FAIR Data Management: Provenance and the Nagoya Protocol

A report on the provenance requirements on FAIR metadata arising from the Nagoya Protocol and prospects on a possible implementation of ABS regulations on access and utilisation of DSI. Version 3.

Authors

Katrina Exter (VLIZ; EMBRC)
Belén Martín Míguez (ECIMAT; EMBRC)
Maria Huete Ortega (ECIMAT; EMBRC)
Arnaud Laroquette (EMBRC-ERIC)
Ibon Cancio (PiE-UPV/EHU; EMBRC)
Anne Emmanuelle Kervella (EMBRC-ERIC)

Audience

EOSC-Life (in particular WP4, 6)

Date

Version 1: 31-10-2019
Version 2: 22-03-2022
Version 3: 01-04-2023
Version 4: 07-08-2023

Disclaimer: the recommendations that appear in this document take into account only the present state of the art regarding Access and Benefit Sharing (ABS) regulations on the utilisation of genetic resources under the umbrella of the Nagoya protocol. Negotiations on incorporation of data (Digital sequence information) into an ABS mechanism are still pending after successive delays in conveying the After the 15th Conference of the Parties (COP15) of the Convention of Biological Diversity and the resulting Kunming-Montreal Global Biodiversity Framework agreement the approach towards the digital sequence information is still pending, although it will not be dealt through any bilateral mechanism. The same needs to be said about the Draft agreement under the United Nations Convention on the Law of the Sea on the conservation and sustainable use of marine biological diversity of areas beyond national jurisdiction.

1 Scope of this document

This document has been prepared for WP4 and WP6 of EOSC-Life. This report is done under Task 6.4: *FAIR data management standards for protecting genetic resources* (Lead: EMBRC). This version was written to summarise the subsequent developments in the international application of the Nagoya Protocol (NP) and the ongoing negotiations on the possible incorporation of Digital Sequence Information into the Nagoya definition of genetic resources that have been incorporated into de Kunming Kunming-Montreal Global Biodiversity Framework and the Draft agreement under the United Nations Convention on the Law of the Sea on the conservation and sustainable use of marine biological diversity of areas beyond national jurisdiction.

This document aims to pinpoint the requirements that are imposed on RIs when providing access to genetic resources[1], as well as utilising[2] them, as a result of the legal frameworks developed by the UN Convention on Biological Diversity (CBD), and in particular the NP on Access and Benefit Sharing (ABS) arising from the use of genetic resources. This report distils the requirements that are imposed on scientific data and data management practices as a result of the ongoing discussions in the CBD.

The aim of the NP is to ensure that the use of genetic material results in the protection of biodiversity, respecting the sovereign rights of countries over their natural resources and the traditional knowledge associated with their use. This protocol has been signed by many countries (134 parties as of February 2021) and has been adopted by the EU as part of its regulatory framework (e.g. EU Regulation 511/2014).

The ABS regulation requires that permission from the country of origin of the genetic resources used in research is sought whenever they are sampled from the wild or obtained from collections or any third party, and further that the use of those resources can be tracked. Humans and model organisms are out of scope, but microorganisms collected on or in human bodies (or in model organisms) may be in scope of the ABS regulation. This tracking places certain requirements on the metadata associated with the (digitised) genetic resources, in particular on the *provenance* metadata. Although the NP (as yet) does not deal with data (for instance, digital sequence information, aka DSI), the national ABS regulations of some countries already include DSI in the definition of genetic resources. In this document we will summarise what the NP is and its scope, explain what you should do to comply with the NP, make a first recommendation on the provenance metadata necessary to satisfy the NP, and indicate which RIs in EOSC-Life are likely to be affected. New decisions on DSI of the CBD (Kunming-Montreal GBF) and UNCLOS (conservation and sustainable use of BBNJ) have delineated a multilateral approach towards ABS, avoiding inclusion of DSI in the definition of genetic resources of the Nagoya protocol.

In this report we distil information on the NP obtained from the ABS Clearing House, EU reporting, and various publications. However, bear in mind that aspects of the NP are still being discussed, globally between the signatories and individually within each country, and the requirements are likely to change in

[1] The genetic resources referred to are non-human biological material and natural compounds (animals, plants, microbes or their derivatives) removed from their habitat and either obtained in situ (fieldwork), ex situ (via collections, resource centres, biobanks), or through a transfer of material (from colleagues or purchased). Specific definitions are laid out in part 3 of the document.

[2] Utilisation of genetic resources means to conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology as defined in Article 2 of the CBD.

the future. Please also remember that this report is concerned only with the metadata implications related to the NP, and not those which would result from any other similar regulations within each country.

2 What is the Nagoya Protocol?

The concept of ABS stems from the CBD, signed by 168 countries, which, amongst other objectives, seeks to ensure the fair and equitable sharing of benefits arising from the utilisation of genetic resources (GR). The discussion on accessing and using GRs in a fair and equitable way has its roots in the North–South debate[3], which goes as far back as the de-colonisation era. The developing countries (including newly industrialised), many of them home to the most bio-diverse ecosystems and providers of biological resources, have for a long time demanded respect for their sovereign rights over their resources and the need for equitable sharing of benefits.

The Nagoya Protocol to the Convention on Biological Diversity is a supplementary agreement to the CBD. It provides a transparent legal framework for implementing ABS and for the effective implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilisation of GR. The Nagoya Protocol on ABS was adopted on 29 October 2010 in Nagoya, Japan and entered into force on 12 October 2014.

The ABS rules complement the national and international rules regulating access to natural resources that are already in place. In this document, we *only* consider the requirements on provenance metadata enforced by the NP.

The NP applies to GR under national jurisdiction that are covered by the CBD, and to the benefits arising from their utilisation. The Protocol also covers traditional knowledge associated with the utilisation of GR, although in this report we will focus on GR only. The NP does not deal with genetic resources in Areas Beyond National Jurisdiction (ABNJ) or in Antarctica, although the genetic resources in the high seas are under discussion in the UN Intergovernmental Conference for an International Legally Binding Instrument on the conservation and sustainable use of marine biological diversity in ABNJ under UNCLOS. The NP establishes, for the first time, obligations to support ABS compliance by signatory countries, and sets out the principles of Mutually Agreed Terms (MAT) and Prior Informed Consent (PIC)[4] to be applied to the collecting and use of GR. In addition, Art. 8 of the NP calls for states to promote and encourage non-commercial research by using simplified measures, taking into account the need to address any change of intent. The basic outcomes of the NP are:

- States have sovereign rights over their natural resources within their territory and their Exclusive Economic Zone (EEZ).
- States have rights over scientific research, data and samples in the remit of their territory (EEZ included).
- States may (or not) adopt a national legal framework on ABS.

[3] The North−South debate is the term used to describe the process through which the “third world countries” engaged the rich Northern countries (North America and Western Europe) in negotiations over changes to the international economic system during the 1970s (New International Economic Order). This process had serious impacts on the development of international environmental law.

[4] MAT: specific benefit sharing conditions must be "mutually agreed" between providers and users of genetic resources. PIC: refers to the administrative permit given by the competent national authority of a provider country to a user, prior to accessing genetic resources. For full definitions, see this glossary.

- Benefit sharing may be monetary or non-monetary and should be agreed in MATs between the country providing the GR and the user of GR.
- Users must check if a PIC from the provider country is required for using GR, in addition to the permit of the State to access the GR for scientific research.

A list of the parties to the NP can be found on the ABSCH site, with links to the legislation and contact points for each country given. National implementation of the NP is generally a complex process. This list includes the signatories of the NP, and those who have not signed but do have NP focal points and/or ABS regulations. The parties are required to develop the necessary legislation to deal with the NP, and this national implementation may go beyond the minimum standards set out in the NP. At the same time, not all parties have yet fully adopted the NP in their national legislation, and not all have yet set up offices to deal with NP-related requests.

2.1 EU implementation of the NP

EU Regulation 511/2014 (EU ABS Regulation) on compliance measures for the NP came into force in 2014 in the European Union. The EU ABS Regulation calls for users to comply with applicable ABS legislation and regulatory requirements, the so-called "Due Diligence". This means that all users of GR must exercise due diligence to ascertain that (i) the GR being used was obtained in accordance with applicable ABS legislation or regulatory requirements (i.e. that a PIC was obtained when the GR were sampled/obtained), and (ii) that benefits are fairly and equitably shared upon mutually agreed terms (i.e. that a MAT had been obtained).

The EU has published a handy guide to the scope of the NP within the EU: the *Horizontal guidance document on scope of the Basic Regulation (EU) No511 / 2014 and the Implementing Regulation (EU)2015/1866 (Official Journal, 12 January 2021)*. Some basic points regarding the scope of the EU ABS regulation from that document are copied here and in Table 1 (see end of report).

- The EU ABS Regulation applies to GR over which States exercise sovereign rights from 12 October 2014 (the date of the NP).
- The EU ABS Regulation applies to GR that were accessed from a country of origin that is party to the NP or from another provider country that is party to the NP.
- In cases where GR are obtained by the user indirectly (e.g. from culture collections) the national legislation or other requirements (including national environmental protection law) of the provider country still apply.
- If a collection holds material that exists in natural habitats in a country party to the NP and that country adopts an ABS law, and the material is accessed from the culture collection after the entry into force of the NP, this falls within the scope of the Regulation.
- Traditional Knowledge associated with GR, in general and independently from the utilisation of said GR, falls within the scope of the EU ABS Regulation. It must be covered by the relevant contractual agreements.
- Derivatives are within the material scope of the EU ABS Regulation depending on their utilisation.
- "Utilisation" means to conduct research and development (R&D) on the genetic and/or biochemical composition of GR, including the application of biotechnology. It may include both basic and applied research.
- Archiving, culturing or biobanking does not constitute utilisation of GR.
- A user is obliged to comply with the ABS legislation in the country where the fieldwork is planned.

- A user is obliged to demonstrate whether requirements for PIC and MAT were met when the resources were originally accessed.
- Due diligence needs to be exercised in the whole course of work starting before accessing the material.
- The EU ABS regulation does not apply to GR for which access and benefit-sharing is governed by specialised international instruments that are consistent with, and do not run counter to the objectives of the Convention and the NP.

There are check-points to be decided by each country to verify that the terms established in their ABS regulations are met by users. In Europe, one of these check-points occurs at the moment of applying for a research project or at the moment of issuing a patent. The EU has created the online tool DECLARE, to make it easier for users to declare that they have exerted due diligence during the use of GRs (see Sec. 4).

3 The scope of the EU ABS Regulation

Essentially ABS means that it is necessary for anyone collecting GR in the wild (*in-situ GR*) or accessing GR from a collection (*ex-situ GR*) or from any third party, or even collecting micro-organisms on or in the human body, to do so in accordance with the country's legislation on ABS in place[5], and furthermore to be able to at any subsequent time to prove compliance (i.e. by providing copies of the related documentation). It is necessary for anyone who then uses the GR to check that the necessary documents are obtained before they can (re)use the resource, and to ensure that the use they are making of the GR falls under the conditions set out in the ABS permits.

We consider the following terms to be important in understanding the scope of the NP.

- **A genetic resource (GR)**. A GR means all genetic material of actual or potential value, from living organisms (plants, animals and microbes), which are taken from the wild, domesticated or cultivated, and are sourced from natural environments or collections (e.g. botanical gardens, gene banks, seed banks and microbial culture collections)[6]. The term "genetic resource" includes also its "derivatives", i.e. naturally occurring biochemical compounds.
  Note that each country can define GR differently in its regulations.
  For more on GR (definition and scope) and on the consideration of digital data obtained from gene sequencing, see the *Horizontal guidance document* (sec. 2.3.1, 2.3.3): for the latter point, there is still some lack of clarity on the definition of a GR once it has been sequenced and digitised.
- **Utilisation of GR**. This means to conduct research and development on the genetic and/or biochemical composition of genetic resources, including through the application of biotechnology as defined in Article 2 of the CBD (Article 3(5) of the Regulation; Article 2(c) of the Protocol). More text can be found in the *Horizontal guidance document* (sec. 2.3.3). The definition of "research and development" is also given in this document and examples of activities that fall under, or not, the definition are given. Note that provider countries may have established different conditions for different types of utilisation in their access legislation. This also applies to the definition of GR and the utilisation of GR.

[5] Any similar national legislation that was created independently of the NP should clearly also be conformed with, but we do not address these possibilities in this report
[6] Modified from the Glossary of Terms

- **Traditional knowledge**. This means traditional knowledge held by an indigenous or local community that is relevant for the utilisation of genetic resources. More text can be found in the *Horizontal guidance document* (sec. 2.3.2).
- **Access**. This means the acquisition of GR or of traditional knowledge associated with GR. This includes in-situ ("in the field") and ex-situ (e.g. in collections) acquisition. For more on ex-situ acquisition of GRs, see the *Horizontal guidance document* (sec. 2.1.3).
- **Derivatives**. This is defined as "a naturally occurring biochemical compound resulting from the genetic expression or metabolism of biological or genetic resources, even if it does not contain functional units of heredity". See the *Horizontal guidance document* (sec. 2.3.4) for more text on this term. At present, derivatives do not apply to those derived from sequences extracted from the GR.
- **Due diligence**. This applies to all users, who are obliged to "exercise due diligence to ascertain that the GRs … have been accessed in accordance with the applicable NP legislation". That means that the user has taken all reasonable measures to act according to the ABS expectations, and that they can show that they have taken these measures: all documentation must be kept for up to 20 years after the sampling or access. It requires that all necessary information on the date, place, and circumstances of the sampling of the GR or TK can be provided, along with a description of the GR/TK including their source and later use.
  The concept is further explained in the *Horizontal guidance document* (sec. 3.1, 3.5, also sec 3.7 with reference to obtaining GR from registered collections, and sec 4 with reference to the checkpoints at which due diligence should be exercised in the EU).
- **Users**. Both individuals and institutes are required to act with due diligence. However, a person who only transfers material is not a user in the meaning of the Regulation, and a person or entity which only commercialises products which have been developed based on utilisation of genetic resources or associated traditional knowledge is neither. See the *Horizontal guidance document* (sec. 2.4).
- **Temporal scope**. In EU, genetic resources accessed before the 12 October 2014 fall outside of the EU ABS Regulation. However, some Parties to the NP have put in place rules that apply also to genetic resources accessed prior to that date. Those rules still apply and should be respected even if not covered by the EU ABS Regulation. See the *Horizontal guidance document* (sec. 2.2). As a general recommendation, we can consider that the utilisation of GR triggers the ABS mechanisms, regardless of when the GR was sampled.

The scope of the EU ABS Regulation has many exceptions. It is important to note that human samples are out of scope of this regulation. Model organisms developed and commonly used in laboratories are also out of scope, see the *Horizontal guidance document* (sec. 7.5).

Regarding the human microbiota, conditions apply to define whether studies will be in or out of scope of the EU ABS regulation. Note that the other usual national legal and ethical requirements still apply.

- When the human microbiota is studied in or on the human body, the studies are considered to be out of scope of the EU ABS Regulation, because of the symbiotic interaction between the microbiota and the human body.
- When the microbial communities are taken from a human body or body products obtained from an individual, the studies are considered to be out of scope of the EU ABS Regulation, because of the unique composition of the microbiota in each individual.
- When the studies are carried out on individual taxa isolated from a sample of the human microbiota, the studies are considered within scope of the EU ABS Regulation, since the isolate no longer

represents the unique microbial characteristic of an individual human. However, the mere taxonomic identification of the organisms is considered to be out of scope.

In the last situation, from where the ABS permit should be requested depends not on where the human subject was situated when their microbial community was sampled, but where the microbial community was formed. Some examples:

- The subject in country A flies over to country B to be sampled upon arrival: the microbes come from country A.
- The subject flies to country B and stays there for a few weeks before being sampled: the microbes come from country B.

In this scenario, it will be necessary to ask sufficient questions of the subjects being sampled, and to consult the regulations of the concerned countries to find the details of timeframes.

The EU ABS regulation does not deal with data (for instance digital sequence information) although:

- The history of utilisation of a biological sample needs to be traceable and tracking this requires provenance data to be carried out along the whole process.
- The inclusion of the DSI data in some kind of ABS mechanism is under discussion in the CBD at the moment and is very relevant to some Global South countries before approving and signing the Post-2020 Global Biodiversity Framework. Such discussions were channelled in COP15 in Montreal after four postponements for three years and partially settled with the agreement around the Kunming-Montreal Global Biodiversity Framework. This framework considers, without defining them, multilateral approaches towards ABS and the use of DSI. Note that some countries already include DSI in their definition of genetic resources in their national ABS regulations.

To labour the point: for *ex-situ* access, it is necessary to identify the original sampling person/institution and/or the country from where the genetic resource was collected, and the date the genetic resource was collected (the entry into force of ABS regulations changes from country to country as well as their requirements on retrospective compliance with access regulations). If that country is signatory to the NP (or has separate ABS regulations), you need to check that the necessary permits were obtained by the sampler *before* you can utilise that material.

The onus of exercising due diligence in checking whether access and use of GRs is in, or out of scope of the EU ABS Regulation, and to obtain all necessary permits*, is on the user*; but the onus is on the nations to set up the processes by which this can be established. Any country that has ratified the NP should set up the necessary focal points/ABS offices which you can consult (see the information on the ABSCH site). If a country has not ratified the NP, these permits are not necessary, *however* they may have other ABS/sampling permits that are necessary, and this will also need to be checked.

If material must be transferred after being sampled/taken from a collection, an additional set of material transfer agreements will be necessary (the Material or Data Transfer Agreements). Transferring material is not making use of it, *and it is the receiver (and eventual user) of the material that requires the ABS permits*. The purpose of these documents is to provide legal certainty regarding transfer of ownership or custodianship, including all associated documents and necessary information. Users are also advised to keep a record of all benefits that have been shared in relation to the GR (both monetary and non-monetary).

In order to better understand the scope of the NP, it is helpful to understand some of its key terms. While many of the terms are not, in fact, *precisely* defined in the NP, we have pieced together definitions obtained from various ABS and EU documentation, as well as the text of the NP itself. Particularly useful again is the Horizontal guidance document on scope of the Basic Regulation (EU) No511 / 2014 and the Implementing Regulation (EU)2015/1866 (Official Journal, *12 January 2021)*, which contains many clearly-written definitions, explains the scope of the various terms, and gives useful examples. While this resource is a report of the EU and does not aim to cover the NP as applied in other countries, it is nonetheless useful for anyone to consult.

4 What do you have to do to comply?

To enact the NP at the local level, national focal points (NFPs) and competent national authorities (CNAs) have been (or should be) established by the parties, to serve as contact points for information, to grant access or to cooperate on issues of compliance. NP-related documentation is stored at the ABS Clearing House (ABSCH), allowing for the sharing of information such as domestic regulatory ABS requirements or information on NFPs and CNAs. This means that all PICs and MATs obtained and/or negotiated with the CNAs from the country where the GR originates, are archived in the ABSCH: when the CNA of a Providing Country publishes a permit or equivalent (e.g. PIC and MAT) on the ABSCH, this is registered and given a unique identifier (Internationally Recognised Certificate of Compliance: IRCC) by the ABSCH and provides legal surety of the GR covered, and assists in the monitoring of its utilisation. Note that it is always necessary to request these permits from the country from where the GR originates. If no permits are required (if the proposed use "falls out of scope"), *it is still necessary* to retain the written instructions (e.g. email communication) to that effect from the NFP/CNA. It is always necessary to exercise due diligence (see Sec. 5 for more detail on this point). Documentation from the country about applicable access and benefit-sharing measures should be provided on the ABSCH and these can be consulted to check against your proposed desired use. If there is no documentation provided on the site, then the NFPs and CNAs should be contacted instead.

According to the EU (the *Horizontal guidance document*), if reasonable attempts have been made to contact an NFP, but no answer could be obtained, the potential users need to decide for themselves whether or not to access or utilise the GR in question. If it is later established that the GR falls "in scope" of the NP and that NP-permits should have been obtained, the user will then need to obtain the permits or discontinue use. In other words, as long as it is possible to prove "due diligence", i.e. all attempts were made to establish contact/obtain permits and proof of this can be produced, a user can decide to continue (or not) with the access or utilisation. *This is admittedly a bit of a grey area.*

In making data arising from the use of GRs FAIR, EOSC-Life should be concerned with the traceability of the use of GRs and the data generated during the process, ensuring that at each step – from acquisition of biological material (in the wild or from a collection), through its processing and storage, to the data generation and analysis – the handling of that data is NP-compliant. In the current ABS regulatory landscape this is always in connection to the actual physical genetic resource whose utilisation is regulated to fulfil the track and trace requirements of the NP – i.e. this track and trace documentation always needs to be connect to the physical genetic material that is utilised. However, future developments on access and utilisation of DSI may add another layer into the tracking and tracing of data generated from genetic resources. In this respect and importantly, the INSDC recently announced a new policy

(https://www.insdc.org/spatio-temporal-annotation-policy-18-11-2021) requiring spatiotemporal information, demonstrating their commitment to scientific transparency and openness to change. A quick summary of the steps and the permits and consents that must be granted when utilising genetic resources and from whom is available in the EMBRC guide: A step-by-step guide to ABS compliance when utilizing marine genetic resources. The next paragraphs explain those steps.

Prior to sampling

1. Notify the primary NFP of the provider state: this will be the "NFP on the Convention on Biological Diversity" or the "NFP on the ABS for the Nagoya Protocol" Some countries allow free access to their GRs, while others require that permission is sought. A list of the countries concerned with the NP/ABS can be found on the ABSCH site. A more detailed explanation of the role of the NFPs can be found on this page of the ABSCH site, and country profiles including contact details can be found via this page. A list of designated competent authorities in the EU can be found via this page on the EC site.
   If you consult a Focal Point, *make sure to document the conversation* and link this information to the specimen records within your data management system. This is important for traceability: if permits are required then these will be stored in the ABSCH and traceability is done through them, but if none are required it is the conversation that must form this part of the traceability.
2. Ask the NFP for advice on the specific requirements to be fulfilled prior to sampling activities. In certain cases, those requirements may be simple (notification of the sampling), in other cases, they may include a Prior Informed Consent (PIC) and an ABS negotiation defining Mutually Agreed Terms (MAT).
3. When asking for your PIC and MAT, it is important to remember that *only* what is set out in these agreements is covered, e.g. the science you want to do must be specified, and *only* that science can be done. For example, some countries allow you to take DNA sequences from the sample after collection, while for others this must be specified in a new permit or may only be done within its borders. It is recommended to include details of all R&D activities you wish to carry out with/from the sample. You should also specify who will be doing this: bear in mind that the MAT and PIC are *bilateral* agreements, they are only between the parties signing them and only covers what is specified therein. Parties can be individuals or institutes.
4. After negotiating and obtaining the PIC and MAT, these are registered in the ABSCH and given an IRCC code. As the PIC and MAT are private documents, only the title and the details of the signatories is given in the IRCC. Keep the IRCC code together with your copies of communication and your sample information.
   If it was not necessary to obtain a PIC and MAT, keep digital copies of the communication to this effect. This is part of what we call the "alternative" traceability information. Keep these somewhere where they can always be accessed upon request. More detail about what other pieces of information you should retain can be found in Sec. 5.

Post-sampling

5. The sample (the GR) must be identified with a unique code. Identifying samples is common practice in science, and choosing the identifier (e.g. MaterialSampleID) can be done using whatever system you commonly use. This sample ID should be kept with the physical sample and with any kind of digital information about it, obtained from it, and with any derivatives thereof.

6. Ensure that the following information are recorded and kept together with the sample and with the digital form(s) it takes (e.g. when stored in a collection): the date and location of sampling (longitude and latitude), the unique sample ID, and the IRCC code. If no IRCC was necessary (i.e. no written permits were required), it is still necessary that the sample ID can be linked to the alternative traceability information (see point 1). For more on these recommended metadata, see Sec. 5.

Data archiving, and data or scientific publication

7. Any data publication or scientific paper or report published by users as consequence of the utilisation of the GR should quote the IRCC code (or state that permits were not necessary because the use was “out of scope”) and it is good practice to also give the ID of the sample in whatever collection/biobank it is stored (e.g. in the acknowledgements or “used resources” section). Any secondary publication or scientific paper or report that is based on data or results obtained from the GR is advised, and it would be considered good practice, to quote the IRCC code for the utilisation of the original GR and provide the ID of the sample if available in any collection/biobank.
8. When the data coming from a GR are archived (e.g. placed in a public data repository) and a metadata record of the dataset created in a public data catalogue (so the data are findable), it should be ensured that that the NP-necessary metadata are also included (point 6).

Transfer and re-use of the GR (i.e. the sample)

9. The MAT that was obtained will be *specific to that sample and for the use indicated in the MAT* (name of users, what kind of processing, storage, etc). Any change of use (intent) *must* be re-negotiated. To be able to re-use the sample, either by someone not specified on the original MAT or for a different purpose, it will be necessary to negotiate this with the sovereign country from where the sample (the GR) was obtained. As the MAT is a private document, it will probably be difficult to establish when you are a new user/wish to make a new use of the sample. Consultation with the originating country will be necessary. If the necessary metadata (e.g. longitude and latitude of the sample site, IRCC code) are not included with the sample (as well as with its digital version), it will be very difficult to do this!
10. Similarly, if the sample has to be transferred to a third party (for instance, for bio-sequencing or archiving), or if the data and scientific results are going to be published, this can only be done in accordance with the requirements set out in the permit/ABS agreement. In this case, a Material or Data Transfer Agreement must be signed (MTA/DTA).
11. If a new set of permits is negotiated, the original and the new metadata should be kept with the sample and its digital version.

Re-use of the derivatives of the GR (e.g. sequences)

12. Currently the NP does not consider the sequences derived from a sample (a GR) to be subject to the NP, and hence those using sequences, e.g. as held in public archives, are not required to exercise due diligence or to carry forward the NP-related metadata. The sequence taken immediately from the sample must ensure due diligence was met and all necessary metadata are recorded, but subsequent users of the sequence do not have to. **However, this is currently under consideration by the parties to the NP and may well change in the near future**. It is therefore good practice to at least ensure that all necessary metadata are recorded with all archived sequences, and the onus here would fall on the archives themselves.

For EU countries, a web-based application, DECLARE, was created to allow users' submission of due diligence declarations online. DECLARE is an EU-wide web-based tool which enables users of genetic resources to submit the due diligence declarations required by Article 7 of Regulation (EU) No 511/20142 (the EU ABS Regulation) and further specified in the Commission Implementing Regulation (EU) No 2015/18663 to the relevant competent authorities responsible for their implementation. The competent authorities also use DECLARE to transfer non-confidential information from the due diligence declarations to the ABS Clearing House. Links to the tool and its documentation can be found from here. Some countries do not use DECLARE and have their own specific platform for the user due diligence declaration.

A flow-chart describing the steps to take along the path-of-use of GR is given in Fig. 1 (see end of report), and Table 3 lists the documentation, permits, actions, or metadata steps.

5 Recommendations on provenance metadata

These are the ABS-relevant information that should be kept with any material samples and digital data, to prove due diligence with respect to ABS. These recommended metadata should be provided *whether or not you actually had to obtain ABS permits yourself*: these metadata will provide the proof that you did check (you did your "due diligence") and to allow a subsequent user of the material to know if they should apply for ABS for whatever utilisation *they* will want to make of that material.

1. Necessary: Name and affiliation of the collectors/users of the GR, including at least one set of contact details
2. Necessary: The IRCC code of the documentation (usually the MAT and PIC) stored in the ABSCH. If no ABS was necessary, you need to specifically indicate this fact ("no ABS necessary").
3. Necessary: If no ABS application was necessary, i.e. your use of the GR "falls out of scope", or the country is not a signatory to the NP, and no actual permits (and no IRCC code) exist, then keep copies of all communication at hand or other kind of proof (e.g. copies of the documents that state that your use falls out of scope or that the country has not signed up to ABS) that show the due diligence was exercised. You will need to indicate that such documentation exists and give the necessary contact details (or a URL of a webpage where those documents can be found/requested).
4. Necessary: The unique sample identifier that the original sample material was given (following whichever protocol is normal for the sampling institution), and the name of that sample collection or name of the source institution/team/person from which the resource was directly obtained (so it can subsequently be located by anyone)
5. Necessary: A description of the GR (e.g. the "genus / species" if known, or "unidentified resources", or "environmental water sample") and ideally also the source of the GR (i.e. from what part of the sampled organism did the GR come from)
6. Necessary: Spatiotemporal information of access of the GR (date and country, but if possible longitude and latitude)
7. Necessary: The reference or IDs of other permits (e.g. material or data transfer agreements)
8. Necessary: A description of the scope of the utilisation that was requested (whether or not a permit was actually required)
9. Useful: Is this sample subject to the NP: yes or no

6 Which EOSC-Life partners are affected?

The EU ABS Regulation does not cover human samples or model organisms produced and utilised in laboratories. Listed in Table 2 (see end of report) shows the level in which the ABS requirements affect the EOSC-Life RIs at this point in time, following an online survey conducted in March 2022. However, it is *always* the responsibility of any researcher, project, RI, to check this for themselves. Bear in mind, RI or their service providers can be directly affected by ABS requirements, when they provide services that can be considered as a utilisation of GR under the EU ABS Regulation (such as extraction of proteins or DNA, their sequence determination, or the identification of bioactive compounds). Thus, to consider services to be out of scope, the RI and user's liabilities shall be well defined in a service agreement (*Horizontal guidance document.* sec. 3.5.2). In addition, RIs are indirectly affected by ABS requirements when supplying biological material. Thus, RI shall seek, keep and transfer the provenance information associated to samples supplied, and as a good practice to DSI supplied, in order not to cut the traceability chain from sampling to utilisation. Otherwise, RI users would not be able to comply with the ABS requirements.

7 Practical aspects: an EMBRC use case

The European Marine Omics Biodiversity Observation Network (EMO BON) was launched in 2021, and is a European initiative from the European Marine Biological Resource Centre (EMBRC) to establish a coordinated, long-term biodiversity observatory. The goal for EMO BON is to support the individual marine biodiversity observatories within EMBRC and connect them under one centrally-coordinated network, with shared protocols, data, and metadata standards. EMBRC aims to build a long-term genomic observatory, generating cost-effective, high-quality, baseline genomic biodiversity data that will be produced in the long term. EMO BON includes marine stations from polar regions to the Red Sea that will sample for genomic marine biodiversity at frequent intervals. This network will contribute to the United Nations Decade of Ocean Science for Sustainable Development and aims to be a component to the global ocean observation networks.

Collection of marine water, sediment and organisms takes place every few months at 16 EMBRC participating observatory stations according to predefined protocols in the EMO BON Handbook, setting a minimum standard for participation to the network. The EMO BON Handbook is available from the Ocean Best Practices System (OBPS) repository and it may also be found on the EMO BON webpage. The samples (plus accompanying paperwork and metadata) are sent by each station to EMBRC headquarters, and from there to the sequencing facility (Genoscope) for DNA extraction and sequencing. The data (metabarcoding and metagenome data) generated within this initiative will follow the FAIR data principles.

A description of the provenance flow of this use case has been created under EOSC-Life WP6, as part of its distributed provenance model work. A preliminary version was created for an EOSC Life FAIR Hackathon for the Open Call projects and the presentation can be found on the hackathon drive; the current repository for this model (which is still a work in progress) can be found on Github). This covers the flow of data from sampling preparation to data publication, and it will include components related to the access to the genetic resources and to the mutually-agreed terms on their utilisation in relation to ABS regulations, as they are dealt with in this document. The flow for this use-case (Figures 2-5) has these components:

1. **Sampling preparation**: SOPs and permits (including ABS)
2. **Sampling activities (field and lab), sample processing, biobanking and shipping between institutions**: collecting, processing, storing material; measuring and collect environmental parameters (shared spreedsheets)
3. **Sequencing and archiving of sequence data and metadata**: material turned into sequences (digital data) archiving in ENA and describing the sequences, gathered metadata.
4. **Publishing**: primary publications by the users of the genetic resources, secondary publications of DSI users.
5. **Re-using of the genetic material**

In relation to the NP, the type of data to be managed are permits that must be linked to the GRs during their life, and as a general recommendation beyond their life (in databases). In this use case, only two countries (France and Spain) out of nine (eight European countries and Israel) in the EMO-BON project have implemented restrictions related to the access and utilisation of physical genetic resources. However, in Spain those restrictions do not concern the material accessed only for taxonomic purposes. Only the Spanish ABS focal point strongly requests (suggests) including the country tag with any DSI generated from Spanish EMO-BON GRs when uploading to public databases. In all countries certificates of compliance (in form for instance of e-mail communications with national Nagoya Focal points indicating that proposed utilisation is out of scope) is present, in addition to sampling permits and material transfer agreements among partners in the process.

In the scheme of EMO-BON provenance flow these are the research institutions (actors) that are involved and their duties along the process:

- **Local sampling institution or station**: sampling, obtaining sampling permits, biobanking, shipping samples for utilisation
- **EMO-BON HQ**: user of GRs, obtaining ABS papers, centralising the reception of sample from local sampling institutions to resend them to the sequencing service.
- **Genoscope**: subcontracted sampling centre that processes the samples and submits sequence data to ENA
- **ENA (European Nucleotide Archive)**: archiving of sequences and biosample associated metadata.
- **Re-users**: users requesting access to the biobanked GRs.

These are the master documents and archives along the process where all the provenance data and information is archived as it is generated along the process:

- **Shared EMO-BON spreadsheet**. This master document is shared by the EMO-BO HQ and the sampling institutions and incorporates all the data (sampling logsheets with spatiotemporal information, physicochemical measurements, permit documents and IDs, sample IDs,…) related to the GRs along the process as it is made available.
- **Biobank catalogues**: These catalogues or records of the biobanked material are dependent on each sampling institution, who is holding the material and the associated information during the life of the GR as this is stored in its biobank.
- **ENA EMON-BON archives**: Biosample file containing all the DSI coming from each GR accessed from every sampling site and over the entire projects, with its associated sample metadata.

Together with these are the following documents: standard operating procedures, sampling and sequencing protocols that are specific to the EMO-BON use case, and the EMO-BON handbook.

The flow of activities for the EMO BON use-case, with the actors and the particular ABS steps highlighted, are shown in Figs. 2 to 5.

Tables and Figures

Table 1. Overview of conditions for applicability of the EU ABS Regulation[1] conditions taken from the Guidance document on the scope of the EU Regulation No511/2014

		Within scope (cumulative conditions)*	Outside scope
Geographic scope (provenance of GR)	*Access in ...*	Area within a country's jurisdiction	Area beyond national jurisdiction e.g. high seas or covered by the Antarctic treaty system
	Provider country is …	Party to the Nagoya Protocol	Not a party to the Protocol
	Provider country has ...	Applicable access legislation	No applicable access legislation
Temporal scope	*Access ...*	On or after 12 October 2014	Before 12 October 2014
***	*Utilisation ...*	On or after 12 October 2014 (recommend to assume this even when the access was prior to this date)	Before 12 October 2014
Material scope	*Genetic resources ...*	Not covered by a specialised ABS instrument**	Covered by a specialised ABS instrument**
		Non-human	Human
		Obtained as commodities but subsequently subject to R&D	Used as commodities
	Utilisation ...	R&D on genetic and/or biochemical composition	No such R&D
Personal scope		Natural or legal persons utilising GR	Persons *only* transferring GR or commercialising products based on it
Geographic scope (utilisation)	*R&D ...*	Within the EU	Exclusively outside of the EU

* To be within scope, all conditions must be fulfilled
** e.g. certain pathogens and genetic resources for food and agriculture, see ABS Clearing House and National Focal Points
*** This raw has been added to the official table taken from the Guidance document on the scope of the EU Regulation n°511/2014, annex I

1 Commission Notice (2016/C 313/01).

Table 2. List of EOSC-Life consortium RI, domains of activity, and whether they are affected by the NP as providers of GRs (data gathered through a survey shared with all EOSC-Life consortium RI. All 13 involved RIs responded.

RI	Domain of activity	Affected by NP?
BBMRI	Biobanking & biomolecular resources (www.bbmri-eric.eu)	Yes
EATRIS	Translational research (www.eatris.eu)	No
ECRIN	Clinical trials (www.ecrin.org)	No
ELIXIR	Curated databases (www.elixir-europe.org)	Yes
EMBRC	Marine Biology and Ecology (www.embrc.eu)	Yes
EMPHASIS	Plant phenotyping (emphasis.plant-phenotyping.eu)	No
ERINHA	Highly pathogenic microorganisms (www.erinha.eu)	Yes
EuBI	Biological/medical imaging (www.eurobioimaging.eu)	No
EU-Openscreen	Screening & medical chemistry (www.eu-openscreen.eu)	No
INFRAFRONTIER	Functional genomics (www.infrafrontier.eu)	No
Instruct	Structural biology (www.structuralbiology.eu)	Yes
ISBE	Systems biology (www.isbe.eu)	Yes
MIRRI	Microorganisms (www.mirri.org)	Yes

* **Providers of GRs under definition of NP**: EMBRC, MIRRI, ELIXIR, ERINHA, EATRIS
** **Providers of microbiomes (humans, model organisms)**: ELIXIR, BBMRI, EATRIS, ISBE
*** **Users of GRs**: EMBRC, MIRRI, ELIXIR, ERINHA, EATRIS, BBMRI, EMPHASIS, INSTRUCT
****ABS protocols in place**: EMBRC, MIRRI, ELIXIR, ERINHA, EATRIS, EMPHASIS,
***** **Do not utilise DSI**: EU-Openscreen, ECRIN, EMPHASIS, INSTRUCT, EuBI

Table 3. Documentation, permits, actions, and metadata necessary for the NP

When	What	Why
At the stage of applying for EU research funding	*Action* DECLARE (EU only)	DECLARE is an EU-wide web-based tool which enables users of genetic resources to submit the due diligence declarations. Due diligence declaration is obligatory but some countries implemented their own platform for declarations.
Before sampling	*Permits* PIC, MAT, IRCC	If the sampling requires ABS/NP/collecting permission. Documents may cover: what you can sample from where and when, what you will do with the samples to store them, the R&D that will be performed, and where and by whom
	Documentation Proof of due diligence	If you checked and were told no permits were necessary, keep copies of these communications If you have tried very hard and could not get a response from the country where your sampling should take place, it is a grey zone whether or not to continue with the sampling. In any case, keep copies of all communication attempts.
After sampling	*Metadata*	Ensure you have the IRCC code, or a link to proof of due diligence, collection location and data
During storage	*Metadata*	Add to the existing metadata the unique sample ID and collection name (so it can be identified by anyone else)
Before shipping (transfering) samples	*Permits* MTA	If you need to transfer samples to somewhere or someone else, a MTA may be necessary: the samples and all associated documentation must be transferred.
Performing R&D (e.g. sequencing, extracting…)	*Metadata*	The physical or digital versions of the R&D done to the sample must include all the collected metadata
Before using a sample from a collection	*Permits* MAT, PIC, IRCC	If the use (R&D) you wish to make of the sample is not covered by the original MAT, and/or if you are not included as a permitted user in the original MAT, you will require a new set of permits from the country from where the sample originates. If your use and yourself are covered in the original MAT, you do not need new permits.
Before issuing a patent or completing a product	*Action* DECLARE (EU only)	DECLARE is an EU-wide web-based tool which enables users of genetic resources to submit the due diligence declarations. Due diligence declaration is obligatory but some countries implemented their own platform for declarations.

Figure 1. Flow chart to show the steps involved in sampling and utilising GRs

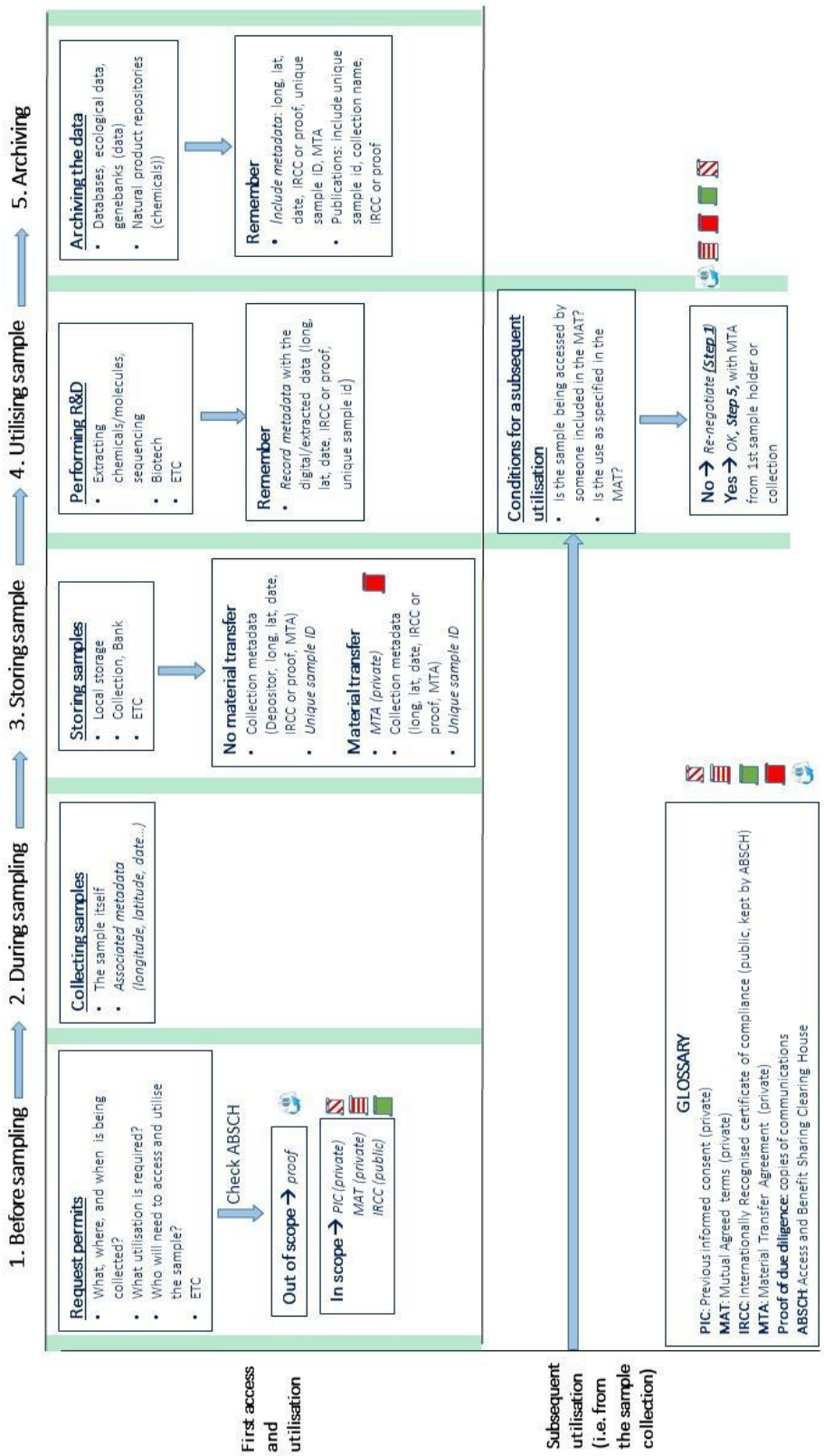

Figure 2. This provenance flow shows the pathway followed by ABS-linked data (pink frame) before the beginning of the sampling (EMO-BON use case, section 7) linked to the steps in section 4 and the recommendations in section 5 of this document.

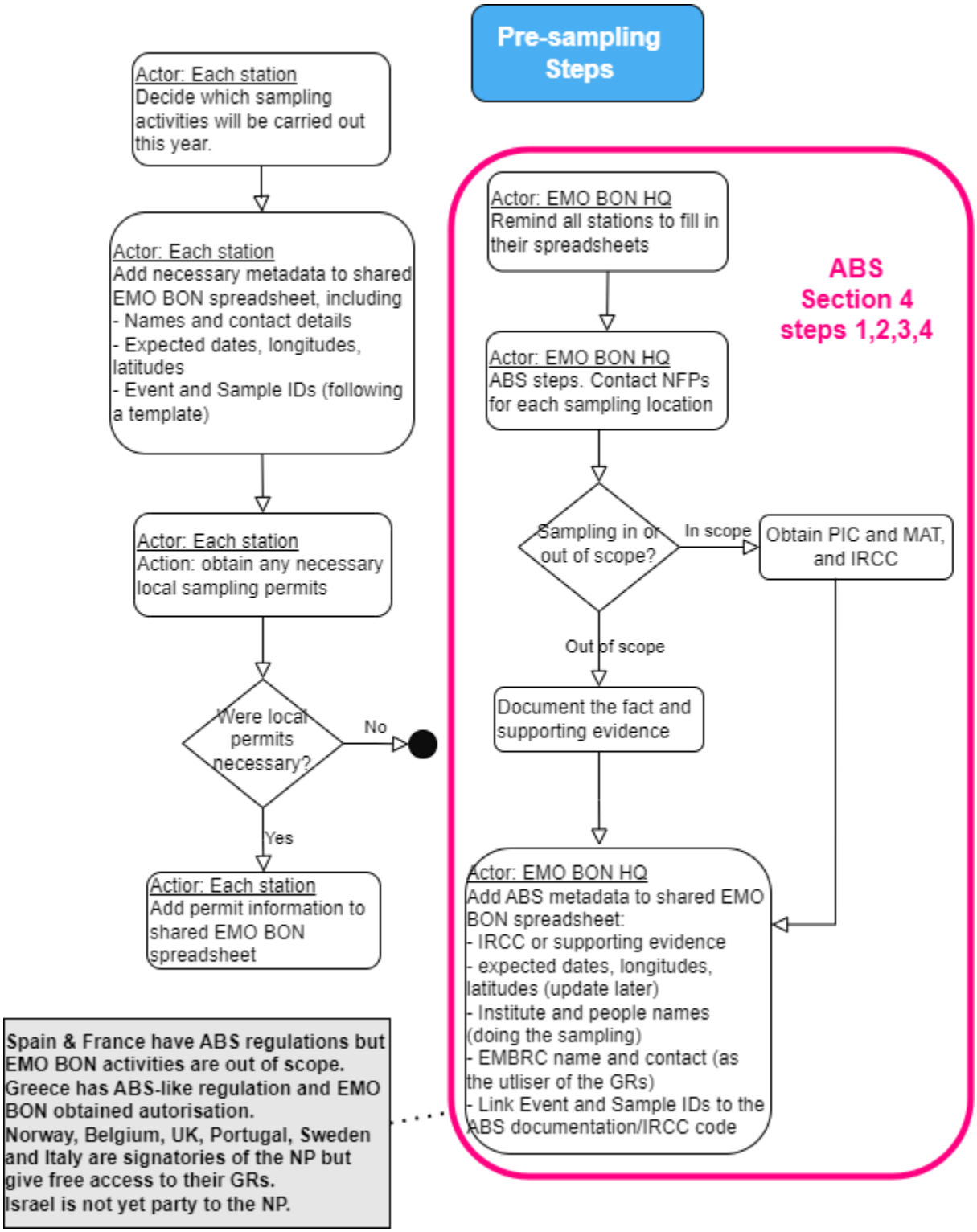

Figure 3. This provenance flow shows the pathway followed by ABS-linked data (pink frame) during sampling, sample processing, sample biobanking and sample shipping to the sequencing centre (EMO-BON use case) linked to the steps in section 4 and the recommendations in section 5 of this document.

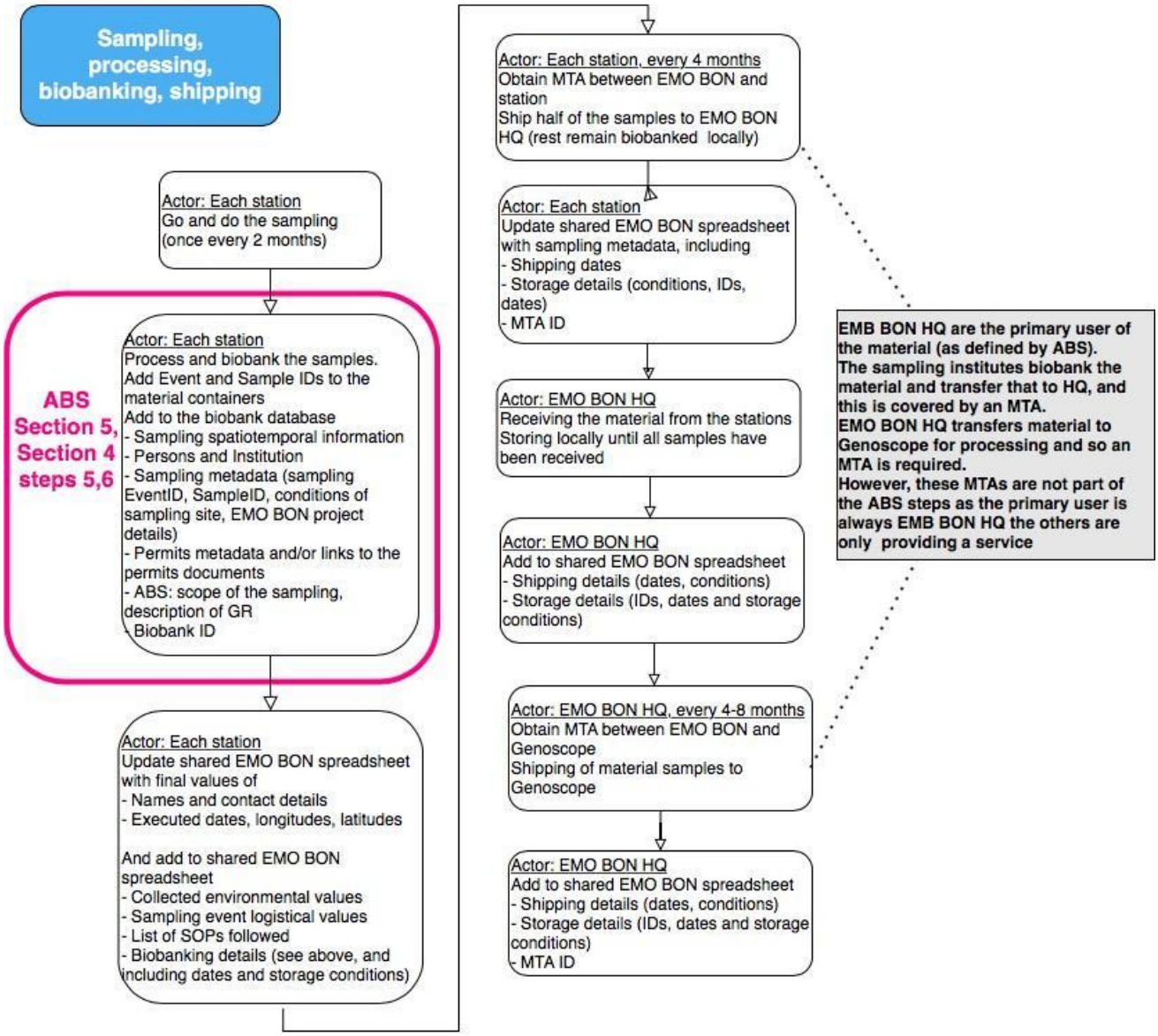

Figure 4. This provenance flow shows the pathway followed by ABS-linked (pink frame) data during sequencing in the subcontracted sequencing centre and archiving of the sequence data and metadata in ENA (EMO-BON use case) linked to the steps in section 4 and the recommendations in section 5 of this document.

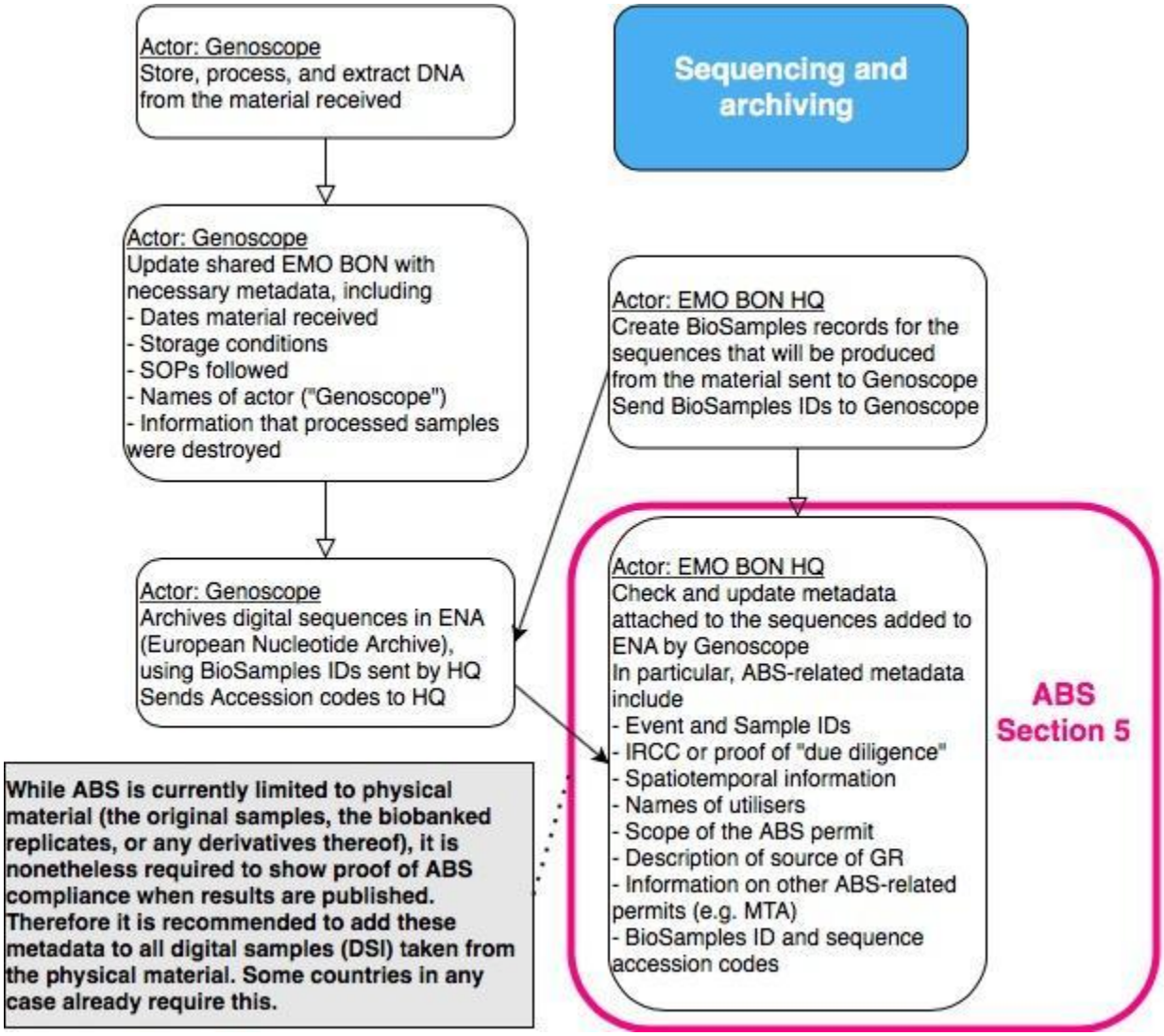

Figure 5. This provenance flow shows the pathway followed by ABS-linked data (pink frame) during the publishing process and the steps to be followed for re-utilization of the genetic resource (EMO-BON use case) linked to the steps in section 4 and the recommendations in section 5 of this document.