

FIAS REGISTRY OF RESEARCH SPECIFICATIONS

Minimum Replication and Decision Rules for Presence/Absence Calls from Soil eDNA Metabarcoding

REGISTRY IDENTIFIER

FOAR-CON2X-KHW78203P

OBJECT TYPE

Procedure or protocol

DOCUMENT MATURITY

Stable

PUBLISHED

21 September 2015, 12:00 UTC

AUTHORITATIVE VERSION

Registration records the outcome of the Registry review for this version: the submission was sufficiently identified, documented, and internally determinate for the stated procedure. Scope, adoption, and competing conventions are documented separately where relevant.

VERSION

1.0

SUBJECT AREA

Biology and biodiversity

SUBMITTED

06 July 2015, 12:00 UTC

LICENCE

Creative Commons Attribution 4.0

REGISTRY STATUS**Accepted for registration**

21 September 2015, 12:00 UTC

Core description

SUBMISSION KIND

Revision of an unregistered submission

EARLIER SUBMISSION OR REGISTRY IDENTIFIER

FIAS-SUB-2013-BIO-01

FULL TITLE

Minimum Replication and Decision Rules for Presence/Absence Calls from Soil eDNA Metabarcoding

SHORT TITLE OR ACRONYM

Soil-eDNA Calls 1.0

OBJECT TYPE

Procedure or protocol

STATUS CLAIMED BY THE SUBMITTER

Project standard

MATURITY OF THIS VERSION

Stable

VERSION

1.0

PRINCIPAL SUBJECT AREA

Biology and biodiversity

KEYWORDS

- environmental DNA
- soil metabarcoding
- PCR replication
- occupancy model
- false positives
- presence/absence

PROBLEM, SCOPE, AND EXCLUSIONS

Version 1.0 replaces the fixed triplicate/single-positive decision in 0.8 with an error-aware protocol. It governs how a soil-eDNA biodiversity project chooses PCR replication, preserves each replicate, and reports present, absent, indeterminate, or invalid calls. It excludes abundance inference, field sampling, primer design, clinical material and taxonomic revision.

INTENDED USERS AND CONTEXTS

- Soil-eDNA biodiversity survey teams planning laboratory replication
- Analysts fitting detection and false-positive occupancy models
- Data curators reproducing categorical calls from replicate evidence

RELATION TO EXISTING PRACTICE

Replaces an earlier specification

SCIENTIFIC CONTEXT, SOURCES, AND SCHOLARLY DISCUSSION

This section describes the scientific background, underlying work, and sources of the specification, together with the research groups and scholarly settings in which the proposed solution or its formalisation was presented or discussed. It documents the context and nature of these scholarly exchanges.

Environmental DNA provides a way to study biodiversity through molecular traces retained in environmental samples. P. Taberlet et al. (2012) establish the methodological setting for this approach. B. R. Schmidt et al. (2013) and D. A. W. Miller et al. (2011) show how repeated observations can be analysed while accounting for imperfect detection and false-positive error. These findings make the probability of detecting a taxon, the evidence retained for each replicate and the interpretation of negative results central to a reproducible presence/absence protocol.

G. F. Ficetola et al. (2015) examine replication levels and erroneous presence calls in eDNA metabarcoding. Their results link the required replication effort to detection probability and motivate greater replication in low-detection settings. Version 1.0 incorporates this reasoning into the selection of PCR replication and the recording of present, absent, indeterminate and invalid outcomes. Retained observations and controls allow each categorical call to be regenerated from the evidence and the declared decision rules.

The scientific setting is the environmental-DNA research developed around LECA, including work by G. F. Ficetola, J. Pansu, A. Bonin, E. Coissac and P. Taberlet. This collaboration connects molecular biodiversity methods with ecological questions about detection and prevalence. The protocol makes that methodological relationship explicit for soil metabarcoding projects, supporting comparison of results and examination of uncertain calls across research teams.

References:

- Taberlet, P.; Coissac, E.; Hajibabaei, M.; Rieseberg, L.H. (2012), Environmental DNA, *Molecular Ecology* 21: 1789–1793. DOI: <https://doi.org/10.1111/j.1365-294X.2012.05542.x>
- Schmidt, B.R. et al. (2013), Site occupancy models in the analysis of environmental DNA presence/absence surveys: a case study of an emerging amphibian pathogen, *Methods in Ecology and Evolution* 4: 646–653. DOI: <https://doi.org/10.1111/2041-210X.12052>
- Miller, D.A.W. et al. (2011), Improving occupancy estimation when two types of observational error occur: non-detection and species misidentification, *Ecology* 92: 1422–1428. Source: <https://pubs.usgs.gov/publication/70033830>
- Ficetola, G.F. et al. (2015), Replication levels, false presences and the estimation of the presence/absence from eDNA metabarcoding data, *Molecular Ecology Resources* 15: 543–556. DOI: <https://doi.org/10.1111/1755-0998.12338>

EXTERNAL STANDARDISATION STATUS

None intended

LICENCE FOR THE SUBMITTED SPECIFICATION

Creative Commons Attribution 4.0

PUBLICATION DECLARATION

Yes

Procedure or protocol

INPUTS, PREREQUISITES, AND MATERIALS

- A documented pilot or prior run from the same matrix and marker
- Soil extract identifiers and extraction blanks
- At least eight separately indexed PCR replicates in low-detection contexts
- Positive amplification control and no-template controls
- Frozen primer, filtering-pipeline and taxonomic-reference versions
- Per-replicate detections or read counts, never only pooled counts

ORDERED PROCEDURE

1. Freeze the marker, primer, filtering pipeline, taxonomic reference and target taxonomic scope.
2. Estimate per-replicate detection from a pilot or comparable retained run; record the estimate and uncertainty.
3. Use at least eight PCR replicates when detection probability is low or DNA degradation is expected; any lower count requires a power-based justification.
4. Process extraction blanks, no-template controls and a positive control with the batch.
5. Assign each sequence independently per PCR and retain the complete replicate matrix.
6. Where sample size permits, estimate true occurrence, detection and false-positive parameters with the declared occupancy/error model.
7. Call PRESENT only when the preregistered model or replicate rule is satisfied and controls remain acceptable.
8. Route a single low-support positive to INDETERMINATE—REPEAT and keep it separate from presence or absence calls.
9. Call ABSENT only when the model-based detection condition is met with no supported detection.
10. Call INVALID when the preregistered blank or positive-control rule fails.
11. Export the call table, model parameters, replicate matrix, controls, raw-data identifiers and software versions.

PARAMETERS, UNITS, TOLERANCES, AND ALLOWED RANGES

Replication is selected from estimated per-replicate detection and the project's target probability of detecting a truly present taxon. Eight is the default minimum for low-detection or degraded-DNA contexts, following Ficetola et al. (2015). The selected replicate count remains justified by the project-specific detection model. Thresholds, priors, false-positive structure and convergence criteria are frozen before final calls and stored with the result.

EXPECTED OUTPUTS AND RECORDS TO RETAIN

PRESENT, ABSENT, INDETERMINATE—REPEAT or INVALID per taxon-site pair; an unpooled replicate matrix; controls; estimated detection and false-positive parameters with uncertainty; model and software versions; frozen reference database; raw sequence accession or checksum; and a machine-readable audit trail from evidence to categorical call.

CONTROLS, CALIBRATION, AND ACCEPTANCE CRITERIA

Negative extraction and PCR controls must remain below the preregistered contamination threshold; the positive control must fall within its declared acceptance range. Replicate counts below the selected design require written justification before a final call. Model diagnostics and any failed convergence are retained alongside any manual call.

FAILURE MODES AND INVALID RUNS

Failed positive control invalidates the affected assay. Blank-associated target signal invokes the declared contamination branch. Insufficient replication, unversioned reference data, non-convergent model, missing raw files, or irreconcilable replicate labels prevents PRESENT/ABSENT release. Isolated low-support detections remain indeterminate until repeat testing.

SAFETY, ETHICS, AND LEGAL CONSTRAINTS

Environmental samples only; protected-site permissions and biosafety for soil handling remain project responsibilities. Human and clinical DNA are excluded. Sequence access must be screened and governed if incidental human reads could be retained. Contamination controls are both a safety practice and an integrity requirement.

INDEPENDENT REPRODUCIBILITY TEST

An independent analyst receives raw per-replicate assignments, control results, frozen code and database releases, model formula, parameters and decision thresholds. Conformance requires regeneration of the same model diagnostics and categorical calls, including every indeterminate and invalid branch.

Registry review record

Submission received

06 July 2015, 12:00 UTC

Submission received.

Accepted for registration

21 September 2015, 12:00 UTC

Version 1.0 resolves the principal defect in 0.8: replication is linked to detection probability, false negatives and false positives are represented, and categorical calls can be regenerated from retained per-replicate evidence. Registration records a determinate, reproducible project protocol. Broader recommendations about replicate counts remain study-specific.