

Annex 5. Topic 5: Innovative molecular techniques for taxonomy: integrating genomic tools for the development of cost-effective genetic markers for species identification and delimitation

1. Challenges to address by the Proposals

Recent advances in the development of genomic tools that screen genome-wide variation in non-model organisms allow taxonomists to identify and delineate species and discover new species at an unprecedented level. These methods are expected to have a strong added value when traditional morphological or standard single-locus genetic markers (e.g. COI, chloroplast genes) provide insufficient resolution to discern species within complexes of closely related species. Especially in biodiversity hotspots, the undetected presence of such species complexes can underestimate their biodiversity. Although current genomic techniques in principle allow for accurate detection of previously unrecognised species or species relationships, the application of these tools requires considerable expertise that is still poorly integrated within taxonomic research. Therefore, it remains challenging for many taxonomists to select and apply the most appropriate and cost-effective method from the plethora of technologies currently available and to gain the required molecular and bioinformatic expertise. TETTRIs aims to address this issue by facilitating the development of innovative, easily applicable and reliable genomic marker sets, targeted to a specific taxonomic group, that allow taxonomists with more limited genomic and bioinformatic experience to discover species complexes and identify species in European biodiversity hotspots and protected areas with maximal resolution. In particular, TETTRIs experts are developing multilocus amplicon sequencing pipelines, which allows to screen custom designed multilocus marker sets using e.g. highly portable and easy-to-use genome sequencers. Complementary to this, additional high-resolution approaches based on genome-wide screens using RAD-Seq or low-coverage sequencing that allow to delimitate particularly complex taxonomic groups are also under development by TETTRIs.

2. Practical impact on site

To further advance genomic marker sets, the envisaged methods need to be previously tested and validated. TETTRIs supports applications that aim to implement more advanced genomic and bioinformatic tools currently under development by TETTRIs experts. A major goal of this call is to test and validate a user-friendly multilocus marker sequencing pipeline that facilitates taxonomic research of species complexes in European biodiversity hotspots and protected areas. Target groups include species complexes of plants and insects, with a priority for pollinators, in European biodiversity hotspots and protected areas.

Among other topics, the implementation of these new methods will be articulated in the project.

Selected projects will:

- implement and validate, in close collaboration with the TETTRIs consortium, a molecular and bioinformatic protocol to develop a multilocus marker set for the identification of species in taxonomic complexes. Preferred taxa include species complexes of plants (in collaboration with BG Meise) and insects (in collaboration with RBINS and NHM Vienna), with a priority for pollinators, in biodiversity hotspots.
- validate the resolution of the multilocus marker set by comparing the obtained results with those from more advanced tools that screen genetic variation, e.g. single nucleotide polymorphisms (SNPs) at a genome-wide scale (e.g. low coverage whole-genome sequencing, genotype-by-sequencing, RADseq). Application of these advanced genomic tools will be supported by the TETTRIs consortium.
- Implement and validate an on-line user-friendly species delimitation platform, developed by NHM Paris within the TETTRIs consortium, based on the results

obtained from the multilocus marker set and SNP data as well as pre-existing morphological and/or genetic data.

- contribute to the development of comprehensive guidelines to disseminate the methodology to other taxonomists.
- generate an example reference dataset that can be used in workshops and courses on genomic data analysis in a taxonomic context.
- demonstrate that the implemented protocol improves the taxonomy of the proposed species complex in a biodiversity hotspot.

3. Collaborative Approach (compulsory activities)

Proposals will involve collaboration between the applicant and TETTRIs specialists (from WP6). By working together, both parties will benefit from the mutual expertise and knowledge, resulting in a more comprehensive and effective development of the molecular pipeline. TETTRIs specialists will facilitate lab work by providing lab and bioinformatic guidelines. Lab work and bioinformatic analysis will be carried out by the applicant in close collaboration and within the labs for molecular systematics of the TETTRIs partners.

The exchange of information will then:

- provide a testing mechanism that evolves with the intermediate results obtained in the application of the methods,
- postulate potential enhancements and improvements of the tested methods,
- articulate in-situ modifications to accommodate the specificities of the lab platforms where the application will be used, and
- ultimately, facilitate the development of a thoroughly tested and validated documented guidebook by the TETTRIs partners.

4. Innovation

Some complementary expertise to increase the innovative aspects of the outcomes may be added to the Proposal. Examples include:

- Specify why the use of genomic tools is expected to leverage the taxonomic knowledge of the focal species complex and why current morphological traits and more classical genetic markers are currently insufficient to classify and identify species within the proposed species complex.
- Putative proposals for additional innovations in molecular techniques such as (but not limited to):
 - new innovative complementary bioinformatics tools specifically tailored for genomic data analysis in a taxonomic context. This could involve the use of artificial intelligence and machine learning algorithms to improve the accuracy and efficiency of species identification.
 - Exploration of potential emerging cost-effective tools for genomic library preparation or sequencing that are relevant for taxonomic exploration.
- A plan how the project results will be disseminated to the local/amateur taxonomists that are less familiar with the interpretation and use of molecular tools for taxonomy (e.g. through taxon specific symposia, societal journals, newsletters,...).
- Interlink the outcomes of the Proposal to improve the reference collection of pollinators (Topic 1).

5. Expected outcomes from the Proposals

Based on the applications, thorough guidelines and workflows on how to use the different approaches will be generated and integrated in the outcome of the TETTRIs project. To that end, the expected outcomes of the 3PPs are:

- a proposed set of multilocus genomic markers obtained using a molecular genetics and bioinformatics protocol developed by the TETTRIs consortium

- sequence data from this multilocus marker set of multiple species and individuals of the focal species complex.
- a genome wide SNP dataset, partly obtained from the same individuals, that will be used to compare the resolution of the multilocus markers set with those of genome-wide SNP variation.
- an in-depth taxonomic clarification of the focal species complex by integrating the results obtained from morphological, multilocus and SNP data using the on-line species delimitation platform developed by NHM Paris within the TETTRIs consortium.

6. Specific conditions

Expertise and experience:

The applicant should have:

- Demonstrated experience with basic molecular lab work and bioinformatics (DNA extraction, PCR, library preparation, basic genomic data analysis...),
- demonstrated taxonomic knowledge of the proposed taxonomic group in biodiversity hotspots.

Supporting elements:

Required in the project description – the applicant shall:

- Collaborate and/or be integrated in a network of local (amateur) taxonomists to perform its taxonomic/faunistic work.
- Specify why increased taxonomic knowledge of the focal taxonomic group(s) is relevant for the conservation of biodiversity hotspots.
- Provide all necessary material. A basic set of specimens suitable for molecular analyses should be available, but a collection of additional specimens suitable for molecular analysis can be included in the project.
- Provide an overview of the currently available genomic resources of the species complex.

General Instructions to applicants:

To be considered for funding, proposals should clearly address all three key aspects of the topic, i.e. 1) demonstrating a strong practical impact, 2) implement a collaborative approach, and 3) integrate innovative dimensions to biodiversity identification, monitoring and/or conservation.

Proposals should clarify how they will accommodate the genomic and bioinformatic expertise requirements during the execution of the project.

Proposals should provide a detailed budget and timeline, as well as clear metrics for measuring project success.

Applicants should also demonstrate relevant experience and expertise in areas such as biodiversity science, citizen science, data analysis, and stakeholder engagement.

Data and sequences generated during the project should be published openly. Projects must comply with all national and international regulations regarding access and benefit sharing of genetic resources.

Citizen science aspects can be involved in all topics. However, proposals focused primarily or exclusively in Citizen science engagement will fall under Topic 7.

See the Call text for further detailed information.