BIODIVERSITY GENOMICS EUROPE

TRAINING HANDBOOK

Instructions on how to Prepare and Deliver a BGE Joint Network Training Workshop

Photo from the "Train the trainers" event
that was held in Thessaloniki in October 2023
and inspired the production of the present handbook

Funded by the EU's Horizon Europe Research and Innovation Action

CONTENTS

Contributing Authors

Last Name	First Name	ORCiD	Affiliation
Bourlat	Sarah	0000-0003-0218-0298	LIB/ZFMK
Buzan	Elena	0000-0003-0714-5301	University of Primorska
Dankova	Gabriela	0000-0003-4453-6248	Naturalis Biodiversity Center
de Guttry	Christian	0000-0002-4341-1200	Swiss Institute of Bioinformatics
Hall	Dan	0000-0002-1102-9293	Natural History Museum, London
Mamos	Tomasz	0000-0002-0524-3015	University of Lodz
Mouton	Alice	0000-0002-0598-8789	University of Florence, Italy
Papakostas	Spiros	0000-0002-5563-0048	International Hellenic University
Price	Ben	0000-0001-5497-4087	Natural History Museum, London
Rewicz	Tomasz	0000-0002-2085-4973	University of Lodz
S. Marins	Luísa	0000-0002-5931-3533	Leibniz Institute for Zoo and Wildlife Research
Sheerin	Edel	0000-0002-3664-193X	Wellcome Sanger Institute
Shokoohi	Roya	-	Naturalis Biodiversity Research Center
Triantafyllidis	Alexander	0000-0003-0469-011X	Aristotle University of Thessaloniki
Kaitetzidou	Elisavet	0000-0002-1931-5876	Aristotle University of Thessaloniki
Minoudi	Styliani	0000-0003-2991-8483	Aristotle University of Thessaloniki
Vos	Rutger	0000-0001-9254-7318	Naturalis Biodiversity Center
Waterhouse	Robert	0000-0003-4199-9052	SIB Swiss Institute of Bioinformatics
Zinenko	Oleksandr	0000-0001-5228-9940	V. N. Karazin Kharkiv National University
Łukasz	Trębicki	0000-0002-6384-226X	University of Lodz

Preface

Biodiversity loss is one of humanity's most pressing challenges. With extinction rates significantly higher than historical baselines and an estimated 25% of species threatened globally, the urgent need to understand, monitor, and conserve Earth's biological diversity has never been greater. This handbook emerges from the Biodiversity Genomics Europe (BGE) project and it aims to equip the next generation of scientists with the tools and knowledge necessary to address this critical issue.

Biodiversity genomics stands at the forefront of this endeavor, offering unprecedented insights into life's complexity. By harnessing genomic technologies, we can delve deeper into species' genetic makeup, understand their interactions within ecosystems, and assess their responses to environmental changes. Two key methodologies in this field—DNA barcoding and high-quality reference genome generation—serve as complementary pillars for species monitoring and conservation.

DNA Barcoding: This technique involves sequencing short, standardized regions of organisms' DNA to generate a reference library of species-specific genetic barcodes to be later used for identifying those species to environmental samples, much like a database of supermarket barcodes that are used to efficiently scan products at the cashier. DNA barcoding thus accelerates species identification, enabling rapid and accurate biodiversity assessments. It is particularly valuable for cataloging life forms, monitoring ecosystem health, and detecting invasive species or pathogens. By building comprehensive reference libraries of these barcodes, we enhance our capacity for large-scale biomonitoring and contribute to completing the inventory of life on Earth.

High-Quality Reference Genomes: While DNA barcoding provides a snapshot for species identification, high-quality reference genomes offer a complete genetic blueprint of an organism. These genomes reveal detailed information about genetic diversity, evolutionary history, and adaptive potential. They are indispensable for understanding the functional aspects of genes, how species respond to environmental stresses, and for developing conservation strategies at the genomic level. Reference genomes also facilitate advanced studies in comparative genomics, aiding in the discovery of genetic factors that contribute to resilience or vulnerability among species.

Together, DNA barcoding and reference genome sequencing form a synergistic approach to biodiversity genomics. DNA b arcoding allows for rapid species assessment and monitoring across large spatial scales, providing immediate data for conservation actions. Reference genomes explore into the genetic makeup, offering insights that inform long-term conservation planning and species recovery programs. By integrating these methodologies, we can achieve a more holistic understanding of biodiversity, from species identification to the preservation of their genetic integrity.

This handbook is designed to guide you through the practical aspects of these techniques. It offers a modular training program that covers the main steps of DNA barcoding workflows using Oxford Nanopore Technologies. It also covers the extraction and quality control of high molecular weight DNA for genome sequencing. While it does not explore genome assembly and annotation, it provides the foundational laboratory skills and theoretical knowledge essential for any biodiversity genomics endeavor.

Our collective goal is to democratize access to genomic technologies, promoting collaboration across Europe and beyond. By standardizing protocols, sharing knowledge, and building expertise networks, we aim to lower barriers to participation in biodiversity genomics research. This aligns with the broader objectives of the BGE project: to enhance understanding of biodiversity, monitor its changes, and guide interventions to halt its decline.

We invite you to engage with the material, participate actively in the sessions, and contribute to this vital field. The challenges are immense, but so are the opportunities. Through combined efforts in DNA barcoding and reference genome generation, we can make significant strides toward conserving the rich tapestry of life on our planet.

Christian de Guttry

Case Study: Conservation of Endangered Riverine Fish

To illustrate the complementary roles of DNA barcoding and high-quality reference genome sequencing in conservation, consider the case of an endangered fish species inhabiting a European river system. This species, once abundant, has experienced a significant decline due to habitat degradation, pollution, and overfishing.

Phase 1: Initial Assessment with DNA Barcoding (Timeline: Year 1)

Conservationists initiated a biodiversity assessment of the river using environmental DNA (eDNA)* barcoding techniques. By collecting water samples and extracting DNA, they employed DNA barcoding to detect the presence of various fish species without physical capture. This non-invasive method confirmed the continued, albeit reduced, presence of endangered fish in certain river segments.

**Environmental DNA (eDNA) refers to genetic material obtained directly from environmental samples—such as soil, water, air, or sediment—without the need to capture or observe the organisms from which the DNA originates. By collecting and analyzing eDNA based on DNA barcodes, scientists can detect the presence of species in a particular environment.*

Phase 2: Monitoring Population Decline (Timeline: Years 2-3)

Regular monitoring using e-DNA barcoding over the next few years revealed a troubling trend: the concentration of endangered fish's DNA in water samples was decreasing, indicating a decline in population size. This method allowed for rapid, cost-effective monitoring across different seasons and river sections, providing timely data on species' status.

Phase 3: High-Quality Reference Genome Sequencing (Timeline: Year 4)

In response to the alarming decline, researchers sequenced a high-quality reference genome of the endangered fish. Leveraging this comprehensive genetic blueprint, they conducted a population genomics study using high-throughput genotyping technologies such as SNP chips and pooled sequencing (PoolSeq) etc. These methods enabled the analysis of genetic diversity, population structure, and adaptive potential. Genomic data revealed low genetic diversity and identified specific genetic bottlenecks that could hinder species' resilience to environmental changes.

Phase 4: Implementation of a genetic rescue program (Timeline: Years 5)

Armed with genomic information, conservationists designed a genetic rescue program. By identifying genetically robust individuals from neighbouring river systems or captive populations, they developed a breeding program to enhance genetic diversity. The reference genome guided selective breeding efforts to introduce beneficial genetic variants without compromising the species' unique genetic identity.

Phase 5: Continued Monitoring and Adaptive Management (Timeline: Years 6 onward)

Following the release of genetically enriched individuals back into the river, ongoing monitoring using DNA barcoding assessed the program's success. The rapid detection capabilities of barcoding allowed for the timely evaluation of population

recovery and the dispersal of introduced genetic variants. Adjustments to the conservation strategy were made based on real-time data, ensuring adaptive species management.

This case study exemplifies the synergy between DNA barcoding and high-quality reference genome sequencing:

1. DNA barcoding provides a quick and non-invasive means to detect and monitor endangered fish's presence and population trends over time.
2. The high-quality reference genome offered information on genetic diversity and vulnerabilities, which was necessary for informing conservation interventions.
3. Integrating data from both techniques enabled the design of a targeted genetic rescue program aimed at enhancing the species' long-term viability.
4. Continuous monitoring through DNA barcoding allows conservationists to evaluate the effectiveness of their interventions and adapt strategies as needed.

This approach showcases how combining DNA barcoding with reference genome sequencing creates a powerful toolkit for conservation biology. DNA barcoding facilitates efficient monitoring and immediate conservation responses, while high-quality genomes provide the genetic understanding necessary for long-term species survival strategies. Together, they enable a holistic approach to preserving biodiversity, particularly for species teetering on the brink of extinction.

Ethics and compliance guidelines

Ensuring ethical conduct and adherence to legal and institutional regulations is essential for the successful execution of this training program. This section outlines the necessary ethical considerations and procedures that **must be** followed by all participants and organizers.

General Data Protection Regulation (GDPR) compliance: When collecting personal data from participants, we must adhere strictly to GDPR guidelines to ensure privacy and data protection. Personal information will be collected only for specific, explicit, and legitimate purposes related to the training program. Participants will be informed about what data is being processed and how it will be used, ensuring transparency and obtaining explicit consent. All personal data will be stored securely, with access limited to authorized personnel involved in organizing the program. Data will not be retained longer than necessary, and participants have the right to access, correct, or delete their personal information at any time.

Consent for photographs and recordings: Photographs and recordings may be taken during the workshop for educational and promotional purposes. Before any recording or photography, participants will be asked to sign a consent form that clearly explains the purpose of the media, how and where it will be used and the duration of its use. A participant may decline participation in photographs and recordings without any negative consequences on their participation in the workshop. They may also withdraw consent at any time, and options will be provided for participants to remain anonymous if they prefer.

Compliance with Institutional Internal Rules: All activities within the workshop must align with the hosting institution's internal rules and policies. This includes following codes of conduct, safety regulations, and any specific guidelines related to facilities and resources. Participants are expected to behave professionally, promoting respect, integrity, and collaboration. Any form of misconduct, harassment, or discrimination is unacceptable and will be addressed promptly according to institutional procedures. Participants should familiarize themselves with these internal policies and seek clarification from organizers if needed.

Ethical Compliance in Laboratory Access and Use: Access to the molecular laboratory is granted under strict adherence to ethical and safety guidelines mandated by the hosting institution. All participants must comply with institutional policies regarding laboratory use, which include proper handling of equipment, use of personal protective equipment (PPE), and adherence to safety protocols. Participants are required to undergo safety training and demonstrate competency before engaging in laboratory activities. Unauthorized experiments or deviations from approved procedures are prohibited. Respect for the laboratory environment and equipment is mandatory, and any incidents or accidents must be reported immediately to the supervising staff.

Adherence to Host Country Regulations: Participants and organizers must comply with all host country laws and regulations. This includes obtaining the necessary permits and licenses for research activities, especially those involving biological samples or genetic material. Ethical considerations extend to environmental responsibility, ensuring that all activities are conducted sustainably and without harm to local ecosystems. Cultural sensitivity is also imperative; participants should respect local customs and traditions, and engage with the community respectfully and ethically.

Environmental impact and sustainability: We are committed to conducting all activities with the utmost respect for the environment, ensuring that our actions do not adversely impact the ecosystems we study. Participants must adhere to sustainable practices during fieldwork, laboratory sessions, and all other workshop activities. This includes minimizing ecological footprints by reducing waste, conserving resources, and avoiding unnecessary disturbances to natural habitats. All sample collections should be performed responsibly, ensuring local populations and ecosystems are not harmed or depleted. Proper waste management protocols must be followed, with all waste—especially hazardous materials—disposed of under local environmental regulations. By prioritizing environmental sustainability, we contribute to biodiversity preservation and set a positive example for responsible scientific research.

Inclusivity and Diversity: We are dedicated to encouraging an inclusive and welcoming environment that values and respects the diversity of all participants. The workshop is committed to providing equal opportunities regardless of gender, race, ethnicity, religion, age, sexual orientation, or disability. Participants are encouraged to share their unique perspectives and experiences, enriching the collective learning experience. We strive to make all activities accessible, providing reasonable accommodations to ensure everyone can fully engage in the workshop. Harassment or discrimination is strictly prohibited and not be tolerated. In embracing inclusivity and diversity, we not only uphold ethical standards but also enhance creativity, collaboration, and the quality of our scientific endeavors.

Our collective commitment to these ethics and compliance guidelines contributes to a safe, respectful, and legally compliant work environment that fosters learning and collaboration. The success of the training program and the advancement of biodiversity genomics research depends upon your commitment to these principles.

For any questions or concerns regarding these guidelines, please contact the program organizers or the designated ethics officer of your institution.

Introduction to the practicals

This training program is the work of members of the Biodiversity Genomics Europe (BGE) project as part of the Joint Network Training task within the BGE Work Package (WP12) focused on aligning activities and sharing knowledge between the reference-genomes and the reference-barcodes communities.

The sessions primarily focus on molecular laboratory techniques to prepare high-quality DNA extractions from biological samples. These practical aspects are contextualised with information on and demonstrations of prior steps e.g. sampling in the field, and downstream steps, e.g. data deposition to public repositories. Note that the sessions do not necessarily follow a standard sequential workflow for one objective (barcodes or genomes), but rather intertwine the different objectives to allow e.g. long incubation times, etc. that need to be performed as part of the protocol.

The program is designed to deliver a learning path, that is, a set of modular sessions that together cover the main steps of the DNA barcoding workflow using Oxford Nanopore (ONT) flongles, and the main steps of the High Molecular Weight (HMW) DNA extraction and quality control for later use in long-read sequencing for building reference genomes. The sessions do not cover genome sequencing to build reference genomes, genome assembly, and annotation, as these require more time, different protocols, and different computational approaches, and are already fairly well represented through other courses, for example via the Training and Knowledge Transfer hub of the European Genome Atlas (ERGA) community (https://www.erga-biodiversity.eu/team-1/tkt---training-and-knowledge-transfer).

The handbook is primarily designed to serve as a comprehensive resource for participants in the training programme, offering detailed guidance, protocols, and insights into the methodologies employed by the Biodiversity Genomics Europe (BGE) project. For any questions on this training material, please do not hesitate to contact us via the official communication form at the BGE website (https://biodiversitygenomics.eu/contact/).

General Notes to help you organise the event

i.e. not specific to a given session

Safety first

- Start with a general safety briefing, with the consent form and allergy alerts (e.g. latex gloves - offer the use of Nitrile gloves instead) - make sure everyone is aware of the dangers, and check if your institute may require some sort of disclaimer to be signed by all participants.
- All the participants who will be engaged in laboratory work should be wearing lab coats.
- Guide the participants through all the safety rules (fire exits, water for eyes etc.) of the building and the floor that they will be working at.
- Specialised work (handling of liquid nitrogen etc.) should be handled by trained personnel.

Things that may help with organising the event

- Start the training event with a comprehensive overview of each day's program, what will be going on in each session, learning outcomes, which parts are being done by the organisers/instructors, and which parts need to be done by the participant. Repeat this briefing at the start of each day in case your event is spanning over multiple days.
- Assign a timekeeper assistant, so instructors can be free to focus on explaining and demonstrating.
- Good to have a bell or ringer to call the attention of the whole class for a general explanation or time to move to the next step.
- Good to have the voice of the instructor projected over the speaker system for all participants to be able to hear well.
- If it is possible to have at least two instructors for each session, it would be easier to help participants.
- In case of a large audience bring assistants that are aware of the protocols to help individual groups to carry out the laboratory work.
- It is useful to have a good idea about the level of experience of participants beforehand, so the instructors would know how to help the participants (the same we do at the beginning of the semester or each new subject with students, to check what they know and what is very new to them).
- It is useful to explain or share beforehand with participants some examples of what can go wrong and how to avoid such problems as much as possible. e.g. how the sample can be contaminated.
- Having an overhead video stream can be important so that everyone in the room can see what the instructor is doing (also for recordings).
- Some certain things that should not be done in the lab, e.g. drinking, eating, make sure there are instructive signs on the lab door (e.g. no food, no drink).
- It may be useful to assess the level of experience of participants in relevant aspects of the training content, such as wet/dry lab experience, barcoding/genomic work experience, and knowledge in bioinformatics. Such knowledge would be beneficial when grouping participants into subgroups so that there is an equal balance of skills and expertise within each group.

- If possible, say a little bit about the theoretical background, for example, explaining what the particular enzyme or buffer added does and what is happening at the molecular level at this step.
- We need to have a code of conduct for participants to agree to abide by, we also may need to provide links to appropriate documentation on ethical considerations that could arise at any point during any of the sessions. Please also be aware of the GDPR guidelines that may apply.
- It is a good idea to have printed copies of the protocols that will be used (one hard copy per group of participants) as they are easier to follow while working in the lab than over a computer screen. A nice idea on top of this would be to have different types of protocols printed in different colours of paper.
- Ideally, all the protocols (pdfs) should be shared with the participants sometime before the course in case they want to prepare or do some reading before the workshop. You may also consider adding not just protocols, but also some useful papers with more explanation. This can be a voluntary reading, or if someone wants to learn more (after each session).
- As organiser, check twice if all the laboratory equipment is working properly before the start. Check if all the plastics are compatible (PCR plates and PCR cap strips). Sometimes workshops are done in laboratories that aren't used routinely by trainers and have lower quality standards (i.e. laboratory for student classes). Ensure there is a pre-prepared set of equipment provided by organisers and checked/accepted by trainers.
- Consider providing a CERTIFICATE of attendance.
- Consider providing the bio of each instructor to the participants (preferably in advance).
- Have stickers to write your name and stick to your top to be easily identified (all instructors and participants).
- Develop an evaluation form to share with the participants after the event.

<u>Things to avoid</u>

- Do not combine two protocols in parallel.

Session Information

0 = the template we will use for all the sessions

Learning Paths for Genomes, Barcodes, and Joint Workflows

Genomes

⇒1. HMW DNA extraction: pulverisation & digestion (90 minutes)
⇒⇒ 4. HMW DNA extraction: purification & QC (30 minutes)
⇒⇒⇒ 6. QC on HMW DNA extraction (60 minutes)
⇒⇒⇒⇒ 15. ENA database (30 minutes)

Barcodes

⇒2. Terrestrial & Aquatic sampling: theory & practical (45 minutes)
⇒⇒⇒ 3. Lysis C DNA extraction for barcoding (180 minutes)
⇒⇒⇒⇒ 5. Finish Lysis C extraction PCR & agarose gels (60 minutes)
⇒⇒⇒⇒⇒ 7. Agarose gel loading & running (60 minutes)
⇒⇒⇒⇒⇒⇒ 8. Pooling & Quantification (90 minutes)
⇒⇒⇒⇒⇒⇒⇒ 9. Training to load the MinION (45 minutes)
⇒⇒⇒⇒⇒⇒⇒⇒ 10. ONT library prep. & load the MinION (45 minutes)
⇒⇒⇒⇒⇒⇒⇒⇒⇒ 11. ONT barcoder (60 minutes)
⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒ 12. Sequence ID (45 minutes)
⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒ 13. Data upload to iBOL (35 minutes)
⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒⇒ 14. On the taxonomy & photographs (30 minutes)

Joint

⇒ 16. Citizen Science projects (30 minutes)

1. HMW DNA extraction: pulverisation & digestion

(C. de Guttry, SIB)

Technical and Learning Objectives: By the end of this wet lab session, participants will be proficient in the extraction of high molecular weight DNA from biological samples using pulverization and liquid nitrogen. They will gain hands-on experience in optimizing sample preparation protocols, including tissue homogenization through pulverization methods and subsequent digestion with specific enzymes to release intact high molecular weight DNA. Participants will also learn to assess the quality and yield of extracted DNA through gel electrophoresis and spectrophotometric analysis, ensuring successful isolation of intact genomic DNA suitable for downstream molecular applications such as long-read sequencing and genomic analysis.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 60 minutes.

Number of Assistants: One assistant per working group of 4-5 participants. Assistants should be familiar with wet lab techniques, and with handling liquid nitrogen.

Estimated Session Time: HMW extraction (45 minutes)
QC on HMW extraction (45 minutes)

Presentation: HMW extraction

Session Protocol: https://docs.google.com/document/d/1D4hZ8oVS6mTNO-ZW16a3FifxtKptgQBu_0uxEMdVkVo/edit

Video of the session: VIDEO

Safety Notices: Special attention should be given when handling liquid nitrogen (special gloves, lab coat, and eyeglasses for eye and skin protection)

2. Terrestrial & Aquatic sampling example techniques

(S. Minoudi & E. Kaitetzidou, AUTh)

Technical and Learning Objectives: During this training session participants will gain proficiency in a variety of sampling techniques relevant to DNA barcoding and environmental DNA (eDNA). The material covered in this training handbook includes water filtration for aquatic sampling, and Malaise Trap for terrestrial sampling. Depending on the expertise of each hosting institute other field collection methods for soil, water, and air samples, may be included. An important learning objective concerns sample preservation, storage, and transportation to maintain DNA integrity. By the end of the session, participants will develop a comprehensive understanding of sampling methodologies for DNA barcoding and eDNA, will enhance their practical skills through hands-on experience, and will gain insights into the potential applications of these techniques for biodiversity monitoring and ecological research.

Target Audience: Early career scientists and citizen scientists.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 15 minutes.

Number of Assistants: 1-2 assistants to help with setting up the malaise trap.

Estimated Session Time: 45 minutes

Presentation: Aquatic sampling with the Vampire sampler
Terrestrial sampling

Session Protocol: Steps for setting up a malaise trap
(also watch the corresponding video)

1. Choose a location spot without many rocks or waters in the surface.
2. Open the malaise trap and check that everything you need is there; the fabric (tent), the skeleton, ground tent stakes, ropes and collecting bottle filled with ethanol. A friend to help you!
3. Open the tent so that you can notice which part is the front side (tall part of tent) and which the back side (short part of tent).
4. Put together the skeleton parts and fit the skeleton in the tent.
5. Use the ground stakes and the ropes to secure the trap and make it stable (hint: Use nearby trees or wooden sticks to tie the trap for extra security).
6. Put the collector in the top of the front side of the trap.
7. Trap is ready to catch insects.

Video of the session: VIDEO

3. Lysis C DNA extraction for barcoding

(E. Sheerin, Sanger)

Technical and Learning Objectives: This session aims to provide participants with a robust and non-destructive approach to extract high-quality DNA suitable for DNA barcoding analysis. Through hands-on training, participants will learn the principles and protocols of Lysis C extraction, including sample preparation, lysis buffer optimization, and DNA recovery. This technique enables efficient DNA extraction without compromising the integrity of delicate arthropod specimens, facilitating downstream applications such as DNA barcoding. The Learning Objective focuses on equipping participants with the skills to apply Lysis C extraction effectively in arthropod biodiversity studies. Participants will gain a comprehensive understanding of the advantages and limitations of this method, enhance their proficiency in sample handling and DNA extraction techniques, and learn to troubleshoot common challenges encountered during the extraction process.

Target Audience: Early career scientists.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 60 minutes.

Number of Assistants: One assistant per working group of 4-5 participants. Assistants should be familiar with wet lab techniques.

Estimated Session Time: 180 minutes (hint: split into two sessions).

Presentation: Lysis C DNA extraction presentation

Session Protocol: https://www.protocols.io/view/sop-lysis-c-plate-based-dna-extraction-cnamvac6.html

Videos of the session: VIDEO

4. HMW DNA extraction: purification & QC

(C. de Guttry, SIB)

Technical and Learning Objectives: This session aims to equip participants with the necessary skills and knowledge to efficiently isolate intact, high-quality DNA suitable for advanced genomic applications. Participants will learn implementing quality control measures to ensure the integrity and purity of extracted DNA. They will gain knowledge in utilizing various methods, such as gel electrophoresis, spectrophotometry, and fluorometry, to evaluate DNA integrity, concentration, and purity.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 30 minutes

Number of Assistants: One assistant per working group of 4-5 participants with knowledge on how to load an agarose gel.

Estimated Session Time: 30 minutes

Presentation: QC on HMW extraction

Session Protocol: In the session we checked the integrity of the DNA extraction using an Agilent 4150 TapeStation System (manual) using the Genomic DNA ScreenTape Analysis. Similar outcomes can be achieved by loading a 1% agarose gel and running them for 90 minutes or more.

Videos of the session: VIDEO

5. Finish Lysis C extraction PCR & prepare agarose gels

(E. Sheerin, Sanger & D. Hall, NHM)

Technical and Learning Objectives: The objective of this session is to equip participants with the technical skills and theoretical knowledge necessary to successfully complete the Lysic C extraction (initiated in Session #3), and then continue with the Polymerase Chain Reaction (PCR) to amplify the target (barcoding) marker), and then to conclude with the preparation of the agarose gel(s) aimed to verify successful amplification (Session #6). Participants will conclude their learning of the principles and procedures involved in Lysic C extraction, enabling them to conclusively isolate DNA from various biological samples. Through hands-on practice, participants will gain proficiency in setting up PCR reactions, amplifying target DNA sequences (barcoding markers), and analyzing PCR products. Additionally, participants will learn how to prepare agarose gels of appropriate concentrations for DNA electrophoresis. Throughout the training, emphasis will be placed on understanding the importance of each step, troubleshooting common issues, and appreciating the applications of Lysic C extraction, PCR, and gel electrophoresis in the context of DNA barcoding.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 120 minutes*
**in case of a tight schedule pre-casted agarose gels may be used*

Number of Assistants: One assistant per working group of 4-5 participants with relevant knowledge (PCR, agarose gel preparation etc.).

Estimated Session Time: 60 minutes

Presentation: Lysate, PCR, and Gel Preparation
1-step PCR schematic
2-step PCR schematic

Session Protocol: Lysis C protocol
2-step PCR protocol

Video of the session: VIDEO

6. Agarose gel loading & running

(D. Hall, NHM)

Technical and Learning Objectives: In this session participants will acquire the technical skills and theoretical knowledge required for proficiently loading and running agarose gels for electrophoresis. Participants will learn the fundamental principles underlying agarose gel electrophoresis and its significance in separating DNA fragments based on size. Through practical demonstration and hands-on exercise, participants will develop the ability to prepare agarose gels of appropriate concentrations, load DNA samples onto the gels accurately, and perform gel electrophoresis under optimal conditions. By the conclusion of the session, participants will have acquired the necessary skills and knowledge to perform agarose gel loading and running with confidence and interpret gel electrophoresis results effectively facilitating their proficiency in conducting DNA barcoding work in the lab.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 60 minutes.

Number of Assistants: One assistant per working group of 4-5 participants with relevant knowledge (loading agarose gels).

Estimated Session Time: 60 minutes.

Presentation: NA

Session Protocol: NA

Video of the session: VIDEO

7. Pooling & Quantification

(D. Hall & B. Price, NHM)

Technical and Learning Objectives: This session aims to equip participants with the technical skills and theoretical knowledge required for effective DNA pooling and quantification to prepare MinION libraries for DNA barcoding. Participants will learn the principles and procedures involved in pooling DNA samples from multiple sources to create a representative pool for downstream sequencing applications. Through practical demonstrations and hands-on exercise, participants will gain proficiency in quantifying DNA concentrations accurately using spectrophotometry and/or fluorometry methods. Technical objectives include mastering the techniques of DNA dilution and normalization to ensure uniform representation of samples in the pooled library and optimizing the pooling ratio to achieve the desired sequencing depth. Learning objectives encompass comprehending the importance of accurate DNA quantification in library preparation, interpreting DNA concentration measurements from spectrophotometry and/or fluorometry readings, and troubleshooting common issues encountered during DNA pooling and quantification processes, facilitating the preparation of high-quality MinION libraries for DNA barcoding applications.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 30 minutes.

Number of Assistants: One assistant per working group of 4-5 participants with basic knowledge of working in a molecular biology laboratory and optimally they have practiced or at least have read the conducted protocol.

Estimated Session Time: 90 minutes.

Presentation: Pooling and Quantification

Session Protocol: NA

Video of the session: VIDEO

8. Training to load the MinION

(E. Sheerin, Sanger & D. Hall, NHM)

Technical and Learning Objectives: This session aims to provide participants with a comprehensive understanding of the MinION sequencing technology, including the structure and functionality of the accompanying flow cell and flongle. Through practical demonstrations and informative discussions, participants will gain insights into best practices, tips, and tricks for preparing the flow cell and flongle correctly, as well as loading the samples efficiently.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 15 minutes.

Presentation: Training_to_Load_the_MinION

Number of Assistants: One assistant per working group of 4-5 participants with knowledge on the practical session.

Estimated Session Time: 45 minutes.

Materials Needed: It is highly recommended to distribute used MinION flow cells and flongles to participants (at least one per group). The used flow cell may be filled with water to allow participants to familiarize themselves with the technique and practice the material taught.

Safety Notices: Gloves should be used when handling the MinION flow cell and flongle.

Session Protocol: NA

Video of the session: VIDEO

9. ONT library prep. & load the MinION

(E. Sheerin, Sanger & D. Hall, NHM)

Technical and Learning Objectives: This session aims to provide participants with the technical skills and theoretical knowledge required for effective DNA pooling based on the DNA quantities of the libraries estimated in a previous session. Essentially, this session completes the ONT library pooling step and proceeds with the preparation of the flow cell using the proprietary application of Oxford Nanopore, namely MinKNOW software, and the loading of the pooled DNA libraries. Additionally, the session covers various settings of the MinKNOW application.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.

Number of Participants: Depends on the facilities of the host institute.

Estimated Setup Time: 45 minutes

Presentation: NA

Estimated Session Time: 45 minutes

Number of Assistants: One assistant per working group of 4-5 participants with knowledge on the practical session.

Materials Needed: ONT libraries that have already been quantified, a MinION flow cell with an adequate number of active pores, or an unused flongle.

Session Protocol: Ligation Sequencing Kit V14 (SQK-LSK114) protocol

Video of the session: VIDEO

10. ONT barcoder

(A. Srivathsan, MFN)

Technical and Learning Objectives: This session aims to equip participants with the technical skills and theoretical understanding of Oxford Nanopore sequencing. It explores the evolution of Nanopore sequencing technology up to its current state-of-the-art, including its limitations and drawbacks such as the evolution of flow cells, basecalling software, and error profiles. Additionally, the session delves into the functionality of the software ONTbarcoder for demultiplexing and barcode sequence identification (consensus calling).

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding. Also, bioinformaticians with interest in analysis DNA barcoding data.

Number of Participants: It depends on the facilities of the host institute. Since this session involves in silico analysis (rather than wet lab work), it may be possible to accommodate more participants, for example, in a conference or seminar room.

Estimated Time: 60 minutes. Please allow 15 minutes for a Q&A of participants with limited knowledge in bioinformatics and/or to account for issues with installations or different laptop speeds.

Number of Assistants: 1-2 assistants may be necessary, depending on the size of the audience, to help with troubleshooting issues (e.g., programs not running or crashing, setting up options properly). Therefore, assistants should have a good knowledge of the practical session beforehand.

Presentation: ONT_barcoder_presentation

Materials Needed: It is advisable to instruct participants beforehand to bring their own laptops so that they can replicate the analytical steps in their own computer environment. Provisions are needed to ensure that participants have a working internet connection.

Cited Protocols/Papers: ONTbarcoder github repository

Video of the session: VIDEO

11. Sequence ID

(B. Price, NHM)

Technical and Learning Objectives: This session aims to equip participants with an understanding of the importance of verifying the validity of generated DNA sequences before uploading them to BOLD. It further discusses potential sources of misidentification, such as sample contamination and sample failures. The session also thoroughly demonstrates some popular online tools for performing sequencing ID checks, including GBIF Sequence-ID and BOLDigger. This includes an overview of their pros and cons, as well as screenshots of the workflow and expected outcomes for both failed and successful identifications.

Target Audience:	Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.
Number of Participants:	It depends on the facilities of the host institute. Since this session involves a presentation and/or *in silico* analysis (rather than wet lab work), it may be possible to accommodate more participants, for example, in a conference or seminar room.
Estimated Time:	About 45 minutes. Please allocate an additional 15 minutes for a Q&A session, especially for participants with limited experience working with the platforms. Additionally, if participants are unfamiliar with DNA sequences, the instructor and assistant(s) may need to prepare extra slides or background information to explain the principles of these topics in simpler terms.
Number of Assistants:	In the event of a live demonstration of any of the presented tools, it may be necessary to have one assistant available to address issues with participants who need help repeating the demonstrated tasks on their laptops.
Presentation:	Sequence_ID_presentation
Materials Needed:	In the case of live demonstration of any of the presented tools, it is advisable to instruct participants beforehand to bring their own laptops should they wish to replicate some of the steps. In this case, provisions are needed to ensure that participants have a working internet connection.
Cited Protocols/Papers:	GBIF_Sequence-ID_tool BOLDigger BOLD workbench BLAST
Video of the session:	VIDEO

12. Data upload to iBOL

(B. Price, NHM)

Technical and Learning Objectives: This session offers a general overview of BOLDSystems, focusing on submitting specimen data to the database. It includes a step-by-step guide on creating a project and uploading specimen data (including data, images, metadata, etc.) either individually or in batch mode. The session covers the most important fields that can be filled in during the submission process. Additionally, participants will receive feedback on their experience using the database, discuss its limitations and advances, and set expectations for future use. Ultimately, learners will formulate realistic expectations for utilizing BOLDSystems in future data submissions.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.

Number of Participants: It depends on the facilities of the host institute. Since this session involves a presentation and/or *in silico* live demonstration (rather than wet lab work), it may be possible to accommodate more participants, for example, in a conference or seminar room.

Estimated Time: 35 minutes. In the case of a live demonstration, please allocate an additional 15-20 minutes to allow participants to execute the presented instructions and to address any troubleshooting issues or additional questions.

Number of Assistants: In the event of a live demonstration of the submission procedure, it may be necessary to have one assistant available to address issues with participants who need help repeating the demonstrated tasks on their laptops.

Presentation: Data_Upload_to_BOLD

Materials Needed: In the case of live demonstration of any of the presented tools, it is advisable to instruct participants beforehand to bring their own laptops should they wish to replicate some of the steps. In this case, provisions are needed to ensure that participants have a working internet connection.

Cited Protocols/Papers: BOLDSystems_handbook
BOLD_user_account_creation_video
BOLD_project_creation_video
BOLD_single_specimen_data_submission_video
BOLD_batch_data_submission_video
BOLD_handbook_video

Video of the session: VIDEO

13. On the taxonomy & photographs

(B. Price, NHM)

Technical and Learning Objectives: This session offers an overview of the importance of photographs accompanying sequencing data in BOLD Systems. It discusses the benefits of providing as much complete information as possible and guides instructors through the steps of uploading photographs to the BOLD database. Additionally, it discusses the significance of providing accurate taxonomical information and additional metadata within BOLD Systems. This discussion includes an overview of the Barcode Audit Grade System (BAGS) for grading specimen information and presents the approach used by the BGE project.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in DNA barcoding.

Number of Participants: It depends on the facilities of the host institute. Since this session involves a presentation, it may be possible to accommodate more participants, for example, in a conference or seminar room.

Number of Assistants: Since this session involves a presentation, a single instructor may be possible to run the session alone.

Estimated Session Time: 30 minutes.

Presentation: Taxonomy_Photographs_presentation

Materials Needed: NA

Cited Protocols/Papers: BOLDSystems_handbook

Video of the session: VIDEO

14. ENA database

(C. de Guttry, SIB)

Technical and Learning Objectives: This session offers an overview of the European Nucleotide Archive (ENA) database, which serves as the European node of the International Nucleotide Sequence Database Collaboration (INSDC). It begins by presenting the BGE-ERGA sample collection and submission standard operating procedure, which includes important aspects of the regulatory framework such as sample permits and the ERGA manifest. Additionally, it covers the COPO metadata platform, which is utilized prior to generating high-quality whole genomes submitted to ENA, ensuring consistency in metadata. The session further explains the structure of both ENA and the ERGA BioProject.

Target Audience: Early career scientists of life sciences, preferably with some background knowledge in molecular biology and interest in whole genome sequencing.

Number of Participants: It depends on the facilities of the host institute. Since this session involves a presentation, it may be possible to accommodate more participants, for example, in a conference or seminar room.

Number of Assistants: Since this session involves a presentation, a single instructor may be possible to run the session alone.

Estimated Session Time: 30 minutes.

Presentation: ENA_database_presentation

Materials Needed: NA

Cited Protocols/Papers: ENA: Guidelines and Tutorials

Video of the session: VIDEO

15. Citizen Science projects

(S. Papakostas, AUTh, S. Bourlat, LIB)

Technical and Learning Objectives: This session discusses the importance of reaching out with the outcomes of the BGE project to the broader society, particularly to Citizen Scientists. It explores the framework through which this can be achieved, presents various ideas, and provides examples of activities that have been undertaken thus far.

Target Audience: Early career scientists of life sciences, with interest in communicating research outcomes and mobilising citizen scientists towards conservation and biodiversity research.

Number of Participants: It depends on the facilities of the host institute. Since this session involves a presentation/discussion, it may be possible to accommodate more participants, for example, in a conference or seminar room.

Number of Assistants: Since this session involves a presentation/ discussion, a single instructor may be possible to run the session alone.

Estimated Session Time: 30 minutes.

Presentation: Citizen Science in BGE presentation

Cited Material: BGE Citizen Science activities

Session Protocol: NA

Video of the session: VIDEO