A blue and black sign

Description automatically generated

Grant Agreement Number | 823827

Acronym | SYNTHESYS PLUS

Call | H2020-INFRAIA-2018-2020

Start date | 01/02/2019

Duration | 54 months

Work Package | WP7

Work Package Lead | Eva Chatzinikolaou

Task Lead | Pete Hollingsworth

Delivery date | 07/03/2023

Deliverable 7.3: Publication providing synthesis of cost-effective + scalable workflows for NH collections DNA sequencing

Giada Ferrari (RBGE), Lore Esselens (RMCA), Michelle L Hart (RBGE), Steven Janssens (BGM), Catherine Kidner (RBGE), Maurizio Mascarello (BGM), Joshua Peñalba (MfN), Flávia Pezzini (RBGE), Thomas von Rintelen (MfN), Gontran Sonet (RBINS), Carl Vangestel (RBINS), Massimiliano Virgilio (RMCA), Peter M Hollingsworth (RBGE)

**SYNTHESYS+**

**JRA2; 7.3: Developing the Protocol Infrastructure for DNA Sequencing Natural History Collections**

*Deliverable: D7.3 Peer review publication providing synthesis of cost-effective and scalable workflows for DNA sequencing natural history collections*

Intentionally preserved biological material in natural history collections represents a vast repository of past biodiversity, with an estimated three billion specimens worldwide. Advances in laboratory and sequencing technologies have made these specimens increasingly accessible for genomic analyses, offering a window into the genetic past of species. Sequencing natural history collections adds a temporal component to conservation and evolutionary biology studies and often permits access to information that can no longer be sampled in the wild. Due to their age, preparation, and storage conditions, DNA retrieved from museum and herbarium specimens is often poor in yield, heavily fragmented, and biochemically modified. This not only poses methodological challenges but also makes such investigations susceptible to environmental and laboratory contamination.

In Task 7.3 we combined empirical studies, literature reviews, and extensive dialogue with the community of researchers sequencing natural history collections to provide a synthesis of developments in protocols and workflows, to address the practical challenges associated with making the recovery of DNA sequence data from museum collections more routine.

Our work in Task 7.3 included a series of case studies, each focusing on protocol practicalities for the application of different mainstream methodologies to museum specimens including (i) shotgun sequencing of insect mitogenomes, (ii) whole genome sequencing of insects, (iii)  genome skimming to recover plant plastid genomes from herbarium specimens, (iv) target capture of multi-locus nuclear sequences from herbarium specimens, (v) RAD-sequencing of bird specimens, and (vi) shotgun sequencing of ancient bovid bone samples.

We also summarised the range of approaches being taken in different institutes to sequence natural history collections, ranging from low-throughput, highly controlled studies using ancient DNA (aDNA) facilities and protocols, through to large scale studies processing museum specimens in standard laboratory facilities. Recognising this heterogenous range of approaches, reflecting heterogenous study objectives, and resource availability, we outlined some pragmatic cost-effective options for laboratory facilities for processing museum specimens, that minimise risks of contamination, without unnecessary constraints on throughput.

We finish by reviewing key operational principles and issues to address, to guide the decision-making process and dialogue between researchers and curators about when and how to sample museum specimens for genomic analyses. This resulted in the development of 10 key principles to guide this decision-making process.

The primary outputs of this work have been submitted as a Methods paper to the *Biodiversity Data Journal*. This includes review and synthesis as well as the case studies. Additional papers relating to the different case studies will be submitted in the coming months.

The submitted pdf to the Biodiversity Data Journal is attached. The PDF has heterogenous font/spacing throughout, and some truncation of the tables. This is an output of the ARPHA writing tool, and we are not able to adjust these formatting issues. Once the formatting issues have been resolved by the journal team, we will pass on the updated file.

Peter Hollingsworth, Royal Botanic Garden Edinburgh

Task lead, 7.3