step would be to zoom in on individual species, especially those that appear to be gaining popularity or are at risk of becoming future targets, and establish a reference panel with geo-referenced population level data to determine where tubers come from as they appear on the market. Targeted interventions could then be made to protect populations locally and prevent future overexploitation. Conversely, the disappearance of species from the market might serve as a warning sign that populations are declining and that conservation efforts are needed to reverse this trend.

Target capture could also assist in identifying whether traded orchids are collected from the wild or artificially propagated, as the latter should be less diverse and have a well-defined genetic composition. While commercially cultivated salep is still far from being realised, efforts taken to create artificial hybrids that do not occur in the wild and are genetically distinguishable from wild hybrids (e.g. Antonetti et al., 2021), are a promising route towards sustainable production and consumption of salep in the future. Consumer preferences for polysaccharide composition should be taken into account in this effort, and the target loci presented in this thesis can help to understand which genetic variation (in which species) is connected to this trait. Lastly, the resource presented in this thesis offers the opportunity to study human-plant relations beyond the realm of legal and illegal trade by exploring past plant use with the use of ethnobotanical and museum collections, and to answer fundamental biological questions about the evolution of this fascinating and valuable group of species.

References

Andermann, T. et al. (2019). “A Guide to Carrying Out a Phylogenomic Target Sequence Capture Project”. *Front. Genet.* vol. 10, p. 1407.

Antonetti, M. et al. (2021). “*Himantoglossum adriaticum* H. Baumann × *Himantoglossum robertianum* (Loisel.) P. Delforge: A New Interspecific Hybrid Assessed by Barcoding Analysis”. *Plants* vol. 10, no. 1.

Baker, W. J. et al. (2022). “A comprehensive phylogenomic platform for exploring the angiosperm tree of life”. *Systematic biology* vol. 71, no. 2, pp. 301–319.

Bashyal, R. and D. L. Roberts (2023). “Assessing the identification uncertainty in plant products traded as traditional Asian medicines”. *J. Nat. Conserv.* vol. 73, p. 126410.

Berdan, F. F. et al. (2009). "Production and use of orchid adhesives in Aztec Mexico: the domestic context". In: *Housework: Craft production and domestic economy in ancient Mesoamerica*, pp. 148–156.

Boivin, N., R. Crassard, and M. Petraglia (2017). *Human Dispersal and Species Movement: From Prehistory to the Present*. Cambridge University Press.

Bolger, A. M., M. Lohse, and B. Usadel (2014). "Trimmomatic: a flexible trimmer for Illumina sequence data". *Bioinformatics* vol. 30, no. 15, pp. 2114–2120.

Booker, A. (2014). "The transformation of traditional Asian medical knowledge into international commodities : the link between traditional medicines and the international market". PhD thesis. University College Londen.

Bradley, P. and S. Yee (2015). *Using the DPSIR Framework to Develop a Conceptual Model: Technical Support Document*. Tech. rep. EPA/600/R-15/154. Environmental Protection Agency.

Brandrud, M. K. et al. (2020). "Phylogenomic Relationships of Diploids and the Origins of Allotetraploids in *Dactylorhiza* (Orchidaceae)". *Syst. Biol.* vol. 69, no. 1, pp. 91–109.

Breitkopf, H. et al. (2015). "Multiple shifts to different pollinators fuelled rapid diversification in sexually deceptive *Ophrys* orchids". *New Phytol.* vol. 207, no. 2, pp. 377–389.

Bulpitt, C. J. (2005). "The uses and misuses of orchids in medicine". *QJM* vol. 98, no. 9, pp. 625–631.

Camacho, C. et al. (2009). "BLAST+: architecture and applications". *BMC bioinformatics* vol. 10, pp. 1–9.

Capella-Gutiérrez, S., J. M. Silla-Martínez, and T. Gabaldón (2009). "trimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses". *Bioinformatics* vol. 25, no. 15, pp. 1972–1973.

Charitonidou, M. et al. (2019). "Implications of salep collection for the conservation of the Elder-flowered orchid (*Dactylorhiza sambucina*) in Epirus, Greece". *J. Biol. Res.* vol. 26, p. 18.

Chinsamy, M., J. F. Finnie, and J. Van Staden (2011). "The ethnobotany of South African medicinal orchids". *S. Afr. J. Bot.* vol. 77, no. 1, pp. 2–9.

CITES Secretariat (2022). *World Wildlife Trade Report 2022*. Tech. rep.

Cooney, R. et al. (2018). *Wild Life, Wild Livelihoods: Involving Communities on Sustainable Wildlife Management and Combating Illegal Wildlife Trade*. Tech. rep. UNEP, Nairobi.

Danecek, P. et al. (2021). "Twelve years of SAMtools and BCFtools". *Gigascience* vol. 10, no. 2.

de Boer, H. J. et al. (2017). "DNA metabarcoding of orchid-derived products reveals widespread illegal orchid trade". *Proc. Biol. Sci.* vol. 284, no. 1863.

Di Franco, A. et al. (2019). "Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences". *BMC Evol. Biol.* vol. 19, no. 1, p. 21.

Dulić, J. et al. (2020). "Implementation of SWOT analysis to evaluate conservation necessity and utilization of natural wealth: terrestrial orchids as a case study". *J. Environ. Planning Manage.* vol. 63, no. 12, pp. 2265–2286.

Ece Tamer, C., B. Karaman, and O. Utku Copur (2006). "A Traditional Turkish Beverage: Salep". *Food Rev. Int.* vol. 22, no. 1, pp. 43–50.
Ellegren, H. (2014). "Genome sequencing and population genomics in non-model organisms". *Trends Ecol. Evol.* vol. 29, no. 1, pp. 51–63.
Emms, D. M. and S. Kelly (2019). "OrthoFinder: phylogenetic orthology inference for comparative genomics". *Genome Biol.* vol. 20, no. 1, p. 238.
Eserman, L. A. et al. (2021). "Target sequence capture in orchids: Developing a kit to sequence hundreds of single-copy loci". *Applications in Plant Sciences* vol. 9, no. 7, e11416.
FairWild Foundation (2010). *FairWild Standard: Version 2.0*. Tech. rep. FairWild Foundation, Weinfelden, Switzerland.
Fariss, B. et al. (2023). "Catalyzing success in community-based conservation". *Conserv. Biol.* vol. 37, no. 1, e13973.
Gatesy, J. and M. S. Springer (2013). "Concatenation versus coalescence versus "concatalescence"". *Proceedings of the National Academy of Sciences* vol. 110, no. 13, E1179–E1179.
Gavin, M. C. et al. (2015). "Defining biocultural approaches to conservation". *Trends Ecol. Evol.* vol. 30, no. 3, pp. 140–145.
Ghorbani, A., B. Gravendeel, F. Naghibi, et al. (2014). "Wild orchid tuber collection in Iran: a wake-up call for conservation". *Biodivers. Conserv.* vol. 23, no. 11, pp. 2749–2760.
Ghorbani, A., B. Gravendeel, S. Selliah, et al. (2017). "DNA barcoding of tuberous Orchidoideae: a resource for identification of orchids used in Salep". *Mol. Ecol. Resour.* vol. 17, no. 2, pp. 342–352.
Givnish, T. J. et al. (2015). "Orchid phylogenomics and multiple drivers of their extraordinary diversification". *Proc. Biol. Sci.* vol. 282, no. 1814.
Gouda, S. et al. (2020). "Wildlife forensics: A boon for species identification and conservation implications". *Forensic Sci. Int.* vol. 317, p. 110530.
Govaerts, R. (2019). "World Checklist of Selected Plant Families in the Catalogue of Life".
Grabherr, M. G. et al. (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome". *Nat. Biotechnol.* vol. 29, no. 7, pp. 644–652.
Griffith, D. M., J. A. Veech, and C. J. Marsh (2016). "cooccur: Probabilistic Species Co-Occurrence Analysis in R". *J. Stat. Softw.* vol. 69, pp. 1–17.
Hale, H. et al. (2020). "Strategies for reducing per-sample costs in target capture sequencing for phylogenomics and population genomics in plants". *Appl. Plant Sci.* vol. 8, no. 4, e11337.
Hedrén, M., R. Lorenz, and D. Ståhlberg (2018). "Evidence for bidirectional hybridization between *Gymnadenia* and *Nigritella*". *Journal Europäischer Orchideen* vol. 50, no. 1, pp. 43–60.
Hinsley, A., H. J. de Boer, et al. (2017). "A review of the trade in orchids and its implications for conservation". *Bot. J. Linn. Soc.* vol. 186, no. 4, pp. 435–455.
Hinsley, A., A. Nuno, et al. (2017). "Estimating the Extent of CITES Noncompliance among Traders and End-Consumers; Lessons from the Global Orchid Trade". *Conserv. Lett.* vol. 10, no. 5, pp. 602–609.

Hollingsworth, P. M., S. W. Graham, and D. P. Little (2011). "Choosing and using a plant DNA barcode". *PLoS One* vol. 6, no. 5, e19254.

Hollingsworth, P. M., D.-Z. Li, et al. (2016). "Telling plant species apart with DNA: from barcodes to genomes". *Philos. Trans. R. Soc. Lond. B Biol. Sci.* vol. 371, no. 1702.

Hughes, A. C. (2021). "Wildlife trade". *Curr. Biol.* vol. 31, no. 19, R1218–R1224.

Inda, L. A., M. Pimentel, and M. W. Chase (2012). "Phylogenetics of tribe Orchideae (Orchidaceae: Orchidoideae) based on combined DNA matrices: inferences regarding timing of diversification and evolution of pollination syndromes". *Ann. Bot.* vol. 110, no. 1, pp. 71–90.

IPBES (2022). *Thematic Assessment Report on the Sustainable Use of Wild Species of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services*. Tech. rep.

Jin, W.-T. et al. (2017). "Phylogenetics of subtribe Orchidinae s.l. (Orchidaceae; Orchidoideae) based on seven markers (plastid matK, psaB, rbcL, trnL-F, trnH-psba, and nuclear nrITS, Xdh): implications for generic delimitation". *BMC Plant Biol.* vol. 17, no. 1, p. 222.

Johnson, M. G., E. M. Gardner, et al. (2016). "HybPiper: Extracting coding sequence and introns for phylogenetics from high-throughput sequencing reads using target enrichment". *Appl. Plant Sci.* vol. 4, no. 7, p. 1600016.

Johnson, M. G., L. Pokorny, et al. (2019). "A Universal Probe Set for Targeted Sequencing of 353 Nuclear Genes from Any Flowering Plant Designed Using k-Medoids Clustering". *Syst. Biol.* vol. 68, no. 4, pp. 594–606.

Kasparek, M. and U. Grimm (1999). "European Trade in Turkish Salep with Special Reference to Germany". *Econ. Bot.* vol. 53, no. 4, pp. 396–406.

Kembel, S. W. et al. (2010). "Picante: R tools for integrating phylogenies and ecology". *Bioinformatics* vol. 26, no. 11, pp. 1463–1464.

Kimura, M. (1980). "A simple method for estimating evolutionary rates of base substitutions through comparative studies of nucleotide sequences". *J. Mol. Evol.* vol. 16, no. 2, pp. 111–120.

Kosakovsky Pond, S. L. et al. (2020). "HyPhy 2.5—a customizable platform for evolutionary hypothesis testing using phylogenies". *Molecular biology and evolution* vol. 37, no. 1, pp. 295–299.

Kreziou, A., H. de Boer, and B. Gravendeel (2016). "Harvesting of salep orchids in north-western Greece continues to threaten natural populations". *Oryx* vol. 50, no. 3, pp. 393–396.

Kühn, R., H. Pedersen, and P. Cribb (2019). *Field guide to the orchids of Europe and the Mediterranean*. Richmond, England: Kew Publishing.

Kurt, A. (2021). *Salep Glucomannan: Properties and Applications*.

Latorre, S. M. et al. (2020). "Isolation, Library Preparation, and Bioinformatic Analysis of Historical and Ancient Plant DNA". *Curr Protoc Plant Biol* vol. 5, no. 4, e20121.

Lavorgna, A. et al. (2018). "CITES, wild plants, and opportunities for crime". *European Journal on Criminal Policy and Research* vol. 24, no. 3, pp. 269–288.

Leão, T. C. C., D. Lobo, and L. Scotson (2017). "Economic and Biological Conditions Influence the Sustainability of Harvest of Wild Animals and Plants in Developing Countries". *Ecol. Econ.* vol. 140, pp. 14–21.

Li, H. and R. Durbin (2009). "Fast and accurate short read alignment with Burrows–Wheeler transform". *Bioinformatics* vol. 25, no. 14, pp. 1754–1760.

López-Giráldez, F. and J. P. Townsend (2011). "PhyDesign: an online application for profiling phylogenetic informativeness". *BMC Evolutionary Biology* vol. 11, no. 1, pp. 1–4.

Mandel, J. R. et al. (2014). "A target enrichment method for gathering phylogenetic information from hundreds of loci: An example from the Compositae". *Appl. Plant Sci.* vol. 2, no. 2.

Manzanilla, V. et al. (2022). "Using target capture to address conservation challenges: Population-level tracking of a globally-traded herbal medicine". *Mol. Ecol. Resour.* vol. 22, no. 1, pp. 212–224.

Margulies, J. D. et al. (2019). "Illegal wildlife trade and the persistence of "plant blindness"". *Plants People Planet* vol. 1, no. 3, pp. 173–182.

Maxim, L., J. H. Spangenberg, and M. O'Connor (2009). "An analysis of risks for biodiversity under the DPSIR framework". *Ecol. Econ.* vol. 69, no. 1, pp. 12–23.

Maxwell, S. L. et al. (2016). "Biodiversity: The ravages of guns, nets and bulldozers". *Nature* vol. 536, no. 7615, pp. 143–145.

Minh, B. Q., M. W. Hahn, and R. Lanfear (2020). "New Methods to Calculate Concordance Factors for Phylogenomic Datasets". *Mol. Biol. Evol.* vol. 37, no. 9, pp. 2727–2733.

Minh, B. Q., H. A. Schmidt, et al. (2020). "IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era". *Mol. Biol. Evol.* vol. 37, no. 5, pp. 1530–1534.

Molnár V, A. et al. (2017). "Turkish graveyards as refuges for orchids against tuber harvest". *Ecol. Evol.* vol. 7, no. 24, pp. 11257–11264.

Molnár V., A. et al. (2017). "Using traditional ecological knowledge in discovery of rare plants: a case study from Turkey". *Acta Soc. Bot. Pol. Pol. Tow. Bot.* vol. 86, no. 3.

Morton, O. et al. (2021). "Impacts of wildlife trade on terrestrial biodiversity". *Nat Ecol Evol* vol. 5, no. 4, pp. 540–548.

Nauheimer, L. et al. (2021). "HybPhaser: A workflow for the detection and phasing of hybrids in target capture data sets". *Appl. Plant Sci.* vol. 9, no. 7.

Oksanen, J. et al. (2022). *vegan: Community Ecology Package*.

Orlando, L. et al. (2021). "Ancient DNA analysis". *Nature Reviews Methods Primers* vol. 1, no. 1, pp. 1–26.

Phelps, J. and E. L. Webb (2015). ""Invisible" wildlife trades: Southeast Asia's undocumented illegal trade in wild ornamental plants". *Biol. Conserv.* vol. 186, pp. 296–305.

Phillips, S. J., R. P. Anderson, and R. E. Schapire (2006). "Maximum entropy modeling of species geographic distributions". *Ecol. Modell.* vol. 190, no. 3, pp. 231–259.

Pridgeon, A. M. et al. (2001). *Genera Orchidacearum: Volume 2. Orchidoideae (Part 1)*. Oxford University Press.
— (2003). *Genera Orchidacearum: Volume 3. Orchidoideae (Part 2), Vanilloideae*. Oxford University Press.
Quinlan, A. R. and I. M. Hall (2010). "BEDTools: a flexible suite of utilities for comparing genomic features". *Bioinformatics* vol. 26, no. 6, pp. 841–842.
Ranwez, V. et al. (2018). "MACSE v2: toolkit for the alignment of coding sequences accounting for frameshifts and stop codons". *Molecular biology and evolution* vol. 35, no. 10, pp. 2582–2584.
Roe, D. and F. Booker (2019). "Engaging local communities in tackling illegal wildlife trade: A synthesis of approaches and lessons for best practice". *Conserv. Sci. Pract.* vol. 1, no. 5, e26.
Ross, H. A., S. Murugan, and W. L. S. Li (2008). "Testing the reliability of genetic methods of species identification via simulation". *Syst. Biol.* vol. 57, no. 2, pp. 216–230.
Saslis-Lagoudakis, C. H. et al. (2012). "Phylogenies reveal predictive power of traditional medicine in bioprospecting". *Proc. Natl. Acad. Sci. U. S. A.* vol. 109, no. 39, pp. 15835–15840.
Sayyari, E. and S. Mirarab (2018). "Testing for Polytomies in Phylogenetic Species Trees Using Quartet Frequencies". *Genes* vol. 9, no. 3.
Scheffers, B. R. et al. (2019). "Global wildlife trade across the tree of life". *Science* vol. 366, no. 6461, pp. 71–76.
Schippmann, U., D. J. Leaman, and A. B. Cunningham (2002). "Impact of cultivation and gathering of medicinal plants on biodiversity: global trends and issues". *Biodiversity and the ecosystem approach in agriculture, forestry and fisheries*.
Sezik, E. (2002). "Turkish orchids and salep". *Acta Pharmaceutica Turcica* vol. 44, pp. 151–157.
Smith, M. D. et al. (2015). "Less is more: an adaptive branch-site random effects model for efficient detection of episodic diversifying selection". *Mol. Biol. Evol.* vol. 32, no. 5, pp. 1342–1353.
Springer, M. S. and J. Gatesy (2016). "The gene tree delusion". *Molecular phylogenetics and evolution* vol. 94, pp. 1–33.
Staats, M., A. J. Arulandhu, et al. (2016). "Advances in DNA metabarcoding for food and wildlife forensic species identification". *Anal. Bioanal. Chem.* vol. 408, no. 17, pp. 4615–4630.
Staats, M., A. Cuenca, et al. (2011). "DNA damage in plant herbarium tissue". *PLoS One* vol. 6, no. 12, e28448.
Teixidor-Toneu, I., F. M. Jordan, and J. A. Hawkins (2018). "Comparative phylogenetic methods and the cultural evolution of medicinal plant use". *Nat Plants* vol. 4, no. 10, pp. 754–761.
Trienekens, J. H. et al. (2012). "Transparency in complex dynamic food supply chains". *Advanced Engineering Informatics* vol. 26, no. 1, pp. 55–65.
Veldman, S. et al. (2018). "Trade in Zambian Edible Orchids—DNA Barcoding Reveals the Use of Unexpected Orchid Taxa for Chikanda". *Genes* vol. 9, no. 12, p. 595.

Wetterstrand, K. A. (2023). *DNA Sequencing Costs: Data from the NHGRI Genome Sequencing Program (GSP)*. Accessed: 2023-7-24.

Wu, S., S. Song, et al. (2013). "Reply to Gatesy and Springer: The multispecies coalescent model can effectively handle recombination and gene tree heterogeneity". *Proceedings of the National Academy of Sciences* vol. 110, no. 13, E1180–E1180.

Wu, T. D., J. Reeder, et al. (2016). "GMAP and GSNAP for Genomic Sequence Alignment: Enhancements to Speed, Accuracy, and Functionality". *Methods Mol. Biol.* vol. 1418, pp. 283–334.

Xu, B. and Z. Yang (2016). "Challenges in Species Tree Estimation Under the Multispecies Coalescent Model". *Genetics* vol. 204, no. 4, pp. 1353–1368.

Yang, Y. and S. Smith (2014). "Orthology inference in nonmodel organisms using transcriptomes and low-coverage genomes: improving accuracy and matrix occupancy for phylogenomics". *Mol. Biol. Evol.* vol. 31, no. 11, pp. 3081–3092.

Youssef, S. et al. (2019). "Wild orchids of the Kurdistan Region areas: a scientific window on the unexpected nature of the North-Western Zagros".

Zhang, C. et al. (2018). "ASTRAL-III: polynomial time species tree reconstruction from partially resolved gene trees". *BMC bioinformatics* vol. 19, no. 6, pp. 15–30.

Papers

Paper I

Wildlife trade

I

Mehrdad Jahanbanifard, Margaretha A. Veltman, Sarina Veldman, Ida Hartvig, Carly Cowell, Frederic Lens, Steven Janssens, Erik Smets

Published in: De Boer, H., Rydmark, M. O., Verstraete, B., & Gravendeel, B. (eds). *Molecular identification of plants: from sequence to species.* Advanced Books. DOI: 10.3897/ab.e98875.

Chapter 25
Wildlife trade

Mehrdad Jahanbanifard[1,2], Margaretha A. Veltman[3], Sarina Veldman[1], Ida Hartvig[4], Carly Cowell[5], Frederic Lens[1], Steven Janssens[6], Erik Smets[1]

1 Naturalis Biodiversity Center, Leiden, The Netherlands
2 Leiden Institute of Advanced Computer Science (LIACS), Leiden University, Leiden, The Netherlands
3 Natural History Museum, University of Oslo, Oslo, Norway
4 University of Copenhagen, Copenhagen, Denmark
5 Royal Botanical Garden, Kew, United Kingdom
6 Meise Botanic Garden, Meise, Belgium

Mehrdad Jahanbanifard **mehrdad.jahanbanifard@naturalis.nl**
Margaretha A. Veltman **margret.veltman@nhm.uio.no**
Sarina Veldman **sarina.veldman@naturalis.nl**
Ida Hartvig **ihla@ign.ku.dk**
Carly Cowell **c.cowell@kew.org**
Frederic Lens **frederic.lens@naturalis.nl**
Steven Janssens **steven.janssens@plantentuinmeise.be**
Erik Smets **erik.smets@naturalis.nl**

Citation: Jahanbanifard M, Veltman MA, Veldman S, Hartvig I, Cowell C, Lens F, Janssens S, Smets E (2022) Chapter 25. Wildlife trade. In: de Boer H, Rydmark MO, Verstraete B, Gravendeel B (Eds) Molecular identification of plants: from sequence to species. Advanced Books. https://doi.org/10.3897/ab.e98875

Introduction

Wildlife trade and its effects

Wildlife trade is the trading of living or dead wild plants, fungi, or animals, either as whole organisms or as parts and the products derived from them. This varies from rare animal and plant species for collectors, to ingredients made of wild organisms for medicinal or cosmetic purposes, to wood for timber, paper, craftwork, and construction, and various animals, plants, and mushrooms for nutritional purposes. Although conservation concerns about the unsustainable use of wildlife became more prominent from the 1960s onward, evidence shows that large-scale wildlife trade is older than the Roman Empire and ancient Greek civilisations ('t Sas-Rolfes et al. 2019). International wildlife trade is a billion-dollar industry, and together with illegal wildlife trafficking, it has become a substantial threat to global biodiversity and the preservation of endangered species (Smith et al. 2017). In addition, the overall impact of wildlife trade on national economies as well as public health is largely underestimated (Kurland et al. 2017; Rosen and Smith 2010).

The impacts of wildlife trade are substantial with both conservation and socio-economic importance. Unsustainable trade could lead to (local) extinction of populations or even entire species. For plants that occupy a specialised niche, it can destabilise interactions with other species, with potential consequences for the entire ecosystem. Therefore, after habitat loss, wildlife trade is the second-biggest threat to species survival (WWF, 2020). Not only does illegal wildlife trade threaten biodiversity due to consistent overexploitation, it also competes with legal use of natural resources and results in a substantial loss of income for both local communities and governments (Cooney et al. 2015). Many source countries rely on the products and/or income generated from wildlife trade, meaning that the livelihoods of the people that depend on it would be compromised if these species go extinct or if trade would be banned. In some areas in Tanzania, for example, illegal chikanda orchid gathering is the primary economic activity for vulnerable HIV/AIDS-affected households (Challe and Price 2009), although resellers further down the supply chain actually profit the most from this trade (Veldman et al. 2014). The best-known examples of wildlife trade in plants can be found in timber commerce (e.g., rosewood and ebony wood), for which the legal market has an annual value of around $200 billion and the illegal market an estimated annual $30–$157 billion (Jenkins et al. 2018; World Bank 2019). Furthermore, it is estimated that 60–90% of medicinal and aromatic plants are harvested from the wild, among which several high-value species, such as sandalwood (*Santalum* spp.), agarwood (*Aquilaria* spp.), African cherry (*Prunus africana*), and American and Chinese ginseng (*Panax* spp.) (Jenkins et al. 2018). Moreover, several groups of plants are traded for ornamental purposes, including species from threatened taxa such as cycads, cacti, aloes, conifers, euphorbs, and orchids. An overview of the global hotspots for wildlife trade, with some examples of plant groups targeted, is given in Figure 1.

Regulating wildlife trade

In order to regulate the trade in vulnerable wildlife, the Convention on International Trade of Endangered Species of Wild Fauna and Flora (CITES) was established in 1975. Species at risk of overexploitation due to international trade are listed on one of three appendices depending on how much they are threatened by unrestricted trade. Appendix I lists the most endangered species, for which commercial trade is not permitted - except for pre-convention material - and

for which non-commercial trade is strictly regulated. Appendix II lists the species that may become extinct if trade is not carefully controlled, which therefore requires a proper permit. Finally, Appendix III lists species that are protected in at least one country and other CITES Parties assistance is required to control the trade. Listing species on Appendix III helps to establish international cooperation in order to control trade in the species according to the laws and regulations of that country. Species can be added to Appendix I and II or removed from them, or shifted from Appendix I to II and vice versa only by voting at a Conference of the Parties (CoP), which is a meeting of the CITES Parties to review the implementation of the Convention. Species can be added to Appendix III or removed from it at any time and by any Party unilaterally (CITES, n.d.).

At the moment, roughly 39,000 species, including ca. 6000 species of animals and ca. 33,000 species of plants (395 species in Appendix I, 32,364 species in Appendix II, and 9 species in Appendix III) are protected by CITES (CITES, n.d.). In countries that are signatories to the convention, import and export permits must be issued for international trade of plants and animals listed in these appendices. Some countries set annual export quotas for certain species to ensure that they will not be traded beyond the sustainable limits for species survival. Non-compliance with CITES regulations can lead to confiscation of the material as well as fines and prison sentences, and in some cases trade sanctions against a country (CITES, n.d.). Since 2017, CITES has also facilitated the Wildlife Cybercrime Working Group that has coordinated national responses to the threat posed by online trade (Sajeva et al. 2013).

Other international and national regulations have been put into place to support the implementation of and in some cases expand on CITES regulations. Examples are the EU Action Plan Against Wildlife Trafficking (European Commission 2016), the EU Wildlife Trade Regulations (European Commission 2010), European Union Timber Regulation (EUTR), United States LEMIS wildlife trade data (Eskew et al. 2020), and the United States Lacey Act (Anderson 1995). Under the National Legislation Project (NLP), various domestic measures need to be implemented in order to meet the four CITES criteria, without which the CITES regulations are not in force at the national level: countries need to designate at least one Management Authority and one Scientific Authority; prohibit trade in specimens in violation of the Convention; penalise such trade; or confiscate specimens illegally traded or possessed. Diverse governmental and non-governmental programmes exist that implement enforcement in source, transit, and consumer countries, and are used to increase the risks of being involved in illegal wildlife trade as well as to decrease the rewards. In terms of global law enforcement, INTERPOL examines websites and social media posts offering wildlife products for sale. This happens annually and a number of seizures and arrests take place every year.

Challenges in combating wildlife trade

Despite the fact that plant species far outnumber animal species on the CITES appendices, in the public discourse on wildlife trade and conservation, charismatic mammals such as elephants, rhinos, tigers, and lions usually take centre stage. Smaller animals (e.g., insects, molluscs), but also most plant groups, receive less attention and generate less funding in discussions regarding wildlife trade and conservation. And although plants appear frequently in national and international regulations, regulatory enforcement and additional conservation measures still primarily target iconic megafauna (Margulies et al. 2019). The relative 'invisibility' of plants as organisms of importance for our lives and worthy of conservation is called "plant blindness", and is one of the biggest challenges in combating illegal plant trade (Box 1).

Chapter 25: Box 1. Example of a challenge in depth: plant blindness

Plant blindness is a psychological bias that leads us to notice (large) animals, and take plants largely for granted, reducing them to background vegetation for other organisms. The term was coined by Wandersee and Schussler (1999) and refers to a number of common problems in the perception of plants: not noticing plants in one's environment; ignoring plants' aesthetic and unique biological features; not recognising the importance of plants (e.g., food production, absorbing carbon dioxide and releasing oxygen, etc.); and considering plants as inferior to animals. Plant blindness has both a physical and a psychological component. The human eye picks-up the colour green more easily than other colours, and hence does not focus on it quite as much (Knapp 2019). Green is also experienced as safe and therefore warrants limited attention. Furthermore, our eyes perceive movements more readily than static objects, which probably stems from an evolutionary function in spotting (attacking) predators and (fleeing) prey.

Plant blindness has been institutionalised throughout society, from (higher) education to governance and wildlife management (Margulies et al. 2019; Wandersee and Schussler 1999), leading to a focus on animals in biology courses, natural history museums, research funding, and conservation policies. Plant blindness is therefore one of the biggest challenges in combating illegal wildlife trade.

Apart from the limited attention that plants receive in research, education, and conservation, effective control of trade in plant species is hampered because some of the traded goods are difficult to recognise, either because they are processed or because they contain only parts of the organism, which lack the morphological characters needed for identification (Lavorgna et al. 2018). Plant products are therefore often harder to identify than living animals or animal parts, and to identify them routinely requires standardised and scalable technologies, many of which are still being developed (for more details, see Methods).

Other challenges are posed by the growing use of the internet for transactions, which makes wildlife material more readily accessible and at lower costs, while preserving anonymity. The internet is not only increasingly used to sell and obtain specimens, but even to organise poaching events (Lavorgna 2014). Rare and exotic plant species can be ordered with ease from a range of online retailers, shipping of plants in the postal system is relatively easy and the search for plant material in these systems is limited. In addition, the scale of the internet and speed at which online marketplaces proliferate make the monitoring of online criminal activities costly and time consuming (Lavorgna et al. 2020; TRAFFIC, 2019). The online market thus facilitates participation in illegal wildlife trade, making it more attractive due to potentially high sales and profits and reduced detection rate (TRAFFIC, 2019). The challenges for curbing illegal online trade are therefore manifold, and only exacerbate existing challenges with law enforcement by enabling covert activities and thereby increasing the volume of illegally traded goods. Distinguishing legal from illegal trade is difficult even with specialist knowledge or extensive training (Vaglica et al. 2017). Mixing legal and illegal shipments, nontransparent supply chains and lack of institutional monitoring capacity in biodiversity rich countries are some of the practical challenges underpinning this difficulty (Engler and Parry-Jones 2007). International conventions such as CITES can also have unintended loopholes that allow wildlife traffickers to circumvent restrictions or to present their information in a way that gives the impression of legal trade. For example, newly discovered rare species that have not yet made their way onto one of the CITES appendices can often be traded freely, despite detrimental effects, if there is no national legislation in place to protect the species. Another commonly observed practice is the export of wild harvested or poached wildlife as captive bred (in the case of animals) or artificially propagated

(in the case of plants) organisms. Verification of legal acquisition can be challenging without sufficient documentation, opening up space for laundering of illegally obtained specimens.

Lastly, since international wildlife trade per definition transcends borders, enforcement of legal trade requires coordinated action between multiple countries to address the whole supply chain. While there are already many institutional collaborations that work across international borders to help track and catch illegal wildlife trafficking syndicates - including financial institutions, NGOs, customs and police forces and online tech platforms - one of the main bottlenecks to combating wildlife trade will be to sustain sufficient international attention to allow the detection and prevention, not just of single illegal transactions, but of organised trade networks operating at larger scales.

The importance of wildlife and the impacts of unsustainable trade on biodiversity are undeniable, which highlights the urgency of developing high-throughput methods that are widely applicable. The next section presents some of the most commonly used methods in illegal trade identification today. In the final section, we provide recommendations on which techniques to use for the identification and tracking of illegally traded plants, and discuss future developments that could improve global wildlife trade monitoring and control.

Methods for identification of plants in trade

Traded plant materials come in all shapes and sizes and in different stages of processing, ranging from complete living plants to raw timber logs and to engineered wood products. There is a wide variety of molecular and non-molecular methods for illegal wildlife trade monitoring, from DNA (meta) barcoding and genetic methods, to chemical identification, and computer vision and pattern recognition tools. Each of these methods is applicable to certain types of materials and requires knowledge about different aspects of the traded product that determines its legality, including species identity, geographic origin, source population (wild or cultivated), and the sample age. Here we describe the most commonly used methods to identify each of these aspects, and why they are important.

Species identity

Methods for species identification are used to ascertain whether the organism being traded is CITES-listed or not. Depending on the taxonomic rank that is listed, it may be necessary to identify the exact species (e.g., *Panax ginseng*), genus (e.g., *Aloe* spp.), or family (e.g., Orchidaceae) to which an organism belongs. Species identification methods include genetic based methods (based on DNA sequencing information), chemical methods (based on molecular mass spectra), and computational methods (based on image recognition). Each of these methods require suitable reference data against which to query an unknown sample. The availability of reference data and the nature of the sample will dictate which method is most suitable for species identification.

Mass spectrometry

The main chemical method used to identify species is Direct Analysis in Real Time (DART) coupled with time-of-flight (TOF) mass spectrometry (DART-TOF MS). DART-TOF MS consists of two parts: DART is an ionisation source that ionises ambient atmospheric molecules by using electronically excited-state helium which reacts with the molecules in the investigated sample to produce analyte ions (Gross 2014). These ions are then sucked into the AccuTOF mass spec-

trometer. Spectral data on molecular masses and their relative intensities (so called chemical fingerprint) can be analysed to identify timbers (Deklerck et al. 2020; Evans et al. 2017; Lancaster and Espinoza 2012), keratin fibres of camelids (Price et al. 2020), rhinoceros keratin (Price et al. 2018), explosives (Lennert and Bridge 2018), and narcotics (Lian et al. 2017). DART-TOF MS is fast and has a simple sample preparation procedure. The accuracy of the result is however dependent on the reference database - as is the case for all other species identification methods - and whether the investigated samples have enough variation in molecular composition to be distinguished with their chemotype (Deklerck et al. 2017).

Computer vision and pattern recognition

Thanks to machine learning and computer vision, expert systems are playing an increasingly important role in identification of a wide variety of wildlife related objects, such as medicinal leaves (Sabu et al. 2017), herbarium specimens (Lorieul et al. 2019; Pearson et al. 2020), wood identification (Lens et al. 2020), mulberry ripeness detection (Ashtiani et al. 2021), pollen grains (Polling et al. 2021), corn seed varieties detection (Javanmardi et al. 2021) and wildlife monitoring (Di Minin et al. 2019, 2018). The concept of this method is pretty simple: train a model using a reliable database (usually an image database) to recognise specific objects such as humans, cars, trees, etc, in an image that the model has not seen before. Not only images (e.g., light microscopic images) can be used as input data, but also Near infrared (NIR) spectroscopy and X-ray micro computed tomography (CT) data can be used for automated pattern recognition. These are nondestructive alternative methods that can be useful when the conventional methods (such as light microscopy or DNA-based methods) are not acceptable or difficult to use, as is often the case in the investigation of registered cultural objects (Kobayashi et al. 2019). The main advantage of using computer vision methods is that it is accurate and applicable on a wide range of materials, such as wood, leaves, flowers, and pollen grains. The main drawback of computer vision, apart from a general lack of reliable databases, is the insufficient resolution of many morphological traits for species recognition, especially amongst closely related species. In some cases, better algorithms, more powerful machines, and high-quality reference databases can mitigate this challenge. However, in the cases where morphological traits do not provide distinctive features, pattern recognition cannot be used.

DNA barcoding and metabarcoding

DNA-based identification methods can use different genomic markers that offer different levels of identification, from universal loci such as conserved genes or intergenic spacers, to neutrally evolving markers with sufficient variation to resolve specific taxa, such as microsatellites and genome-wide Single Nucleotide Polymorphisms (SNPs). In addition to these markers, which require information about genomic context, it is also possible to identify species and populations using alignment-free shotgun data (see Chapter 17 Species delimitation).

For species identification, DNA barcoding (see Chapter 10 DNA barcoding) is often the method of choice. It can effectively identify traded plant species in a number of cases, including the identification of rosewood (*Dalbergia* spp.), species used in Ayurvedic medicine (*Decalepis* spp.), and cycads (*Encephalartos* spp.) (Hartvig et al. 2015; Mishra et al. 2017; Williamson et al. 2016). In addition, DNA metabarcoding (see Chapter 11 Amplicon metabarcoding) detects multiple species in mixed products such as traditional medicine and processed foods (Arulandhu et al. 2017; Veldman et al. 2017). An advantage of DNA barcoding is that, for the core land plant barcodes such as *rbcL*, *matK*, and nrITS, reference data is readily and freely available in public databases such as NCBI's GenBank or BOLD (barcodinglife.org). Tropical species are generally under-represented in these databases, and NCBI GenBank is known to contain er-

roneous sequences due to limited quality control. Species-level discrimination using standard barcodes has proven to be difficult among closely related and hybridising species, as well as taxa with low rates of evolution (Hassold et al. 2016; Veldman et al. 2017). An alternative in these cases is to develop custom barcodes. This provides researchers with more control over choosing genomic features that are informative for their plant group, but requires generating novel reference data, raising both the financial costs and time investment.

Source population and geographic origin

Neutral genetic markers

An advantage of DNA barcoding is that the sequence data is universally comparable among labs and large numbers of species. But since DNA barcoding was originally meant to distinguish between species and not within species, this method often falls short when higher resolution is needed. Identification below the species level may be useful if the legality of trade is determined by the source population. In some cases, the country of origin determines the legal status of traded plants, which requires population level data for a collection of reference samples spanning the species range. Cost-effective traditional population genetic methods use a number of species-specific variable markers, typically simple sequence repeats (SSRs) or inter simple sequence repeats (ISSRs), which can be highly variable and show fine-grained population structure. More recently developed high-throughput sequencing methods cover larger sections of the genome, such as reduced representation sequencing methods (RAD-seq, target capture, or low coverage whole genome shotgun sequencing (also known as genome skimming, see Chapter 16 Whole genome sequencing).

These methods can generate large numbers of SNPs that allow inference of geographic origins at various scales. Although the increased costs for library preparation and sequencing means that these methods are not economically feasible in all cases, they offer the added advantage that functional analyses of genes or markers linked to genes with adaptive significance is possible.

Geographic origins have even been identified at the level of continents using genome skimming (Schroeder et al. 2016), at the level of countries with SNPs generated by target enrichment of nuclear loci (Manzanilla et al. 2022) and RAD-seq (Blanc-Jolivet et al. 2017; Pakull et al. 2020), and even at the level of individual forest concessions with microsatellites (Vlam et al. 2018). Population genetic methods could potentially also be useful in detecting laundering of illegally harvested plants that are claimed to be cultivated. Genetic diversity analysis of the same neutral markers that are used to infer geographic origin, could then point out whether the plants were indeed sourced from a particular plantation or rather from the wild - in which case their genetic composition would be much more diverse than expected from artificially propagated material.

Stable isotope analysis

While population genetic markers can offer unmatched resolution of spatial variation, a general disadvantage is that many of them (with the exception of those used in RAD-seq and shotgun sequencing) need to be tested or developed specifically for each species, and reference data must be generated for populations across the distribution range to be tested. Stable isotope analysis can also infer geographic origin of samples, and does not depend on species-specific reference data to the same extent as genetic methods do. Stable isotope analysis is based on the principle that the presence of stable isotopes in the environment depends on both climate and geography. This creates a correlation between the stable isotope profile and its geographic location (Hermes et al. 2018). Since plants generally incorporate the stable isotopes into their

tissue at the same ratios as they occur in their environment, stable isotope analysis of plant material can be used to infer its geographic origin and be a tool in wildlife forensics (Matos and Jackson 2019). Stable isotope analysis however does not have a geographic resolution as high as population genetic methods have (Gori et al. 2015; Horacek et al. 2009). Georeferenced data is also required for stable isotope analysis, and global isotope databases are currently not freely available yet (Camin et al. 2017), limiting broad application of this method.

Harvesting pre- or post CITES legislation

Radiocarbon dating

There are two methods to measure radiocarbon abundance: radiometric dating and accelerator mass spectrometry (AMS). These methods can be used to date samples based on the decay of carbon isotopes. The estimated age gives an indication of whether or not the traded sample is a pre-convention material, meaning that the traded material predates the convention or listing of the species (e.g., Kalt-O'Bannon 1994; Uno et al. 2013; Cerling et al. 2016). While both radiometric dating and AMS provide high quality results, they are fundamentally different. AMS quantifies the number of carbon 14 (^{14}C) atoms in the investigated samples, while radiometric dating methods are based on the detection of beta particles resulting from the ^{14}C decay. AMS requires a much smaller sample size (20–500 mg) compared with radiometric methods (10–100 g). AMS is also faster and usually gains higher precision results than radiometric methods. Samples can be analysed in a few hours with AMS, while it can take one or two days with radiometric methods.

Recommendations to improve wildlife trade monitoring

Currently, no genetic methods for inferring sample age can compete with radiocarbon dating, and while DNA fragment sizes tend to be shorter for older and more degraded plant tissues, this alone cannot be used to determine the plant age (see Chapter 2 DNA from museum collections). For other purposes, genetic markers are the method of choice to infer species identity and geographic origin, whenever DNA extraction is a realistic option. Any genetic method will however be limited by the quality and quantity of DNA that can be extracted, which can be notoriously difficult for some materials, especially timber and processed products (Jiao et al. 2020; Lo and Shaw 2018). The obtained DNA quality and quantity will influence the range of techniques that can be applied downstream. High-copy regions such as chloroplast markers or nuclear ITS, for example, are easier to retrieve from samples with highly degraded DNA than low copy nuclear markers. For applications that require broader genomic coverage, amplification of low copy nuclear target regions can be achieved even with highly fragmented DNA, making target capture preferable over untargeted RAD-seq or genome-wide shotgun sequencing for degraded samples. However, for fresher material RAD-seq or WGS libraries may be easier to prepare and require less time for the bioinformatic analyses needed to develop markers prior to sequencing.

Despite significant progress in methods and computational analyses, applications for most methods are still limited by the lack or incompleteness of suitable reference data. As shown in Table 1, reference databases are currently under development or need further development for nearly all the methods currently used. The ForeST database for CITES protected timbers, the U.S. Fish & Wildlife Service Forensics Laboratory (Ashland, Oregon, USA), CITESwoodID by the

Thünen Institute (Hamburg, Germany), and the ebony wood microscopic database (Jahanbanifard et al. 2020, 2019) are examples of ongoing projects that are developing databases for the identification of CITES protected species.

When one method lacks sufficient reference data or is not sensitive enough to infer species identity or population of origin, multiple identification techniques tools (e.g., DNA barcoding, machine learning, and DART-TOF MS) can be combined to improve identification accuracy. Developing an integrated identification framework, which links reference databases and connects multiple sources of data for taxa of interest, is expected to play a major role in the future of regulating wildlife trade, though this would rely on standardisation and equitable distribution to enforcement agencies around the world. Coupled with new technologies that ensure quality control and compliance across the supply chain of wildlife products, the tools available for wildlife trade monitoring can aid not just the detection and confiscation of illegally traded goods, but also the transparency and traceability of legally traded commodities.

With blockchain for example, it may eventually be possible to develop a secure and robust infrastructure to register and track wildlife-related products from source to destination (Chang et al. 2020; Pournader et al. 2020). A blockchain is a database, consisting of several distributed nodes called blocks that are connected to one another using cryptography. Each block contains a cryptographic hash of the previous block, a timestamp, and transaction data (Narayanan et al. 2016). Blockchain provides an immutable and decentralised network which increases its reliability and security as no single party has full control of the system and no one can manipulate the transactions (Aimin and Yunfeng 2019; Saurabh and Dey 2021; Zheng et al. 2020).

The technology has already proven its relevance in agriculture and fisheries, where the WWF Blockchain Tuna Project demonstrates it is possible to track the history of a fishing product from ocean to plate with just a QR Code (WWF, 2018). The customisable and scalable features of blockchain make it a promising technology for application to traded timber and other wildlife-related products (MoonX, 2019). Once it is possible to keep track of all steps taken throughout the commercialisation of wild harvested plants, the checkpoints for identification will no longer be restricted to points of entry or sales, enabling monitoring of wildlife trade from the source.

Table 1. A comparison of the methods used for identifying plants in trade with an indication of their applications and limitations.

	DNA (meta) barcoding	Population genetic markers	Computer vision and pattern recognition	DART-TOF MS	AMS/ ^{14}C dating	Stable isotope
Material input	Whole plants, organs, tissues, powder	Whole plants, organs, tissues, powder	Timber, leaves, flowers, pollen	All	Anything containing organic matter	Anything containing organic matter
Purpose of application	Determine taxonomic identity from genus to species level	Determine population or region of origin	Determine taxonomic identify, from genus to (sometimes) species level	Determine taxonomic identity at species level	Determine age of material	Determine the region of origin
Availability of reference data	Well-developed for temperate species, less for tropical species and regions	Needs to be developed and referenced for each species separately	Being developed for CITES protected timber and plants	Being developed for CITES protected timber	Calibration might be required depending on the sample	Needs to be developed for each region separately

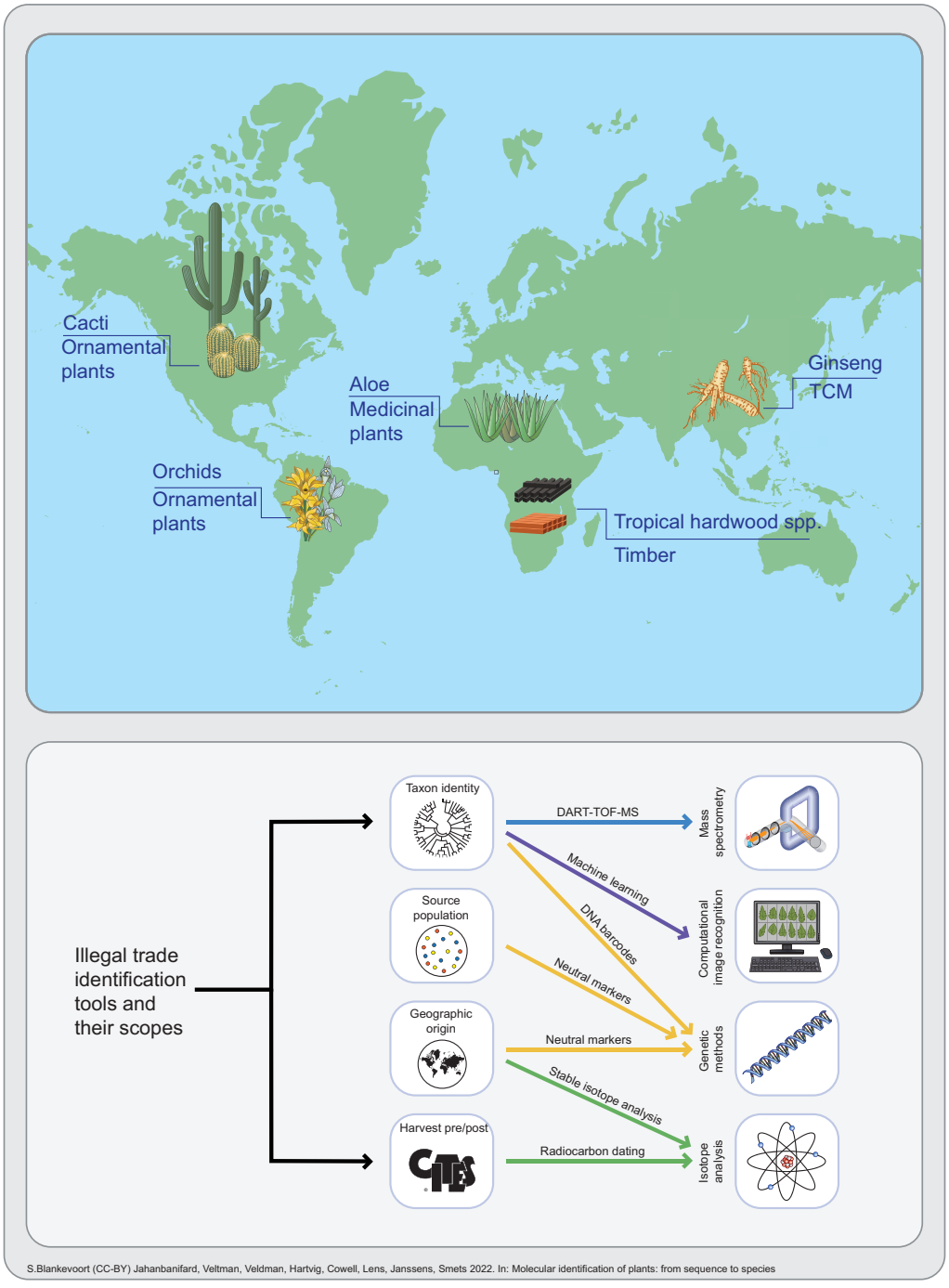

Figure 1. Chapter 25 Infographic: Global wildlife trade hotspots and some examples of traded plants from those areas, and their respective uses (ornamental, medicinal, or timber).

Questions

1. Customs officers often come across cultural heritage such as sculptures made from economically costly, legally protected wood (such as Brazilian rosewood). Which method could they use to find out whether the sculpture is made from CITES-listed species? Motivate your answer.
2. What is "plant blindness" and why is it hampering the battle against illegal plant trade?
3. Provide two advantages of AMS over radiometric dating when investigating illegal wildlife trade. Motivate your answer.

Glossary

Accelerator Mass Spectrometry (AMS) – A form of mass spectrometry that accelerates ions to extraordinarily high kinetic energies before mass analysis.
Ayurvedic medicine – A medical system from India that aims to cleanse the body and to restore balance to the body, mind, and spirit by using diet, herbal medicines, exercise, meditation, breathing, physical therapy, and other methods.
Blockchain – A decentralised and distributed network that is used to record transactions across many computers.
Computer vision – An interdisciplinary scientific field that deals with how computers can gain high-level understanding from digital images or videos.
Expert systems – In artificial intelligence, an expert system is a computer system emulating the decision-making ability of a human expert.
Inter-simple sequence repeats (ISSRs) – ISSRs are regions in the genome flanked by microsatellite sequences. PCR amplification of these regions using a single primer yields multiple amplification products that can be used as a dominant multilocus marker system for the study of genetic variation in various organisms.
Near infrared spectroscopy – A spectroscopic method that uses a certain range of the electromagnetic spectrum from 780 nm to 2500 nm which is called the near infrared region.
Pattern recognition – The automated recognition of patterns and regularities in data.
Restriction site Associated DNA Sequencing (RAD-Seq) – A fractional genome sequencing strategy, designed to interrogate anywhere from 0.1% to 10% of a selected genome.
Simple sequence repeats (SSRs) – SSRs are DNA tracts in which a short base-pair motif is repeated several to many times in tandem. These sequences experience frequent mutations that alter the number of repeats.
Spectroscopy – The study of the interaction between matter and electromagnetic radiation as a function of the wavelength or frequency of the radiation.
X-ray microtomography – A 3D modelling method uses X-rays to create cross-sections of a physical object that can be used to recreate a virtual model without destroying the original object.

References

't Sas-Rolfes M, Challender DWS, Hinsley A, Veríssimo D, Milner-Gulland EJ (2019) Illegal wildlife trade: patterns, processes, and governance. Annu. Rev. Environ. Resour. 44, 201–228. https://doi.org/10.1146/annurev-environ-101718-033253

Aimin D, Yunfeng L (2019) "Intelligent factoring" business model and game analysis in the supply chain based on block chain. Management Review 31, 231–240.

Anderson RS (1995) The Lacey Act: America's premier weapon in the fight against unlawful wildlife trafficking. Public Land Law Review, United States.

Arulandhu AJ, Staats M, Hagelaar R, Voorhuijzen MM, Prins TW, Scholtens I, Costessi A, Duijsings D, Rechenmann F, Gaspar FB, Barreto Crespo MT, Holst-Jensen A, Birck M, Burns M, Haynes E, Hochegger R, Klingl A, Lundberg L, Natale C, Niekamp H, Kok E (2017) Development and validation of a multi-locus DNA metabarcoding method to identify endangered species in complex samples. Gigascience 6, 1–18. https://doi.org/10.1093/gigascience/gix080

Ashtiani S-HM, Javanmardi S, Jahanbanifard M, Martynenko A, Verbeek FJ (2021) Detection of mulberry ripeness stages using deep learning models. IEEE Access 1–1. https://doi.org/10.1109/ACCESS.2021.3096550

Blanc-Jolivet C, Kersten B, Daïnou K, Hardy O, Guichoux E, Delcamp A, Degen B (2017) Development of nuclear SNP markers for genetic tracking of Iroko, *Milicia excelsa* and *Milicia regia*. Conserv. Genet. Resour. 1–3. https://doi.org/10.1007/s12686-017-0716-2

Camin F, Boner M, Bontempo L, Fauhl-Hassek C, Kelly SD, Riedl J, Rossmann A (2017) Stable isotope techniques for verifying the declared geographical origin of food in legal cases. Trends Food Sci. Technol. 61, 176–187. https://doi.org/10.1016/j.tifs.2016.12.007

Cerling TE, Barnette JE, Chesson LA, Douglas-Hamilton I, Gobush KS, Uno KT, Wasser SK, Xu X (2016) Radiocarbon dating of seized ivory confirms rapid decline in African elephant populations and provides insight into illegal trade. Proc Natl Acad Sci USA 113, 13330–13335. https://doi.org/10.1073/pnas.1614938113

Challe JFX, Price LL (2009) Endangered edible orchids and vulnerable gatherers in the context of HIV/AIDS in the southern highlands of Tanzania. J. Ethnobiol. Ethnomed. 5, 41. https://doi.org/10.1186/1746-4269-5-41

Chang Y, Iakovou E, Shi W (2020) Blockchain in global supply chains and cross border trade: a critical synthesis of the state-of-the-art, challenges and opportunities. International Journal of Production Research 58, 2082–2099. https://doi.org/10.1080/00207543.2019.1651946

CITES ([n.d.] 2022a) The CITES species [WWW Document]. URL https://cites.org/eng/disc/species.php (accessed 5.25.22a).

CITES ([n.d.] 2022b) Convention on International Trade in Endangered Species of Wild Fauna and Flora [WWW Document]. URL https://cites.org/eng/disc/text.php (accessed 5.25.22b).

Cooney R, Kasterine A, Macmillan DC, Milledge SAH, Nossal K, Roe D, 't Sas-Rolfes M (2015) The trade in wildlife: a framework to improve biodiversity and livelihood outcomes. IUCN, Geneva.

Deklerck V, Finch K, Gasson P, Van den Bulcke J, Van Acker J, Beeckman H, Espinoza E (2017) Comparison of species classification models of mass spectrometry data: kernel discriminant analysis vs random forest; A case study of Afrormosia (*Pericopsis elata* (Harms) Meeuwen). Rapid Commun. Mass Spectrom. 31, 1582–1588. https://doi.org/10.1002/rcm.7939

Deklerck V, Lancaster CA, Van Acker J, Espinoza EO, Van den Bulcke J, Beeckman H (2020) Chemical fingerprinting of wood sampled along a pith-to-bark gradient for individual comparison and provenance identification. Forests 11, 107. https://doi.org/10.3390/f11010107

Di Minin E, Fink C, Hiippala T, Tenkanen H (2019) A framework for investigating illegal wildlife trade on social media with machine learning. Conserv. Biol. 33, 210–213. https://doi.org/10.1111/cobi.13104

Di Minin E, Fink C, Tenkanen H, Hiippala T (2018) Machine learning for tracking illegal wildlife trade on social media. Nat. Ecol. Evol. 2, 406–407. https://doi.org/10.1038/s41559-018-0466-x

Engler, M, Parry-Jones, R (2007) Opportunity or threat: the role of the European Union in global wildlife trade. TRAFFIC.

Eskew EA, White AM, Ross N, Smith KM, Smith KF, Rodríguez JP, Zambrana-Torrelio C, Karesh WB, Daszak P (2020) United States wildlife and wildlife product imports from 2000-2014. Sci. Data 7, 22. https://doi.org/10.1038/s41597-020-0354-5

European Commission (2010) Wildlife trade regulations in the European Union. European Union.

European Commission (2016) EU action plan against wildlife trafficking, COM/2016/087.

Evans PD, Mundo IA, Wiemann MC, Chavarria GD, McClure PJ, Voin D, Espinoza EO (2017) Identification of selected CITES-protected Araucariaceae using DART TOFMS. IAWA J. 38, 266–281. https://doi.org/10.1163/22941932-20170171

Gori Y, Wehrens R, La Porta N, Camin F (2015) Oxygen and hydrogen stable isotope ratios of bulk needles reveal the geographic origin of Norway spruce in the European Alps. PLoS ONE 10, e0118941. https://doi.org/10.1371/journal.pone.0118941

Gross JH (2014) Direct analysis in real time--a critical review on DART-MS. Anal. Bioanal. Chem. 406, 63–80. https://doi.org/10.1007/s00216-013-7316-0

Hartvig I, Czako M, Kjær ED, Nielsen LR, Theilade I (2015) The use of DNA barcoding in identification and conservation of rosewood (*Dalbergia* spp.). PLoS ONE 10, e0138231. https://doi.org/10.1371/journal.pone.0138231

Hassold S, Lowry PP, Bauert MR, Razafintsalama A, Ramamonjisoa L, Widmer A (2016) DNA barcoding of Malagasy rosewoods: towards a molecular identification of CITES-listed *Dalbergia* species. PLoS ONE 11, e0157881. https://doi.org/10.1371/journal.pone.0157881

Hermes TR, Frachetti MD, Bullion EA, Maksudov F, Mustafokulov S, Makarewicz CA (2018) Urban and nomadic isotopic niches reveal dietary connectivities along Central Asia's Silk Roads. Sci. Rep. 8, 5177. https://doi.org/10.1038/s41598-018-22995-2

Horacek M, Jakusch M, Krehan H (2009) Control of origin of larch wood: discrimination between European (Austrian) and Siberian origin by stable isotope analysis. Rapid Commun. Mass Spectrom. 23, 3688–3692. https://doi.org/10.1002/rcm.4309

Jahanbanifard M, Beckers V, Koch G, Beeckman H, Gravendeel B, Verbeek F, Baas P, Priester C, Lens F (2020) Description and evolution of wood anatomical characters in the ebony wood genus *Diospyros* and its close relatives (Ebenaceae): a first step towards combatting illegal logging. IAWA 41, 577–619. https://doi.org/10.1163/22941932-bja10040

Jahanbanifard M, Gravendeel B, Lens F, Verbeek F (2019) Ebony wood identification to battle illegal trade. BISS 3. https://doi.org/10.3897/biss.3.37084

Javanmardi S, Miraei Ashtiani S-H, Verbeek FJ, Martynenko A (2021) Computer-vision classification of corn seed varieties using deep convolutional neural network. J. Stored Prod. Res. 92, 101800. https://doi.org/10.1016/j.jspr.2021.101800

Jenkins M, Timoshyna A, Cornthwaite M (2018) Wild at home: an overview of the harvest and trade in wild plant ingredients. TRAFFIC, Cambridge, UK.

Jiao L, Lu Y, He T, Guo J, Yin Y (2020) DNA barcoding for wood identification: global review of the last decade and future perspective. IAWA 41, 620–643. https://doi.org/10.1163/22941932-bja10041

Kalt-O'Bannon J (1994) Scientific dating and the law: establishing the age of old objects for legal purposes. UMKC L. Rev. 63, 93–132.

Knapp S (2019) Are humans really blind to plants? Plants, People, Planet 1, 164–168. https://doi.org/10.1002/ppp3.36

Kobayashi K, Hwang S-W, Okochi T, Lee W-H, Sugiyama J (2019) Non-destructive method for wood identification using conventional X-ray computed tomography data. Journal of Cultural Heritage 1. https://doi.org/10.1016/j.culher.2019.02.001

Kurland J, Pires SF, McFann SC, Moreto WD (2017) Wildlife crime: a conceptual integration, literature review, and methodological critique. Crime Sci. 6, 4. https://doi.org/10.1186/s40163-017-0066-0

Lancaster C, Espinoza E (2012) Analysis of select *Dalbergia* and trade timber using direct analysis in real time and time-of-flight mass spectrometry for CITES enforcement. Rapid Commun. Mass Spectrom. 26, 1147–1156. https://doi.org/10.1002/rcm.6215

Lavorgna A, Middleton SE, Pickering B, Neumann G (2020) FloraGuard: tackling the online illegal trade in endangered plants through a cross-disciplinary ICT-enabled methodology. J. Contemp. Crim. Justice 36, 428–450. https://doi.org/10.1177/1043986220910297

Lavorgna A, Rutherford C, Vaglica V, Smith MJ, Sajeva M (2018) CITES, wild plants, and opportunities for crime. Eur. J. Crim. Pol. Res. 24, 269–288. https://doi.org/10.1007/s10610-017-9354-1

Lavorgna A (2014) Wildlife trafficking in the Internet age. Crime Sci. 3, 5. https://doi.org/10.1186/s40163-014-0005-2

Lennert E, Bridge C (2018) Analysis and classification of smokeless powders by GC-MS and DART-TOFMS. Forensic Sci. Int. 292, 11–22. https://doi.org/10.1016/j.forsciint.2018.09.003

Lens F, Liang C, Guo Y, Tang X, Jahanbanifard M, da Silva FSC, Ceccantini G, Verbeek FJ (2020) Computer-assisted timber identification based on features extracted from microscopic wood sections. IAWA 1–21. https://doi.org/10.1163/22941932-bja10029

Lian R, Wu Z, Lv X, Rao Y, Li H, Li J, Wang R, Ni C, Zhang Y (2017) Rapid screening of abused drugs by direct analysis in real time (DART) coupled to time-of-flight mass spectrometry (TOF-MS) combined with ion mobility spectrometry (IMS). Forensic Sci. Int. 279, 268–280. https://doi.org/10.1016/j.forsciint.2017.07.010

Lo Y-T, Shaw P-C (2018) DNA-based techniques for authentication of processed food and food supplements. Food Chem. 240, 767–774. https://doi.org/10.1016/j.foodchem.2017.08.022

Lorieul T, Pearson KD, Ellwood ER, Goëau H, Molino J-F, Sweeney PW, Yost JM, Sachs J, Mata-Montero E, Nelson G, Soltis PS, Bonnet P, Joly A (2019) Toward a large-scale and deep phenological stage annotation of herbarium specimens: case studies from temperate, tropical, and equatorial floras. Appl. Plant Sci. 7, e01233. https://doi.org/10.1002/aps3.1233

Manzanilla V, Teixidor-Toneu I, Martin GJ, Hollingsworth PM, de Boer HJ, Kool A (2022) Using target capture to address conservation challenges: population-level tracking of a globally-traded herbal medicine. Mol. Ecol. Resour. 22, 212–224. https://doi.org/10.1111/1755-0998.13472

Margulies JD, Bullough L, Hinsley A, Ingram DJ, Cowell C, Goettsch B, Klitgård BB, Lavorgna A, Sinovas P, Phelps J (2019) Illegal wildlife trade and the persistence of "plant blindness." Plants, People, Planet 1, 173–182. https://doi.org/10.1002/ppp3.10053

Matos MPV, Jackson GP (2019) Isotope ratio mass spectrometry in forensic science applications. Forensic Chemistry 13, 100154. https://doi.org/10.1016/j.forc.2019.100154

Mishra P, Kumar A, Sivaraman G, Shukla AK, Kaliamoorthy R, Slater A, Velusamy S (2017) Character-based DNA barcoding for authentication and conservation of IUCN Red listed threatened species of genus *Decalepis* (Apocynaceae). Sci. Rep. 7, 14910. https://doi.org/10.1038/s41598-017-14887-8

MoonX (2019) Blockchain – Future of wildlife conservation. Medium.

Narayanan A, Bonneau J, Felten E, Miller A, Goldfeder S (2016) Bitcoin and cryptocurrency technologies. Network Security 2016, 4. https://doi.org/10.1016/S1353-4858(16)30074-5

Pakull B, Schindler L, Mader M, Kersten B, Blanc-Jolivet C, Paulini M, Lemes MR, Ward SE, Navarro CM, Cavers S, Sebbenn AM, di Dio O, Guichoux E, Degen B (2020) Development of nuclear SNP markers for mahogany (*Swietenia* spp.). Conserv. Genet. Resour. https://doi.org/10.1007/s12686-020-01162-8

Pearson KD, Nelson G, Aronson MFJ, Bonnet P, Brenskelle L, Davis CC, Denny EG, Ellwood ER, Goëau H, Heberling JM, Joly A, Lorieul T, Mazer SJ, Meineke EK, Stucky BJ, Sweeney P, White AE, Soltis PS (2020) Machine learning using digitized herbarium specimens to advance phenological research. Bioscience 70, 610–620. https://doi.org/10.1093/biosci/biaa044

Polling M, Li C, Cao L, Verbeek F, de Weger LA, Belmonte J, De Linares C, Willemse J, de Boer H, Gravendeel B (2021) Neural networks for increased accuracy of allergenic pollen monitoring. Sci. Rep. 11, 11357. https://doi.org/10.1038/s41598-021-90433-x

Pournader M, Shi Y, Seuring S, Koh SCL (2020) Blockchain applications in supply chains, transport and logistics: a systematic review of the literature. International Journal of Production Research 58, 2063–2081. https://doi.org/10.1080/00207543.2019.1650976

Price E, Larrabure D, Gonzales B, McClure P, Espinoza E (2020) Forensic identification of the keratin fibers of South American camelids by ambient ionization mass spectrometry: Vicuña, alpaca and guanaco. Rapid Commun. Mass Spectrom. 34, e8916. https://doi.org/10.1002/rcm.8916

Price ER, McClure PJ, Jacobs RL, Espinoza EO (2018) Identification of rhinoceros keratin using direct analysis in real time time-of-flight mass spectrometry and multivariate statistical analysis. Rapid Commun. Mass Spectrom. 32, 2106–2112. https://doi.org/10.1002/rcm.8285

Rosen GE, Smith KF (2010) Summarizing the evidence on the international trade in illegal wildlife. Ecohealth 7, 24–32. https://doi.org/10.1007/s10393-010-0317-y

Sabu A, Sreekumar K, Nair RR (2017) Recognition of ayurvedic medicinal plants from leaves: a computer vision approach, in: 2017 Fourth International Conference on Image Information Processing (ICIIP). Presented at the 2017 Fourth International Conference on Image Information Processing (ICIIP), IEEE, pp. 1–5. https://doi.org/10.1109/ICIIP.2017.8313782

Sajeva M, Augugliaro C, Smith MJ, Oddo E (2013) Regulating internet trade in CITES species. Conserv. Biol. 27, 429–430. https://doi.org/10.1111/cobi.12019

Saurabh S, Dey K (2021) Blockchain technology adoption, architecture, and sustainable agri-food supply chains. J. Clean. Prod. 284, 124731. https://doi.org/10.1016/j.jclepro.2020.124731

Schroeder H, Cronn R, Yanbaev Y, Jennings T, Mader M, Degen B, Kersten B (2016) Development of molecular markers for determining continental origin of wood from white oaks (*Quercus* L. sect. *Quercus*). PLoS ONE 11, e0158221. https://doi.org/10.1371/journal.pone.0158221

Smith KM, Zambrana-Torrelio C, White A, Asmussen M, Machalaba C, Kennedy S, Lopez K, Wolf TM, Daszak P, Travis DA, Karesh WB (2017) Summarizing US wildlife trade with an eye toward assessing the risk of infectious disease introduction. Ecohealth 14, 29–39. https://doi.org/10.1007/s10393-017-1211-7

TRAFFIC (2019) Combating wildlife crime linked to the internet - Global trends and China's experiences. TRAFFIC.

Uno KT, Quade J, Fisher DC, Wittemyer G, Douglas-Hamilton I, Andanje S, Omondi P, Litoroh M, Cerling TE (2013) Bomb-curve radiocarbon measurement of recent biologic tissues and applications to wildlife forensics and stable isotope (paleo)ecology. Proc Natl Acad Sci USA 110, 11736–11741. https://doi.org/10.1073/pnas.1302226110

Vaglica V, Sajeva M, McGough HN, Hutchison D, Russo C, Gordon AD, Ramarosandratana AV, Stuppy W, Smith MJ (2017) Monitoring internet trade to inform species conservation actions. Endanger. Species Res. 32, 223–235. https://doi.org/10.3354/esr00803

Veldman S, Gravendeel B, Otieno JN, Lammers Y, Duijm E, Nieman A, Bytebier B, Ngugi G, Martos F, van Andel TR, de Boer HJ (2017) High-throughput sequencing of African chikanda cake highlights conservation challenges in orchids. Biodivers. Conserv. 26, 2029–2046. https://doi.org/10.1007/s10531-017-1343-7

Veldman S, Otieno J, van Andel T, Gravendeel B, de Boer HJ (2014) Efforts urged to tackle thriving illegal orchid trade in Tanzania and Zambia for chikanda production. Traffic Bulletin 26, 47–50.

Vlam M, de Groot GA, Boom A, Copini P, Laros I, Veldhuijzen K, Zakamdi D, Zuidema PA (2018) Developing forensic tools for an African timber: regional origin is revealed by genetic characteristics, but not by isotopic signature. Biological Conservation 220, 262–271. https://doi.org/10.1016/j.biocon.2018.01.031

Wanderssee JH, Schussler EE (1999) Preventing plant blindness. Am. Biol. Teach. 61, 82–86. https://doi.org/10.2307/4450624

Williamson J, Maurin O, Shiba SNS, van der Bank H, Pfab M, Pilusa M, Kabongo RM, van der Bank M (2016) Exposing the illegal trade in cycad species (Cycadophyta: *Encephalartos*) at two traditional medicine markets in South Africa using DNA barcoding. Genome 59, 771–781. https://doi.org/10.1139/gen-2016-0032

World Bank (2019) Illegal logging, fishing, and wildlife trade : the costs and how to combat it. World Bank.

WWF (2018) New blockchain project has potential to revolutionise seafood industry. WWF.

WWF (2020) Wildlife crime [WWW Document]. URL https://www.wwf.org.la/what_we_do/wildlife_crime/ (accessed 5.25.22).

Zheng K, Zhang Z, Gauthier J, 2020 Blockchain-based intelligent contract for factoring business in supply chains. Ann. Oper. Res. https://doi.org/10.1007/s10479-020-03601-z

Answers

1. Any non destructive method would be potentially usable such as near infrared spectroscopy or X-ray micro CT, to preserve the samples in their original form.
2. Plant blindness is the bias towards animals, and taking-for-granted plants, which are not recognised as anything but background. The downside of plant blindness is that illegal plant trade is considered as relatively harmless as compared with illegal animal trade.
3. AMS requires a much smaller sample size (20–500 mg) compared to radiometric methods (10–100 g). It is also faster and usually produces higher precision results than radiometric methods. Samples can be analysed in a few hours with AMS, while it can take one or two days with radiometric methods. In case confiscated organisms are still alive, a fast verdict increases the chances of survival as rescued animals or plants can quickly be transferred back to the wild before they die.

Paper II

Orchidinae-205: a new genome-wide custom bait set for studying the evolution, systematics, and trade of terrestrial orchids

Margaretha A. Veltman, Bastien Anthoons, Audun Schrøder-Nielsen, Barbara Gravendeel, Hugo de Boer

Submitted for publication.

II

Paper III

Geographic shifts and increasing species diversity of wild orchid harvesting threatens survival of natural populations

Margaretha A. Veltman, Bastien Anthoons, Audun Schrøder-Nielsen, Jesus Adrian Chimal Ballesteros, Abdolbaset Ghorbani, Arzu Karahan, Esra Öztürk, Sedef Terzioglu, Selay Akan-Küçükaladağ, Susanne Masters, Mark Nesbitt, Michael F. Fay, Alexandre Antonelli, Barbara Gravendeel, Hugo de Boer

Manuscript.

III

Appendices

Appendix A

Supplementary information for Paper II

Appendix B

Supplementary information for Paper III