Final Report

Project title (Acronym)
DNA sequences for reliable identification of arthropod species of plant health importance (ArthCollect)

Project duration:

Start date:	2021-07-22
End date:	2023-02-01

Contents

1. Research consortium partners

Coordinator – Partner 1	
Organisation	Science and Agriculture for Scottish Agriculture (SASA)
Name of Contact (incl. Title)	Katherine Lester Rebecca Cairns
Job Title	Molecular Entomologist Molecular Entomologist
Postal Address	Roddinglaw Road, Edinburgh, EH12 9FJ, Scotland, UK
E-mail	katherine.lester@sasa.gov.scot; katherine.lester@forestresearch.gov.uk rebecca.cairns@sasa.gov.scot
Phone	+44 131 244 8830

Partner 2	
Organisation	Austrian Agency for Health and Food Safety (AGES)
Name of Contact (incl. Title)	Helga Reisenzein Richard Gottsberger
Postal Address	Spargelfeldstr. 191, 1220 Vienna, Austria
E-mail	helga.reisenzein@ages.at richard.gottsberger@ages.at

Partner 3	
Organisation	European and Mediterranean Plant Protection Organization (EPPO)
Name of Contact (incl. Title)	Françoise Petter Charlotte Trontin
E-mail	petter@eppo.int trontin@eppo.int

Partner 4	
Organisation	Department of Agriculture, Food and Marine (DAFM)
Name of Contact (incl. Title)	Maria Laura Destefanis Amanda Brechon
Postal Address	Plant Health Laboratories, Backweston Laboratory Campus, Young's Cross, Celbridge, W23 X3PH, Ireland
E-mail	maria.destefanis@agriculture.gov.ie amanda.brechon@agriculture.gov.ie

Partner 5	
Organisation	Ministry of Primary Industries, (MPI)
Name of Contact (incl. Title)	Disna Gunawardana Dongmei Li
Postal Address	Plant Health and Environment Laboratory, Diagnostic and Surveillance Services, Ministry for Primary Industries, 231 Morrin Road, St. Johns, Auckland 1072, New Zealand
E-mail	Disna.Gunawardana@mpi.govt.nz Dongmei.Li@mpi.govt.nz

Partner 6

Organisation	University of Agronomic Sciences and Veterinary Medicine of Bucharest (USAMV)
Name of Contact (incl. Title)	Roxana Ciceoi Mihaela Iordachescu
Postal Address	Laboratory of Diagnoses for plant protection, 59 Marasti Boulevard, Bucharest, 011464, Romania
E-mail	roxana.ciceoi@qlab.usamv.ro mihaela.iordachescu@qlab.usamv.ro

2. Short project report

2.1. Executive summary

Classical insect and mite taxonomy is a highly specialised skill. Taxonomists generally operate within a few areas of expertise and have limited opportunities to pass on their knowledge to the next generation or colleagues in other countries. DNA sequencing is increasingly used to complement classical taxonomic methods for the rapid and accurate identification of arthropod species and is of particular use to researchers and diagnosticians involved in plant protection. However, this method relies heavily on the accuracy and availability of sequences on public databases which although useful for biodiversity studies can be less reliable at the species level. Ideally, all barcode sequences should have been derived from a vouchered specimen, which was initially identified by a taxonomic expert, however this is not always the case. Errors can also be compounded if any newly generated molecular identification is based upon a previous misidentification.

Some important plant pest groups, such as scale insects (Infraorder: Coccomorpha), can be difficult to distinguish between morphologically. Furthermore, barcoding scale insects using the standard cytochrome c oxidase subunit I (COI) barcoding gene region has proven to be problematic, with alternative gene regions needing to be considered (Park *et al*, 2011)[1]. This Euphresco project aimed to exchange knowledge and expertise between partners on molecular barcoding methods for challenging arthropod pests. In-house studies on extraction and sequencing techniques were shared, some of which have been summarised in this report, including: DNA extraction method comparisons; comparing storage media for insects prior to extraction; and comparing alternative gene primer sets for identifying Coccomorpha specimens. A Test Performance Study (TPS) for the comparison of molecular methods identifying specimens of the Infraorder Coccomorpha (soft scales and mealybugs) was performed. One DNA extraction method using a non-destructive technique was evaluated, and the effectiveness of three barcoding primer sets were assessed in their ability to identify the Coccomorpha specimens tested. The participants found the TPS useful for optimising methods for extracting and sequencing this insect group.

Additionally, this project identified potential institute arthropod collections and sequences suitable for EPPO-Q-bank. Gaps in the database were also identified to assist future submissions.

2.2. Project aims

This project aimed to improve molecular barcoding techniques used for challenging arthropod pests through sharing knowledge and data between partners, and by carrying out a Test Performance Study identifying Coccomorpha specimens.

The main objectives of the project were:

- To exchange knowledge and expertise between partners on molecular barcoding techniques and troubleshooting. Some in-house studies on extraction techniques, storage media, and primer comparisons have been summarised and shared in this report.
- To carry out a Test Performance Study for comparison of molecular methods identifying specimens of the Infraorder Coccomorpha.
- To identify individual institute arthropod sequence collections and metadata that are available to submit to EPPO-Q-bank.

[1] Park *et al.* (2011). DNA barcodes for two scale insect families, mealybugs (Hemiptera: Pseudococcidae) and armored scales (Hemiptera: Diaspididae). DOI: 10.1017/S0007485310000714

2.3. Description of the main activities

2.3.1. Validating and improving molecular tools for identification of arthropods:

2.3.1.1. Comparing DNA extraction and purification methods for arthropods (SASA, GB)

This study aimed to review existing destructive & non-destructive extractions methods, automation of DNA extraction methods for arthropods and methods to handle smaller specimens. Four DNA extraction methods were evaluated: (1) QIAGEN DNeasy Blood & Tissue kit, (2) Macherey-Nagel kit, (3) automated extraction using a Magnetic Particle Processor (Kingfisher mL, Thermo Scientific) using QIAGEN Biosprint 15 DNA Blood Kit and (4) a manual chloroform extraction method. Their effectiveness using both a destructive and non-destruction lysis step was also examined.

For the destructive lysis method, ten aphids (*Megoura viciae*, Family Aphididae) and ten common pollen beetles (*Brassicogethes aeneus*, Family Nitidulidae) were homogenised in lysis buffer using a bead mill and left to incubate for a minimum of 1 hour at 56°C. DNA was then purified following the manufacturer's instructions for each kit or manually using chloroform.

For non-destructive lysis, a second set of ten aphids and ten beetles were pierced using a 0.1 mm pin in two places; through the abdomen and partly through the top of the thorax, and left incubating in lysis buffer overnight at 56°C. DNA was then purified following the manufacturer's instructions for each kit or manually using chloroform.

2.3.1.2. Comparison of storage media for arthropods (SASA, GB)

Successful barcoding of insects is dependent on extracting good quality DNA for sequencing and the first step to ensuring success is understanding the proper storage conditions required to reduce degradation of specimens. Traditionally insects are dried and pinned or preserved in 70% ethanol however, if subsequent extraction of DNA is required higher ethanol concentrations are necessary. Five percent glycerol is often added to ethanol to prevent brittleness and the morphological changes that can occur when storing insects in pure ethanol, but it was not recommended (personal communication) for studies where molecular work was required. No information on the potential effect of glycerol on DNA preservation was found in the literature to support this assumption and so further research was warranted.

For this study we investigated several preservatives to determine what, if any, effect glycerol or lower ethanol concentrations may have on preservation of DNA and sequencing success. We also explored the suitability of 'mini' barcode primers that amplify shorter segments (300-600bp) of the Folmer region for use with specimens that may have degraded or low molecular weight DNA.

Aphids (Family Aphidae) were collected by sweep netting onsite at SASA, in August 2021 and stored in six different preservatives at room temperature (20°C): (1) 95% Ethanol, (2) 75% ethanol, (3) 50% ethanol, (4) 95% ethanol & 5% glycerol, (5) IMS (industrial methylated spirit), and (6) tubes containing silica gel beads (Silica gel orange, Sigma). A final set of insects (7) were stored dry at -20°C.

At timepoints 3 months, 6 months, and 1-year, aphids were removed and DNA extracted to investigate how yield and quality of DNA, and sequencing success was affected over time. At the 3-month timepoint, 10 aphids were removed and extracted. For the 6-month and 1-year timepoint, 5 aphids were removed and extracted. DNA was extracted utilizing non-destructive lysis (piercing insects with a 0.1 mm pin) and purified using a Magnetic Particle Processor and QIAGEN Biosprint 15 DNA Blood Kit. The mitochondrial cytochrome c oxidase subunit I (COI)

gene 'Folmer' region was amplified by PCR and sequenced following the EPPO protocol for DNA barcoding of arthropods (EPPO, 2021)[2].
Secondly, shorter fragments of the 'Folmer' region were then amplified by PCR and sequenced using a selection of 'mini-barcode' primers: Primer set 1: LCO1490 & HCO2198 (EPPO, 2021)[3], Primer set 2: MF1 & LepR (Hajibabaei *et al*, 2006)[3], Primer set 3: LepF & Enh-LepR1 (Hajibabaei *et al*, 2006)[4] and Primer set 4: LCO1490 (EPPO, 2021)[5] & mICO1-intGLR (Park *et al*, 2019)[6].

2.3.1.3. Molecular Identification of infraorder Coccomorpha (SASA, GB)

This study aimed to evaluate the effectiveness of other barcoding primers for molecular identification of Coccomorpha. Review of the literature reveals that using the COI gene in conjunction with nuclear genes; elongation factor 1α (EF-1α), Internal Transcribed Spacer 2 (ITS2) and large ribosomal subunit (28s), can increase the likelihood of a successful molecular identification (Park *et al.* (2011)[7], Sethusa *et al.* (2014)[8]).
Specimens of Coccidae (soft scale), Diaspididae (armoured scale) and Pseudococcidae (mealbugs) were pierced with a 0.1mm pin and extracted using the QIAGEN DNeasy Blood & Tissue kit. Four gene regions and nine primer sets were compared: Mitochondrial cytochrome c oxidase subunit I (COI); Elongation factor 1α (EF-1α); Large ribosomal subunit 28s (D2 & D3 expansion region) and Internal Transcribed Spacer 2 (ITS2), (Table 1).

Table 1. Primer sets tested in Coccomorpha barcoding study

Primer set number	Gene region	Primer set	Reference
COI (1)	Cytochrome c oxidase subunit I (CO1)	LCO1490 & HCO2198	Folmer *et al.* (1994)[9] and EPPO PM 7/129 (2)[6]
COI (2)	Cytochrome c oxidase subunit I (CO1)	Arthropod COI cocktail mix	EPPO PM 7/129 (2)[6]
COI (3)	Cytochrome c oxidase subunit I (CO1)	PC0F1 & Lep-R	Park *et al.* (2011)[8] Hajibabaei *et al.* (2006)[5]
EF-1α (1)	Elongation factor 1α (EF-1α)	Prowler & Shirley	Djernaes & Damgaard (2006)[10]

[2] EPPO (2021). DNA barcoding as an identification tool for a number of regulated pests. PM 7/129 (2) https://doi.org/10.1111/epp.12724
[3] Hajibabaei *et al.* (2006). DNA barcodes distinguish species of tropical Lepidoptera. https://doi.org/10.1073/pnas.0510466103.
[4] Hajibabaei *et al.* (2006). DNA barcodes distinguish species of tropical Lepidoptera. https://doi.org/10.1073/pnas.0510466103.
[5] EPPO (2021). DNA barcoding as an identification tool for a number of regulated pests. PM 7/129 (2) https://doi.org/10.1111/epp.12724
[6] Park *et al.* (2019). Utilising samples collected in an existing biodiversity network to identify the presence of potential insect vectors of Xylella fastidiosa in the UK. https://www.planthealthcentre.scot/publications/identify-presence-potential-insect-vectors-xylella-fastidiosa-scotland
[7] Park *et al.* (2011). DNA barcodes for two scale insect families, mealybugs (Hemiptera: Pseudococcidae) and armored scales (Hemiptera: Diaspididae). DOI: 10.1017/S0007485310000714
[8] Sethusa *et al.* (2014). DNA Barcode Efficacy for the Identification of Economically Important Scale Insects (Hemiptera: Coccoidea) in South Africa. https://doi.org/10.4001/003.022.0218.
[9] Folmer *et al.* (1994). DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. https://pubmed.ncbi.nlm.nih.gov/7881515/
[10] Djernaes & Damgaard (2006). Exon-Intron structure, paralogy and sequenced regions of Elongation Factor-1 alpha in Hexapoda. https://www.senckenberg.de/wp-content/uploads/2019/08/asp_64_1_djernaesdamgaard_45-52.pdf

EF-1α (2)	Elongation factor 1α (EF-1α)	Manto & Phasma	Djernaes & Damgaard (2006)[11]
EF-1α (3)	Elongation factor 1α (EF-1α)	EF-1a(a) & EF2	Norse & Normark (2006)[11]
28s (1)	Large ribosomal subunit 28s	28s-D2-F & 28s-D2-R	Malausa *et al.* (2009)[12]
28s (2)	Large ribosomal subunit 28s	28s-S3660 & 28s-a335	Normark (2019)[13]
ITS2 (1)	Internal Transcribed Spacer 2 (ITS2)	M-ITS2-F & M-ITS2-R	Abd-Rabou *et al.* (2012)[14]

2.3.2. Interlaboratory Test Performance Study (TPS) for comparison of molecular methods identifying specimens of the Infraorder Coccomorpha

The scale insects and mealybugs of the Infraorder Coccomorpha contain numerous pest species found on crop and ornamental plants worldwide. Although adult females can be identified morphologically to species, problems that are routinely encountered with classical taxonomy such as, distinguishing between very similar species and identifying eggs and juvenile specimens, remain. DNA-based identification methods would typically be used in parallel alongside classical taxonomy however molecular identification of specimens from the Infraorder Coccomorpha has also proven to be problematic (Park *et al.*, 2011[15]).
The main aim of the TPS was to compare molecular methods for identifying specimens of the Infraorder Coccomorpha: Coccidae & Pseudococcidae (soft scales and mealybugs). Six laboratories across six different countries took part. One DNA extraction method (QIAGEN DNeasy Blood & Tissue kit using a non-destructive piercing technique) and the effectiveness of two other barcoding primers alongside the LCO1490 & HCO2198 Folmer primers were tested for their usefulness in molecular identification of two different species of Coccomorpha. For further details on the methodologies used, please refer to the TPS report.

2.3.3. Identifying existing arthropod collections for EPPO-Q-bank

Gaps in quarantine arthropod species/groups and some of their closest relatives in the arthropod database on EPPO-Q-bank were identified. Institutions identified vouchers and sequences within their own collections based on these findings. Together, a list of sequences and the metadata required was agreed in collaboration with the EPPO-Q-bank coordination team and the curators of the arthropod database.

2.4. Main results

2.4.1. Validating and improving molecular tools for identification of arthropods:

2.4.1.1. Comparing DNA extraction and purification methods for arthropods (SASA, GB)

DNA quantity varied between methodologies however overall high-quality sequences suitable for barcoding were obtained by all methods (Table 2). There was no significant difference

[11] Norse & Normark (2006). A molecular phylogenetic study of armoured scale insects (Hemiptera: Diaspididae). https://doi.org/10.1111/j.1365-3113.2005.00316.x
[12] Malausa *et al.* (2009). DNA markers to disentangle complexes of cryptic taxa in mealybugs (Hemiptera: Pseudococcidae). DOI: 10.1111/j.1439-0418.2009.01495.x
[13] Normark *et al.* (2019). Phylogeny and classification of armored scale insects (Hemiptera: Coccomorpha: Diaspididae). https://doi.org/10.11646/zootaxa.4616.1.1
[14] Abd-Rabou et al. (2012). Identification of mealybug pest species (Hemiptera: Pseudococcidae) in Egypt and France, using a DNA barcoding approach. https://doi.org/10.1017/S0007485312000041
[15] Park *et al.* (2011). DNA barcodes for two scale insect families, mealybugs (Hemiptera: Pseudococcidae) and armored scales (Hemiptera: Diaspididae). DOI: 10.1017/S0007485310000714

between extraction methods or between destructive and non-destructive lysis that would lead us to favour one method over another.

Table 2. Average DNA yield and percentage of samples that were successfully sequenced for each extraction method.

Extraction method		Chloroform ND	Chloroform D	MagAttract ND	MagAttract D	Qiagen ND	Qiagen D	Macherey-Nagel ND	Macherey-Nagel D
Average DNA yield (ng/µl)	Aphididae	223.6	67.7	75.5	101.5	62.6	38.4	x	x
Standard Deviation		139	96.4	57	53.4	44	16.8	x	x
Sequencing Success		100%	90%	100%	100%	90%	100%	x	x
Average DNA yield (ng/µl)	Nitidulidae	17.7	21.4	31.1	43.8	24.1	22.6	15.6	30.0
Standard Deviation		9.6	22.9	13.6	16.4	11.1	9.2	14.4	9.8
Sequencing Success		60%	90%	100%	80%	100%	100%	100%	90%

"ND" and "D" denote non-destructive lysis and destructive lysis methodologies, respectively. x = not done.

All kits have a typical recovery rate of DNA, depending on the starting material, unlike a manual chloroform extraction which theoretically will retain total DNA. DNA will be released more easily from a specimen if it is first crushed and this will be especially applicable for insects with a carapace such as beetles. Aphids have no carapace and relatively 'soft' bodies, therefore lysis is more efficient. The longer overnight lysis step required for the non-destructive lysis step will also aid with this. Generally more variation in DNA yield is observed using a manual chloroform extraction. This was also reflected in the raw data (not shown) where it is assumed that the DNA pellet was lost during processing if samples returned a quantity of <1ng/µl DNA. In comparison, the kit and automated DNA extraction methods produced more consistent yields and suffer less from operator error. This is also reflected in the percentage of successful sequences generated being higher.
Therefore, we recommend using non-destructive lysis and the QIAGEN DNeasy blood & tissue kit for routine extraction of DNA from arthropods. The DNA quantity and quality is sufficient for barcoding and less prone to operator error. If large quantities of samples are to be processed, automated extraction using a Magnetic Particle Processor e.g. Kingfisher mL (Thermo Scientific), Biosprint 96 (QIAGEN), or equivalent is also suitable.
For extracting DNA from very small insects e.g. thrips or mites, it was suggested within group discussions to try the QIAGEN QIAamp DNA Micro kit if experiencing low DNA yields with the DNeasy Blood & Tissue kit. Increasing the lysis incubation time to overnight and reducing the elution buffer volume may also assist in increasing the DNA concentration. It was also suggested if working with other very small specimens e.g. insect eggs, that a minimum of 5 may need to be extracted together, if feasible, to ensure enough high-quality DNA to sequence.

2.4.1.2. Comparison of storage media for arthropods (SASA, GB)

DNA yields varied between samples and timepoints (Table 3). Specimens stored in silica and 70% ethanol had noticeably lower concentrations of DNA throughout all timepoints compared

to the other storage conditions. DNA quality (not shown) was also poorest in silica stored specimens. PCR's were recorded as successful if a single band of correct size was observed using gel electrophoresis.

Table 3. Average DNA yields and PCR success rate per storage condition over timepoints

Storage conditions	3 months (n = 10)		6 months (n = 5)		12 months (n = 5)	
	Average DNA concentration (ng/µL)	PCR success (%)	Average DNA concentration (ng/µL)	PCR success (%)	Average DNA concentration (ng/µL)	PCR success (%)
50% ethanol	14.66 ng/µL	100%	20.29 ng/µL	0%	6.72 ng/µL	0%
70% ethanol	6.50 ng/µL	100%	9.71 ng/µL	0%	10.3 ng/µL	25%*
95% ethanol	59.59 ng/µL	100%	71.77 ng/µL	80%	58.13 ng/µL	60%
95% ethanol & 5% glycerol	29.98 ng/µL	90%	79.68 ng/µL	20%	90.29 ng/µL	40%
Industrial methylated spirit (IMS)	57.59 ng/µL	100%	28.77 ng/µL	20%	15.81 ng/µL	0%
Silica	3.15 ng/µL	70%	6.74 ng/µL	0%	1.11 ng/µL	0%
Freezer (-20°C)	28.76 ng/µL	100%	17.67 ng/µL	80%	15.12 ng/µL	20%

*n = 4 due to one sample lost.

Generally, the sequencing success rate reflected the PCR success rate (Table 3 & Figure 1). There were however some instances of PCR samples with weak bands that failed to sequence, accounting for the lower sequencing success rate. For silica stored samples at the 3-month timepoint, some of the sequences produced were of very poor quality (HQ < 5%) and therefore were also recorded as unsuccessful. At the 3-month time point, the number of successful PCR's and high-quality sequences generated were roughly similar for each storage method (excepting silica), but this drops off after 6 months for all storage media except: 95% ethanol and -20°C storage, both of which still produced high quality PCR's and sequences (Table 3 & Figure 1).

The 95% ethanol & 5% glycerol samples had similar high DNA concentrations and quality readings to the 95% ethanol samples. However, at the 6-month timepoint, the sequencing success was 0% for the samples stored in 5% glycerol, suggesting that some DNA degradation or PCR inhibition is occurring (Figure 1).

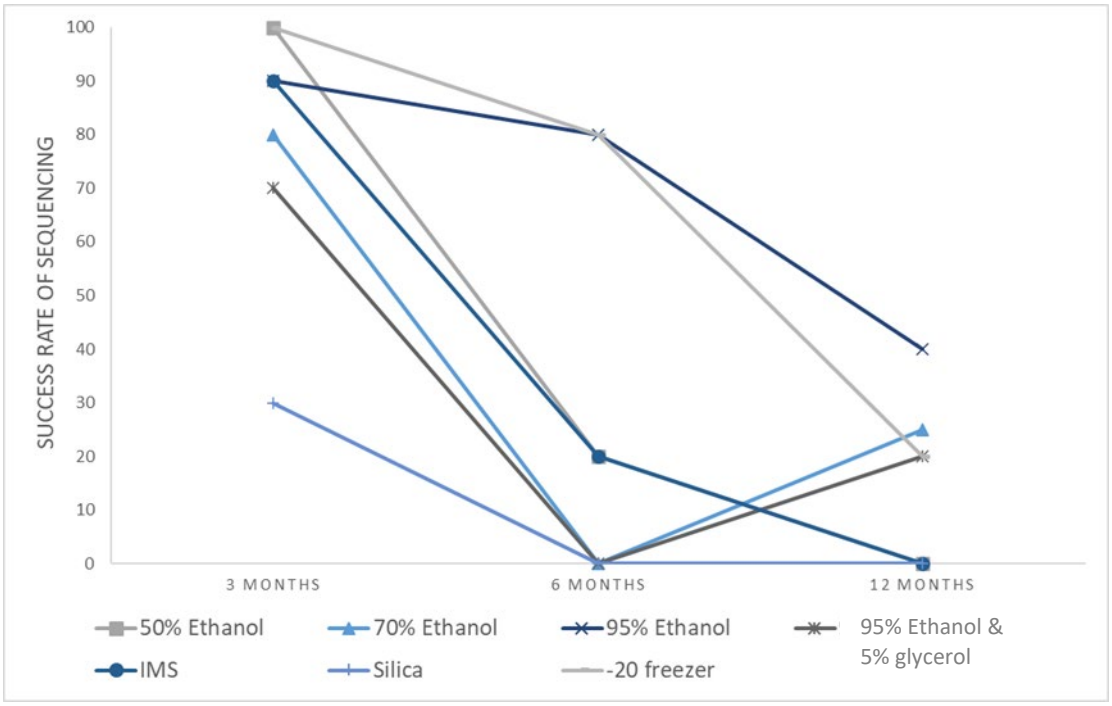

Figure 1. Percentage of specimens stored for 3, 6 and 12-month periods in different storage conditions from which DNA was successfully extracted and sequenced.

Individual sample variation will also be a contributing factor to sequencing success and can provide an explanation as to why there was some success at the 12-month timepoint for 70% ethanol and 95% ethanol & 5% glycerol samples, compared to no success at the 6-month timepoint. Furthermore, some aphids had unknowingly been parasitised by endoparasitoid wasps e.g. *Aphidius* sp., which may have interfered with sequencing quality results. Most previous studies have also found higher concentrations of ethanol are suitable for molecular work and preserve high molecular weight DNA effectively (Ballare *et al.*, 2019; Marquina *et al.*, 2020, Skvarla *et al.*, (2014).)[16,17,18].

The failure of silica gel to preserve DNA was unexpected given that this method of preservation usually leads to preservation of high molecular weight DNA (Quicke *et al.*, 1999)[19]. It is possible that not enough silica gel was used to completely dehydrate the insect so that residual moisture in the specimens led to degradation.

Larger sample sizes and repeats may be needed to help confirm findings, however preliminary results from this pilot study indicate that for storage of arthropods intended for barcoding, 95% ethanol should be used as a preservative or -20°C. Once DNA has been extracted voucher specimens can then be stored in absolute ethanol with 5% glycerol added to help maintain them for future morphological study. Long term specimen storage of 12 months or more could result in degradation of DNA. Other insect orders were not included in this study, however further research on this may be useful.

[16] Ballare *et al.*, (2019). Utilizing field collected insects for next generation sequencing: Effects of sampling, storage, and DNA extraction methods. https://doi.org/10.1002/ece3.5756

[17] Marquina *et al.*, (2020). The effect of ethanol concentration on the morphological and molecular preservation of insects for biodiversity studies. https://doi.org/10.7717/peerj.10799

[18] Skvarla *et al.*, (2014). Pitfalls and preservatives: A review. https://www.jstor.org/stable/26558988

[19] Quicke *et al.*, (1999). Preservation of hymenopteran specimens for subsequent molecular and morphological study. https://doi.org/10.1046/j.1463-6409.1999.00004.x

In several instances mini-barcode sequences were generated from specimens that had previously failed to amplify the 709bp product using the Folmer primers (primer set 1), meaning that amplification was possible from samples with degraded DNA (Figure 2).
If handling historical material in which some DNA degradation has likely already occurred, the 'mini' barcode primers listed within this report could aid amplification. Some of the 'mini' barcodes amplify consecutive sections of the gene region. It is therefore theoretically possible to generate a full 700bp COI sequence if required.

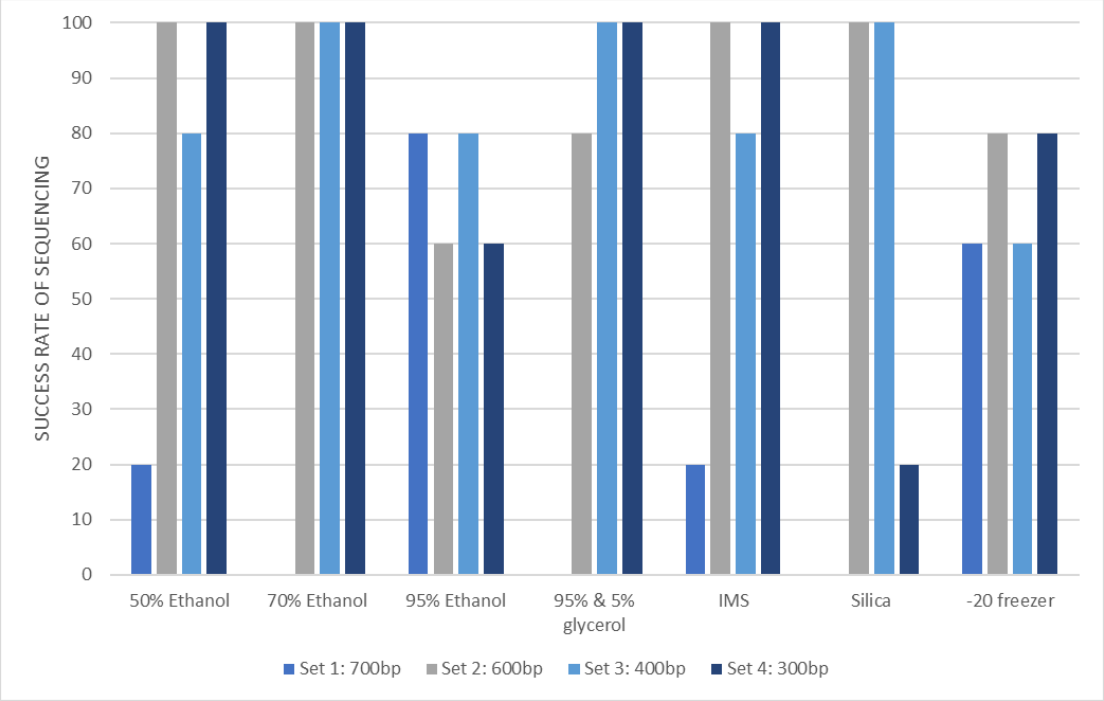

Figure 2. Percentage of samples that were successfully extracted and sequenced for each primer set. DNA was extracted from aphids stored for 6 months in different conditions. Primer set 1: LCO1490 & HCO2198, Primer set 2: MF1 & LepR, Primer set 3: LepF & Enh-LepR1, Primer set 4: LCO1490 & mICO1-intGLR.

2.4.1.3. Molecular Identification of infraorder Coccomorpha (SASA, GB)

Species from three different Coccomorpha families were tested with the primer sets (Table 4). The COI primer set PCOF1 & Lep-R (Park *et al*., 2011)[20] was more successful than the Folmer primers LCO1490 & HCO2198. It also had a higher success rate generating a high-quality consensus sequence than the LCO1490 & HCO2198 and arthropod cocktail primer mix (EPPO, 2021)[21]. We recommend also sequencing the large ribosomal subunit 28s gene region (Normark *et al*., 2019)[22] when a COI sequence cannot be amplified or in cases where the online databases are lacking in COI references sequences for comparison. And finally, although a limited range of species were compared within this study there is early evidence

[20] Park *et al.* (2011). DNA barcodes for two scale insect families, mealybugs (Hemiptera: Pseudococcidae) and armored scales (Hemiptera: Diaspididae). DOI: 10.1017/S0007485310000714
[21] EPPO (2021). DNA barcoding as an identification tool for a number of regulated pests. PM 7/129 (2) https://doi.org/10.1111/epp.12724
[22] Normark *et al.* (2019). Phylogeny and classification of armored scale insects (Hemiptera: Coccomorpha: Diaspididae). https://doi.org/10.11646/zootaxa.4616.1.1.

that the elongation factor (EF-1α) gene region primer sets (Djernaes & Damgaard, 2006)[23] & (Norse & Normark, 2006)[24] are of help when confirming identification of members of the Diaspidae family. We advise that all these different gene regions should be used in conjunction with COI and the morphological identification when possible to be confident of a correct molecular identification.

Table 4. PCR results for primer sets tested on Coccomorpha specimens

Family	Species	COI			EF-1α			28s		ITS2
		(1)	(2)	(3)	(1)	(2)	(3)	(1)	(2)	(1)
Pseudococcidae	*Planococcus citri*	-	-	+	+	-	+	+	+	+
Pseudococcidae	*Chryseococcus arecae*	-	-	-	-	-	nd	nd	+	+
Pseudococcidae	*Pseudococcus viburni*	mb	-	+	+	mb	+	+	+	+
Pseudococcidae	*Pseudococcus longispinus*	+	-	+	+	+	+	+	+	+
Diaspididae	*Aspidiotus nerii*	+	+	mb	+	+	+	+	+	+
Diaspididae	*Diaspis boisduvalii*	-	-	+	+	+	+	+	+	-
Diaspididae	*Gymnaspis aechmeae*	-	-	+	+	+	+	+	+	-
Coccidae	*Saissetia oleae*	-	-	mb	-	+	mb	+	+	-
Coccidae	*Saissetia coffeae*	-	-	+	-	+	-	+	+	-
Coccidae	*Coccus hesperidum*	-	+	mb	mb	+	mb	+	+	-
Coccidae	*Pulvinaria* sp.	-	-	+	-	+	mb	+	+	+

'+' =indicates PCR generated a clear single band present, '-' = PCR failed to produce a product, 'mb' = PCR resulting in multiple bands or smear, 'nd' = not done.

2.4.2. Interlaboratory Test Performance Study (TPS) for comparison of molecular methods identifying specimens of the Infraorder Coccomorpha

The key findings of the TPS found the QIAGEN DNeasy Blood & Tissue kit was successful at extracting DNA from Coccomorpha specimens and that the 28s primer set had the highest sequencing success rate followed by the alternative COI primer set. The Folmer COI primer set failed to produce a valid sequence for both Coccidae and Pseudococcidae specimens. Please refer to the TPS report for the full breakdown of results.

2.4.3. Identifying existing arthropod collections for EPPO-Q-bank

Sequences and metadata from each partner were submitted to the EPPO-Q-bank coordination team for the curators' consideration. Sequences were considered for addition in EPPO-Q-bank if they met standard criteria and fit in within the scope of the EPPO-Q-bank arthropod database e.g. species and origin are of interest and a voucher specimen is available. The final agreed upon submissions included 10 different arthropod species (across 5 orders and 7 families), 28 sequences in total (Table 5).

Table 5. Sequences of arthropod species undergoing submission to EPPO-Q-bank

Order	Family	Species	Loci	No.	Institute

[23] Djernaes & Damgaard (2006). Exon-Intron structure, paralogy and sequenced regions of Elongation Factor-1 alpha in Hexapoda. https://www.senckenberg.de/wp-content/uploads/2019/08/asp_64_1_djernaesdamgaard_45-52.pdf

[24] Norse & Normark (2006). A molecular phylogenetic study of armoured scale insects (Hemiptera: Diaspididae). https://doi.org/10.1111/j.1365-3113.2005.00316.x

Coleoptera	Cerambycidae	*Tetropium fuscum*	COI	1	DAFM
Diptera	Tephritidae	*Bactrocera albistrigata*	COI	1	PHEL/MPI
Diptera	Tephritidae	*Bactrocera dorsalis*	COI	1	PHEL/MPI
Diptera	Tephritidae	*Bactrocera tryoni*	COI	4	PHEL/MPI
Hemiptera	Psyllidae	*Bactericera cockerelli*	COI, ITS2	10	SASA
Hemiptera	Pseudococcidae	*Ripersiella hibisci*	18S/28S	1	AGES
Hemiptera	Diaspididae	*Aspidiotus nerii*	COI, 28S	1	SASA
Lepidoptera	Crambidae	*Maruca vitrata*	COI	1	DAFM
Thysanoptera	Thripidae	*Frankliniella panamensis*	COI	6	PHEL/MPI
Thysanoptera	Thripidae	*Frankliniella insularis*	COI	2	PHEL/MPI

Accepted submissions included species from the Cerambycidae, Tephritidae, Psyllidae, Pseudococcidae, Diaspididae, Crambidae and Thripidae families.
The partners within this project suggested other pest species and lookalikes which would be useful to include, either due to there being limited or no sequences available or to include additional sequences from other geographical locations: Scarabaeidae: *Phyllopertha horticola*; Chrysomelidae: *Diabrotica* sub spp.; Cecidomyiidae: *Enigmadiplosis agapanthi*; Muscidae: *Atherigona orientalis*; Phoridae: *Megaselia scalaris*; Tephritidae: *Bactrocera caryeae*, *B. occipitalis*, *B. pyrifoliae*, *Rhagoletis pomonella*; Ichneumonidae: *Ischnoceros caligatus*; Crambidae: *Cydalima perspectalis*, *Duponchelia fovealis*, *Maruca vitrata*; Pyralidae: *Etiella behri*; *Nephopterix proximalis*, *Elasmopalpus lignosellus*; Tortricidae: *Rhopobota naevana*, *Epiphyas postvittana*; Thripidae: *Thrips alni*, *T. flavus*, *T. palmi*, *T. urticae*.
Attention could also be focused towards different global plant health arthropod quarantine lists. While many arthropod species of plant health importance are available on other public databases e.g. NCBI GenBank and BOLD, they are not always easily identified as being fully validated sequences. As an example, **Appendix 1** details EPPO A1 and A2 arthropod quarantine species absent from EPPO-Q-bank as of 2022 and their presence or absence from other public databases (NCBI GenBank and BOLD).

2.5. Conclusions and recommendations to policy makers:

2.5.1. Validating and improving molecular tools for identification of arthropods

The partners of this project have had success using the QIAGEN DNeasy blood and tissue kit for routine DNA extraction from a variety of arthropod species. For very small samples e.g. insect eggs, a minimum of 5 specimens may be required to obtain enough high quality DNA for sequencing. Alternatively, the QIAGEN QIAamp DNA Micro kit was also recommended for extracting DNA from very small insects e.g. mites and thrips. If large quantities of samples are to be processed, automated extraction using a Magnetic Particle Processor is also suitable. Overnight lysis and reduced elution buffer volume can assist with smaller specimen extractions. The study comparing extraction methods for aphids and pollen beetles found no significant differences in sequencing success between extraction methods, and therefore recommend using the QIAGEN Blood & Tissue kit.
For storage of arthropods intended for barcoding, the consortium recommend 95% ethanol as a preservative or -20°C. If handing historical material in which some DNA degradation has likely already occurred, using ‘mini’ barcoding primers for smaller region amplification can aid

amplification and can act as a useful resource when the traditional Folmer primers have failed. Any species identification based on a smaller section (<400bp) of the COI mitochondrial region needs to be interpreted carefully as it may not be as accurate as an identification based on the full 709bp region. It is however theoretically possible to generate a full ~700bp region using multiple 'mini' barcodes that amplify consecutive sections of the gene region.

2.5.2. Molecular identification of infraorder Coccomorpha: Interlaboratory Test Performance Study (TPS) and additional in-house studies

The TPS comparing molecular methods to identify specimens of the Infraorder Coccomorpha proved useful in confirming that the QIAGEN DNeasy Blood & Tissue kit was successful at extracting DNA from soft scales and mealybugs. Using a non-destructive piercing technique, a voucher specimen could be preserved. However, it was clear that some damage was caused to specimens using this technique, making it difficult to morphologically identify post extraction. If adopting this method, extra care should be taken on the location of the piercing. We recommend using 28s primers for Coccomorpha to aid in identifications or if COI primer sets have failed. The alternative COI primer set was also more successful for Coccidae (soft scales) specimens than the standard Folmer region primers and could be used if the Folmer primers have failed. The TPS also highlighted the limitations of relying on public sequence databases like NCBI GenBank for a species level identification e.g. species not present in the database or misidentified species, which can lead to no matches or potential false matches. We encourage institutions and plant health laboratories to submit alternative barcoding regions e.g. 28s to public databases for species within the Infraorder Coccomorpha alongside COI when possible to aid in identifications.

The in-house study comparing additional primer sets for Coccomorpha species at SASA found similar findings to the TPS: 28s region primers had high sequencing success followed by an alternative COI primer set. There was also early evidence that the elongation factor (EF-1α) gene region primer sets are of help when confirming identification of members of the Diaspididae family (armoured scales).

2.5.3. Identifying existing arthropod collections for EPPO-Q-bank

The partners from this project submitted to EPPO-Q-bank arthropod sequences covering several families across Orders: Coleoptera, Diptera, Hemiptera, Lepidoptera and Thysanoptera of plant health importance. There are however many more species of concern and lookalikes to be added. Submitting other gene regions alongside COI could also aid in identification of species that cannot be readily identified using this gene.

2.6. Benefits from trans-national cooperation

The involvement of partners from across Europe and New Zealand permitted sharing of knowledge and expertise in molecular barcoding techniques for challenging arthropod pests, including DNA extraction methods and alternative primer sets. Online meetings allowed partners to discuss current research work and troubleshoot molecular issues within the group. Working alongside the EPPO-Q-bank team helped the group gain a better understanding of the process and metadata required for submitting sequences to the database. The partners will continue to submit suitable arthropod sequences of plant health importance as they are collected. Transnational cooperation also allowed the sharing of material for the Interlaboratory Test Performance Study. Insect specimens and DNA were shared, allowing six different laboratories across six different countries to take part, a study which the participants found useful for optimising the best methods for extracting and sequencing Coccomorpha specimens. Following on from the TPS, some laboratories have also had preliminary success using the 28s primers to identify species beyond Coccomorpha that they were previously struggling to sequence using standard COI primers, e.g. *Bemisia tabaci*.

3. Publications

3.1. Article(s) for publication in the EPPO Bulletin

None.

3.2. Article for publication in the EPPO Reporting Service

None.

3.3. Article(s) for publication in other journals

None.

4. Open Euphresco data

- Sequences submitted to EPPO-Q-bank
  https://qbank.eppo.int/project/EuphrescoARTHCOLLECT

- Arthcollect 2019-F-323 Interlaboratory Test Performance Study Final report: Interlaboratory Test Performance Study (TPS) for comparison of molecular tests identifying specimens of the Infraorder Coccomorpha: Coccidae & Pseudococcidae (soft scales and mealybugs)
  https://drop.euphresco.net/data/71c97f30-f051-4b92-9641-3e7004a3da96/

Appendix 1. Examples of arthropod quarantine species absent from EPPO-Q-bank as of 2022

Arthropod quarantine species (from EPPO A1 and A2 lists)			COI sequences available in other public databases	
Order	Family	Sequences not available in EPPO-Q-bank (as of 2022)	NCBI GenBank	BOLD
Acarida	Tetranychidae	*Oligonychus perditus*	1	1
Acarida	Eriophyidae	*Aculops fuchsiae*	0	0
Acarida	Eriophyidae	*Phyllocoptes fructiphilus* (Acari: Eriophyidae) vector	50+	0
Coleoptera	Buprestidae	*Agrilus anxius*	5	17
Coleoptera	Buprestidae	*Agrilus bilineatus*	6	10
Coleoptera	Buprestidae	*Agrilus fleischeri*	0	0
Coleoptera	Buprestidae	*Chrysobothris femorata*	50+	31
Coleoptera	Buprestidae	*Chrysobothris mali*	1	6
Coleoptera	Cerambycidae	*Apriona cinerea*	0	0
Coleoptera	Cerambycidae	*Apriona germari*	25	3
Coleoptera	Cerambycidae	*Apriona rugicollis* (*A. japonica*)	1	2
Coleoptera	Cerambycidae	*Massicus raddei*	1	0
Coleoptera	Cerambycidae	*Monochamus nitens*	3	0
Coleoptera	Cerambycidae	*Monochamus obtusus*	4	0
Coleoptera	Cerambycidae	*Monochamus* spp. (vectors of *B. xylophilus*)	*	*
Coleoptera	Cerambycidae	*Oemona hirta*	3	50
Coleoptera	Cerambycidae	*Saperda candida*	0	1
Coleoptera	Cerambycidae	*Tetropium gracilicorne*	1	0
Coleoptera	Cerambycidae	*Turanoclytus* (=*Xylotrechus*) *namanganensis*	2	6
Coleoptera	Cerambycidae	*Xylotrechus altaicus*	0	0
Coleoptera	Chrysomelidae	*Diabrotica virgifera zeae*	50+	1
Coleoptera	Chrysomelidae	*Epitrix subcrinita*	0	0
Coleoptera	Chrysomelidae	*Leptinotarsa decemlineata*	50+	21
Coleoptera	Curculionidae	*Anthonomus bisignifer*	0	1
Coleoptera	Curculionidae	*Anthonomus signatus*	12	23
Coleoptera	Curculionidae	*Conotrachelus nenuphar*	50+	181

EPPO A1 and A2 arthropod quarantine species
COI sequences available in other public databases
Order
Family
Sequences not available in EPPO-Q-bank (as of 2022)
NCBI GenBank
BOLD

Coleoptera	Curculionidae	*Dendroctonus adjunctus*	6	3
Coleoptera	Curculionidae	*Dendroctonus frontalis*	50+	50+
Coleoptera	Curculionidae	*Dryocoetes confusus*	1	8
Coleoptera	Curculionidae	*Gonipterus gibberus*	0	0
Coleoptera	Curculionidae	*Ips calligraphus*	8	2
Coleoptera	Curculionidae	*Ips hauseri*	4	0
Coleoptera	Curculionidae	*Ips lecontei*	2	0
Coleoptera	Curculionidae	*Ips subelongatus*	40	0
Coleoptera	Curculionidae	*Pissodes nemorensis*	4	5
Coleoptera	Curculionidae	*Pissodes terminalis*	1	5
Coleoptera	Curculionidae	*Polygraphus proximus*	1	55
Coleoptera	Curculionidae	*Pseudopityophthorus minutissimus* and *P. pruinosus* (as vectors)	1/1	3/0
Coleoptera	Curculionidae	*Premnotrypes latithorax*	1	1
Coleoptera	Curculionidae	*Premnotrypes suturicallus*	0	2
Coleoptera	Curculionidae	*Premnotrypes vorax*	1	1
Coleoptera	Elateridae	*Pheletes californicus*	50+	105
Coleoptera	Scarabaeidae	*Blitopertha orientalis*	11	9
Coleoptera	Scarabaeidae	*Heteronychus arator*	3	21
Diptera	Anthomyiidae	*Strobilomyia viaria*	12	6
Diptera	Cecidomyiidae	*Prodiplosis longifila*	38	7
Diptera	Tephritidae	*Bactrocera caryeae*	18	3
Diptera	Tephritidae	*Bactrocera kandiensis*	25+	124
Diptera	Tephritidae	*Bactrocera occipitalis*	50+	255
Diptera	Tephritidae	*Bactrocera pyrifoliae*	1	0
Diptera	Tephritidae	*Bactrocera tryoni*	50+	178
Diptera	Tephritidae	*Bactrocera tsuneonis*	50+	66
Diptera	Tephritidae	*Euphranta japonica*	0	0
Diptera	Tephritidae	*Rhagoletis cingulata*	50+	58

EPPO A1 and A2 arthropod quarantine species			COI sequences available in other public databases	
Order	Family	Sequences not available in EPPO-Q-bank (as of 2022)	NCBI GenBank	BOLD
Diptera	Tephritidae	*Rhagoletis indifferens*	26	7
Hemiptera	Psyllidae	*Bactericera cockerelli*	12	52
Hemiptera	Diaspididae	*Lepidosaphes ussuriensis*	0	0
Hemiptera	Diaspididae	*Quadraspidiotus perniciosus*	30	0
Hemiptera	Diaspididae	*Unaspis citri*	30	34
Hemiptera	Diaspididae	*Chionaspis pinifoliae*	50+	50+
Hemiptera	Pseudococcidae	*Ripersiella hibisci*	3	2
Hemiptera	Margarodidae	*Margarodes prieskaensis*	0	0
Hemiptera	Margarodidae	*Margarodes vitis*	0	0
Hemiptera	Margarodidae	*Margarodes vredendalensis*	0	0
Hymenoptera	Siricidae	*Sirex ermak*	0	3
Lepidoptera	Carposinidae	*Carposina sasakii*	50+	156
Lepidoptera	Crambidae	*Leucinodes africensis*	30	34
Lepidoptera	Crambidae	*Leucinodes pseudorbonalis*	1	1
Lepidoptera	Crambidae	*Leucinodes rimavallis*	2	2
Lepidoptera	Crambidae	*Neoleucinodes elegantalis*	4	8
Lepidoptera	Erebidae	*Orgyia leucostigma*	50+	156
Lepidoptera	Gelechiidae	*Keiferia lycopersicella*	8	27
Lepidoptera	Lasiocampidae	*Malacosoma parallela*	0	13
Lepidoptera	Noctuidae	*Spodoptera eridania*	50+	137
Lepidoptera	Nolidae	*Erschoviella musculana*	36	0
Lepidoptera	Tortricidae	*Acleris gloverana*	1	26
Lepidoptera	Tortricidae	*Choristoneura freemani* (= *C. occidentalis* Freeman)	50+	132
Lepidoptera	Tortricidae	*Cydia inopinata*	1	2
Lepidoptera	Tortricidae	*Cydia packardi*	40	81
Lepidoptera	Tortricidae	*Cydia prunivora*	15	27
Lepidoptera	Tortricidae	*Platynota stultana*	0	52
Thysanoptera	Thripidae	*Ceratothripoides brunneus*	18	15
Thysanoptera	Thripidae	*Ceratothripoides claratris*	0	3