A validation scale to determine the readiness of environmental DNA assays for routine species monitoring

Bettina Thalinger[1,2]*, Kristy Deiner[3,4], Lynsey R. Harper[5], Helen C. Rees[6], Rosetta C. Blackman[7,8], Daniela Sint[2,9], Michael Traugott[2,9], Caren S. Goldberg[10] & Kat Bruce[11]

[1] Centre for Biodiversity Genomics, University of Guelph, 50 Stone Road E, N1G 2W1, Guelph, Ontario, Canada

[2] Department of Zoology, University of Innsbruck, Technikerstr. 25, 6020, Innsbruck, Austria

[3] Natural History Museum, Cromwell Road, London, SW7 5BD, UK

[4] Institute of Biogeochemistry and Pollutant Dynamics, ETH Zurich, Universitätstr. 16, 8092 Zurich, Switzerland

[5] Illinois Natural History Survey, Prairie Research Institute, University of Illinois at Urbana-Champaign, Champaign, IL, USA

[6] RSK ADAS Ltd, School of Veterinary Medicine and Science, The University of Nottingham, Sutton Bonington Campus, Loughborough, LE12 5RD, UK

[7] Eawag: Swiss Federal Institute of Aquatic Science and Technology, Department of Aquatic Ecology, Überlandstr. 133, 8600 Dübendorf, Switzerland

[8] Department of Evolutionary Biology and Environmental Studies, University of Zurich, Winterthurerstr. 190, CH-8057 Zurich, Switzerland

[9] Sinsoma GmbH, Lannes 6, 6176 Völs, Austria

[10] School of the Environment, Washington State University, Pullman, WA 99164, USA

[11] Nature Metrics Ltd, CABI Site, Bakeham Lane, Egham, TW20 9TY, UK

***Corresponding author:**

Bettina Thalinger, bettina.thalinger@gmail.com

Centre for Biodiversity Genomics, University of Guelph, 50 Stone Road E, N1G 2W1, Guelph, Ontario, Canada

Running headline: eDNA assay validation scale

Abstract

1. Environmental DNA (eDNA) analysis utilises trace DNA released by organisms into their environment for species detection and is revolutionising non-invasive species monitoring. The use of this technology requires rigorous validation - from field sampling to interpretation of PCR-based results - for meaningful application and interpretation. Assays targeting eDNA released by individual species are typically validated with no predefined criteria to answer specific research questions in one ecosystem. Their general applicability, uncertainties and limitations often remain undetermined. The absence of clear guidelines prevents targeted eDNA assays from being incorporated into species monitoring and policy, thus their establishment will be key for the future implementation of eDNA-based surveys.
2. We describe the measures and tests necessary for successful validation of targeted eDNA assays and the associated pitfalls to form the basis of guidelines. A list of 122 variables was compiled and consolidated into a scale to assess the validation status of individual assays. These variables were evaluated for 546 published single-species assays. The resulting dataset was used to provide an overview of current validation practices and test the applicability of the validation scale for future assay rating.
3. The 122 variables representing assay validation status were classified into 14 thematic blocks, such as "*in silico* analysis", and arranged on a 5-level validation scale from "incomplete" to "operational". Additionally, minimum validation criteria were defined for each level. The majority (30%) of investigated assays were classified as Level 1 (incomplete), and 15% did not achieve this first level. These assays were characterised by minimal *in silico* and *in vitro* testing, but their share in annually published eDNA assays has declined since 2014. The total number of reported variables ranged from 20% to 76% and deviated both between and within levels.
4. *Synthesis and applications:* The meta-analysis demonstrates the suitability of the 5-level validation scale for assessing targeted eDNA assays. It is a user-friendly tool to evaluate previously published assays for

future research and routine monitoring, while also enabling appropriate interpretation of results. Finally, it provides guidance on validation and reporting standards for newly developed assays.

Key words: digital PCR, eDNA, endpoint PCR, https://edna-validation.com/, quantitative PCR, species-specific, species detection

62 **1) Introduction**

63

64 Determining the occurrence of species is essential for ecology and requires sensitive and accurate
65 detection methods. Within the last decade, species detection from environmental DNA (eDNA; i.e.
66 detection of trace DNA released by organisms into their environment) has shown great potential for
67 routine species surveys (Rees, Maddison, Middleditch, Patmore, & Gough, 2014; Goldberg et al., 2016;
68 Deiner et al., 2017). The interest in molecular species detection has fuelled the development of over 500
69 assays, reviewed herein, that utilise PCR of DNA or RNA extracted from environmental samples.
70 Generally, "targeted" eDNA assays must be specific to the species of interest and possess high sensitivity
71 to allow detection at low densities, low DNA concentrations, and across spatiotemporal scales (Goldberg
72 et al., 2016; MacDonald & Sarre, 2017).

73 A targeted eDNA assay encompasses the entire workflow used to detect a species' DNA from an
74 environmental sample, inclusive of field sampling to the interpretation of PCR-based results, and not
75 just the primers and probes. Thus, adherence to workflows will determine the success or failure of an
76 eDNA assay because methodological choices influence performance and sensitivity (e.g. Doi, Takahara,
77 et al., 2015). An assay is often validated within a specific system to answer a set question about the
78 target species. Its applications beyond initial development are hampered by the poor understanding of
79 remaining uncertainties, such as the potential for false positives resulting from non-target amplification
80 or contamination, or false negatives resulting from low sensitivity, sample degradation, low DNA yield
81 protocols or inhibition (Goldberg et al., 2016; Lacoursière-Roussel & Deiner, 2019). In a management
82 context, false positives and false negatives may lead to misuse of resources (e.g. funds and personnel)
83 for issues such as rare species protection and invasive species control. Both scenarios foster inaccurate
84 interpretation of results, fuelling arguments against routine use of eDNA detection for species
85 monitoring (Jerde, 2019).

86 Aside from a few well-validated eDNA assays that have been developed and incorporated into
87 routine monitoring, the application of published assays is a minefield for end-users to navigate. We
88 illustrate this through two examples. The assay for great crested newt (*Triturus cristatus*, a legally
89 protected species in the UK [Natural England, 2015]) was one of the first eDNA assays validated in both
90 laboratory and field trials against conventional tools, demonstrating its potential for routine monitoring
91 (Thomsen et al., 2012; Rees, Bishop, et al., 2014). After successful validation, a national eDNA-based
92 citizen science monitoring scheme was tested and showed that large-scale eDNA sampling can enable
93 distribution modelling (Biggs et al., 2015). These initial studies paved the way for eDNA-based
94 *T. cristatus* detection to inform new policies aimed at providing landscape-level species protection (L. R.
95 Harper, Buxton, et al., 2019). Studies have since investigated optimal methods of eDNA capture, relative
96 abundance and detection probability estimation, and the influence of seasonality as well as biotic and
97 abiotic factors on *T. cristatus* eDNA detection and quantification (Buxton, Groombridge, & Griffiths,
98 2017; Buxton, Groombridge, Zakaria, & Griffiths, 2017; Buxton, Groombridge, & Griffiths, 2018a, 2018b).
99 Due to these combined efforts, the assay has undergone exemplary validation and is operational for
100 management.
101 Conversely, no assays have been successfully applied to routine monitoring of invasive American
102 crayfish. For example, there is a lack of consensus on a single assay for the signal crayfish (*Pacifastacus*
103 *leniusculus*). Larson et al. (2017) developed and tested an assay against conventional trapping, but five
104 other assays have also been proposed with differing degrees of validation and divergent *in silico* and *in*
105 *vitro* approaches (Agersnap et al., 2017; Dunn, Priestley, Herraiz, Arnold, & Savolainen, 2017; K. J.
106 Harper, Anucha, Turnbull, Bean, & Leaver, 2018; Mauvisseau et al., 2018; Robinson, Uren Webster,
107 Cable, James, & Consuegra, 2018). The *P. leniusculus* assays were developed using a variety of strategies
108 for eDNA sampling, capture, extraction, qPCR of different genetic markers, and applied across
109 genetically diverse populations within the species range. Due to this substantial methodological

110 variability, direct comparisons between results obtained from these assays are impossible. Therefore,
111 the *P. leniusculus* assays represent a minefield for end-users, despite the need for accurate and sensitive
112 tools to enable actionable species management of this invasive species in Europe.

113 These case studies exemplify how consensus and dissent in assay validation can influence the
114 implementation of eDNA analysis for species monitoring. Developing guidelines to determine the
115 suitability of eDNA assays for end-users will therefore ensure that ecological insights or management
116 decisions are based on robust molecular analyses with quantifiable uncertainties and clear inference
117 limits (Goldberg et al., 2016; MacDonald & Sarre, 2017; Nicholson et al., 2020). Here, we describe the
118 general validation process for targeted PCR-based methods and examine the extent of assay validation
119 and reporting in the eDNA literature. We present an eDNA assay validation scale, which establishes
120 criteria to enable the classification of assays based on their accuracy and sensitivity for single-species
121 detection. To demonstrate the utility of the scale, we performed a meta-analysis of targeted eDNA
122 assays published in 327 papers as of 11th April 2019 (546 assays). By placing an eDNA assay on the
123 validation scale, end-users can determine the recommended scenarios for application and improve
124 assay performance with further validation.

125

126

127 **2) Criteria and principles of validation**

128

129 *General requirements for an eDNA laboratory*

130

131 All laboratory activities are subject to error. In order to have confidence in results, quality standards and
132 good practices are required in diagnostic laboratory environments (e.g. World Health Organization,
133 2011; Halling, Schrijver, & Persons, 2012). Although few eDNA-processing laboratories will employ

134 'ancient DNA' practices (e.g. hazmat suit and positive air pressure with HEPA filtered inflow), all
135 laboratories conducting eDNA analysis should utilise a unidirectional workflow where pre-PCR steps are
136 performed in separate laboratories dedicated to low DNA quality and quantity (Goldberg et al., 2016).
137 Completely standardised laboratory environments are rare and the use of proficiency tests (as
138 conducted by UK laboratories participating in *T. cristatus* monitoring) can help end-users understand the
139 quality of results obtained among different laboratories. Even results obtained from an extensively
140 validated assay can be questionable when they are not produced within a suitable laboratory
141 environment (Goldberg et al., 2016).
142
143 *Reporting standards for* in silico*,* in vitro *and* in situ *validation of assays*
144
145 Targeted eDNA assay validation is a multi-step process. It can be divided into *in silico* validation (i.e.
146 computer-based tests for primer specificity), *in vitro* validation (i.e. laboratory tests with reference
147 tissue samples) and *in situ* validation (i.e. field tests with eDNA samples) (see Goldberg et al., 2016;
148 MacDonald & Sarre, 2017). Understanding the utility of an assay requires both knowledge of the context
149 in which it has been designed, and a broader understanding of how it was developed. Here, we give a
150 brief overview of what key steps comprise assay validation, with a focus on troubleshooting steps that
151 may be necessary when applying previously published assays to new locations or with modified
152 laboratory practices (Table 1).
153 The first step is *in silico* assay validation, the goal of which is to determine assay specificity based
154 on known sequence diversity. Sequence diversity has three categories, including sequences from: (i)
155 closely related and co-occurring species, (ii) closely related but geographically distinct species, and (iii)
156 distantly related but co-occurring species (i.e. sequences that could co-amplify and produce false
157 positive results for a target species). By checking primer specificity from available sequences, the

158 geographic area of applicability for an assay can be maximised through identifying and removing
159 potential issues of co-amplification. Typically public or custom databases are used for performing *in*
160 *silico* amplification (e.g. ecoPCR [Boyer et al., 2016]; primerBlast [Ye et al., 2012]; PrimerTREE [Cannon et
161 al., 2016]; PrimerMiner [Elbrecht & Leese, 2017]). While reference sequence libraries are often far from
162 complete and many of the factors influencing successful PCR amplification cannot be simulated, *in silico*
163 testing provides a first impression of primer performance and should be conducted.

164 The essential components of *in vitro* assay validation are optimisation, specificity, and sensitivity.
165 Tests with varying PCR chemistry, reaction volume, primer/probe concentration, cycling conditions and
166 technical replication will ensure optimal, standardised and error-free target DNA amplification (Bustin et
167 al., 2009; Wilcox, Carim, McKelvey, Young, & Schwartz, 2015; Goldberg et al., 2016). The assay must
168 then be tested against closely related and co-occurring non-target taxa to ensure specificity (Goldberg et
169 al., 2016). Ideally, tissue-derived DNA samples from multiple individuals spanning a defined geographic
170 area are tested to ensure the assay is robust to genetic variants of target and non-target species.
171 Amplicons should be Sanger sequenced to confirm species identity (Goldberg et al., 2016), but short
172 fragments (<100 bp) have limited sequence length available for species determination (Meusnier et al.,
173 2008). Next, the Limit of Detection (LOD) must be determined to assess assay sensitivity, and the Limit
174 of Quantification (LOQ) determined, if the measurement of eDNA quantity is desired. Generally, these
175 values are obtained using a dilution series of quantified DNA amplicons or synthesized gene fragments
176 (e.g. gBlock, IDT) based on public or *de novo* reference sequences. The LOD and LOQ have various
177 definitions in the eDNA literature, but were recently standardised by Klymus et al. (2019), where LOD is
178 the lowest standard concentration at which 95% of technical replicates amplify and LOQ is the lowest
179 standard concentration for which the coefficient of variation (CV) value is <35%. Unfortunately, the
180 existence of past definitions requires the final LOD and LOQ to be reported as well as the definition
181 used. We note that these metrics apply directly to the assay as developed and assume no interference

182 during PCR from the rest of the species' genome (i.e. if a gBlock is used), other genetic material, or
183 inhibitory compounds.

184 Finally, the assay must be validated *in situ* by surveying sites with and without the target species
185 (Goldberg et al., 2016). It must be tested against conventional tools for presence/absence detection and
186 tests for estimation of relative abundance/biomass are advisable. Assays are deemed successful if eDNA
187 and conventional detections concur at occupied sites and no eDNA detections are observed at
188 definitively unoccupied sites. Sanger sequencing of eDNA amplicons can provide additional evidence but
189 cannot distinguish sample contamination from true detections (Goldberg et al., 2016). Besides screening
190 for the target species, negative eDNA samples (or all eDNA samples if quantification is necessary) should
191 be tested for inhibition. This requires an Internal Positive Control (IPC) assay for synthetic DNA (e.g.
192 ThermoFisher) or an assay for non-target species using exogenous or endogenous DNA (Goldberg et al.,
193 2016; Doi et al., 2017; Furlan & Gleeson, 2017).

194 Advanced *in situ* validation may investigate the influence of biotic (e.g. abundance, biomass, life
195 stages, microbial activity) and abiotic (e.g. temperature, pH, ultraviolet light, salinity) factors influencing
196 eDNA origin, state, fate, and transport (Barnes & Turner, 2016; Lacoursière-Roussel & Deiner, 2019).
197 Assays that account for spatial (e.g. shoreline versus offshore) and temporal (e.g. summer versus winter)
198 variation in eDNA distribution and abundance due to the ecology of a species can be implemented with
199 greater confidence (de Souza, Godwin, Renshaw, & Larson, 2016; Lawson Handley et al., 2019).
200 Occupancy modelling using eDNA data is desirable as it accounts for detection probability while
201 estimating site occupancy, even if all field samples from a site return negative. Hierarchical models that
202 incorporate eDNA occupancy and detection probabilities at site, sample, and technical replicate levels
203 are most accurate and can be implemented in software such as R (e.g. package "eDNAoccupancy"
204 [Dorazio & Erickson, 2018]) or PRESENCE (MacKenzie et al., 2002). However, model assumptions
205 regarding false positives should be carefully considered.

206

207

208 **3) Types and trade-offs of targeted eDNA detection methods**

209

210 Amid the processing chain (i.e. sampling to data analysis) for a targeted eDNA assay, PCR warrants extra
211 consideration as the technological spectrum and potential for variation is enormous. Previous
212 publications have typically defined an assay as the primers (and probe) required for DNA amplification
213 and associated visualisation (e.g. agarose gel electrophoresis, qPCR software). Hence, differences
214 between the multiple detection instruments and chemistry used in combination with species-specific
215 primer (and probe) sets are at the core of targeted eDNA assays. Table 2 provides an overview of
216 amplification types and their associated trade-offs.

217 Many assays - especially those published in earlier years - use endpoint PCR. However, most assays
218 to date employ real-time quantitative (q)PCR allowing for greater sensitivity and quantitative data. More
219 recent publications have used digital (d)PCR for absolute quantification. Alternatively, LAMP and CRISPR
220 were shown to be suitable for eDNA applications, decreasing the requirements of in-field testing
221 equipment (M. R. Williams et al., 2017; M. A. Williams et al., 2019). A few publications use alternative
222 methods such as PCR combined with restriction fragment length polymorphism (RFLP). All amplification
223 types (Table 2) enable distribution and occupancy modelling, provided enough biological and technical
224 replication is employed (Hunter et al., 2015; Goldberg, Strickler, & Fremier, 2018; Wilcox et al., 2018).
225 However, endpoint PCR in combination with agarose gels is the only type where a detection limit cannot
226 be set objectively (low sensitivity) and which does not provide estimates of DNA copy number (Doi,
227 Uchii, et al., 2015; Yamanaka & Minamoto, 2016; Hunter et al., 2017; Thalinger, Wolf, Traugott, &
228 Wanzenböck, 2019). Depending on the management context and study size, the optimal detection
229 instrument can vary, albeit technological advances will slowly shift the focus away from endpoint PCR.

230

231

232 **4) Evaluating the current status of assay validation**

233

234 To understand current validation practices for targeted eDNA assays, we generated an extensive list of
235 variables deemed important for assay validation by 35 experts in the field of targeted eDNA detection
236 whom convened at a DNAqua-Net EU COST Action (Leese et al., 2016) workshop held on 26-27 March
237 2018 at the University of Innsbruck, Austria. The variables in this list consist of 18 categorical variables
238 (e.g., species identity, target gene, sample type) and 104 binomial variables directly associated with the
239 eDNA processing chain, from primer design to interpretation of field study results (SI1).

240 A comprehensive literature database for targeted eDNA assays was built in three steps. First, we
241 included all papers listed on the 'eDNA assays' web page (https://labs.wsu.edu/edna/edna-assays/) as of
242 10 April 2019. Second, we conducted a Web of Science literature search on 11 April 2019, including the
243 search terms "environmental DNA" and "eDNA" but excluding terms associated with microbial
244 organisms and metabarcoding (SI2A). Third, the resulting 660 Web of Science entries were manually
245 checked for suitability (i.e. macrobial target organism, targeted eDNA detection intended) leading to a
246 combined database of 327 papers. For each of the assays contained in these papers, the 122 variables
247 were recorded in a checklist by one of the authors. Before data entry, all authors validated the same
248 four papers (Deiner & Altermatt, 2014; Rees, Bishop, et al., 2014; Thalinger et al., 2016; L. R. Harper,
249 Griffiths, et al., 2019) to ensure recorder standardisation. Validation efforts were classified as 1 for
250 "tests done, or parameter reported", 0 for "variable not reported, or no testing done", and NA in cases
251 where the respective variable did not match the assay type (e.g. filter type when samples were
252 precipitated). When an assay was used in multiple papers, all validation efforts were summarised in one
253 entry and the literature database was extended with the papers reporting primer sequences or other

254 methodological aspects. As the type of amplification is important for assay validation, primer pairs used
255 on multiple detection platforms were given separate database entries per amplification type. However,
256 because most assays developed were presented in one publication, we did not account for slight
257 variations in other aspects of the workflow (e.g. different extraction method, different filter type). After
258 recording the values for each eDNA assay using the validation checklist (SI1), each author scored the
259 assay intuitively based on a preliminary version of the validation scale (see section 5). The resulting
260 database of 122 variables for each assay was the basis for all further analyses using R (R Core Team,
261 2019) and associated packages (SI2B).

262 Altogether, 546 assays from 327 papers were assessed. Of these assays, 227 were designed to
263 detect fish species and 74 were designed to detect amphibian species; hence, it is unsurprising that
264 ~80% of assays utilised water sampling. Fourteen percent of the assays were tested on tissue only and
265 few assays were optimised for other sample types such as aerosol, sediment, snow or soil. More than
266 80% of assays were reported in only one paper, and most were designed for qPCR (~60%) or endpoint
267 PCR (~35%) platforms. The cytochrome c oxidase subunit I (*COI*) gene was the most popular (>40%)
268 genetic marker, followed by the cytochrome b (*cytb*) gene (~23%) (SI3).

269

270

271 **5) The 5-level assay validation scheme**

272

273 To enable standardised assay validation and reporting in the future, we assigned the assay-specific
274 variables to 14 thematic blocks such as "*in silico* analysis", "PCR", or "extensive field testing of
275 environmental samples" (Table 3, Fig. 1). These blocks were placed on a five-level scale enabling the
276 categorisation of assays from Level 1 ("incomplete") to Level 5 ("operational"), and the interpretation of
277 associated field study results (Fig. 1). Each of these blocks contained variables associated with either

278 rudimentary or substantial validation and reporting. For instance, the thematic block "concentration of
279 eDNA from environmental sample" contains the variable "volume/weight of environmental sample",
280 which was reported for almost all assays, but also contains "pressure used for filtration", which was
281 rarely reported and/or measured. Therefore, a minimum criterion was introduced for each variable
282 block functioning as proof of validation. For example, "detection from an environmental sample" was
283 used as evidence that some validation had been undertaken in the block "detection obtained from
284 environmental samples" (Table 3). To reach a level on the validation scale, an assay must fulfil all
285 minimum reporting criteria for that level. For each assay, we calculated a total scoring percentage and a
286 block scoring percentage for the 14 blocks by dividing the number of variables tested or reported by the
287 complete set of variables possible. The total possible for each percentage included only variables
288 relevant to the applied methods (e.g., for assays using filtration, precipitation variables were omitted;
289 see Box 1 for example assays).

290 Of the 546 assays analysed, the majority (30%) were classed as Level 1. Of the remainder, 15%
291 (N = 83) did not fulfil the minimum criteria necessary to reach Level 1, and no assay reached Level 5
292 (Fig. 1 and Fig. 2A). Newer assays published after 2016 were more likely to reach Level 4, and the
293 percentage of assays failing to reach Level 1 gradually declined since 2014 (Fig. 2B). Generally, the total
294 scoring percentage for all variables in a level increased from Level 0 to Level 4, but its variation was non-
295 uniform with most outliers observed at Level 2 (Fig. 2C). Assays not reaching Level 1 exhibited scoring
296 percentages between ~20% and ~55%, clearly showing the difference between incomplete and partially
297 validated assays, which did not achieve higher levels due to one or several missing validation step(s).
298 Generally, variables associated with lower levels on the validation scale were more likely to be reported
299 or tested. Nevertheless, some variables (e.g. "haplotypes of target tissue" or "pressure used for
300 filtration"), were addressed by fewer than 10% of assays (Fig. 2D).

301 The rigour of the minimum criteria was evaluated by tallying the not achieved or reported cases.
302 Specifically, 62 assays did not reach Level 1 because the targeted species sequence(s) used in primer
303 design were not reported, which is a prerequisite for this level. Eight assays did not report the primer
304 sequence. The lack of target detection from an environmental sample and omission of filter type or
305 precipitation chemicals were the most restrictive criteria and were not fulfilled by 80 and 43 assays
306 respectively, most of which were exclusively used for tissue tests (Fig. 2E). For assays ranked at Levels 2
307 and 3, there was agreement between the intuitive assay rating provided by the recorder and that
308 assigned by the objective criteria. For assays placed at Level 1 following objective criteria, authors
309 tended to be more liberal and rated them one or two levels higher (SI4). Finally, a classification tree
310 analysis (De'Ath & Fabricius, 2000) was carried out to identify common characteristics of assays placed
311 at each level of the validation scale (SI5). Most assays failing to reach Level 1 showed distinctly low levels
312 of *in silico* validation. This was also true for most Level 1 assays (*N* = 103), albeit these showed higher
313 levels of target tissue validation (Fig. 3). On the other end of the spectrum, Level 4 assays were
314 characterised by substantial testing or reporting for *in vitro* testing, field sample processing, LOD
315 determination, PCR, and field testing.
316
317
318 **6) Conclusions and outlook**
319
320 The validation scale and reporting standards developed here are an inclusive set of guidelines for
321 targeted eDNA assays. They are directed at both the scientific community and end-users who want to
322 apply a previously published assay or develop and publish a new assay. One needs to acknowledge that
323 a strict standardisation of eDNA assays will not be possible due to applications for manifold taxa in
324 diverse ecosystems combined with technological advances. By checking which of the 122 validation

325 variables have been addressed, it is possible to identify both available and missing information needed
326 to successfully develop or reuse an assay. As a general recommendation, authors should report as much
327 information as possible on the conducted validation steps, either in the main text or in the
328 supplementary material. Additionally, developers should consult existing guidelines for best practices
329 along the validation workflow prior to assay design and fieldwork (Bustin et al., 2009; Goldberg et al.,
330 2016; MacDonald & Sarre, 2017; Klymus et al., 2019). To make the validation process accessible, we
331 provide a checklist in SI1. Furthermore, a website (https://edna-validation.com/) was created to
332 summarise the cornerstones of the validation process and the validation scale. It will be possible to
333 enter all 122 variables and the minimum criteria to rank assays and calculate their scoring percentages
334 on the scale. The website will also serve as a living document when improvements in technology and/or
335 our understanding of eDNA in the environment advance.

336 The 5-level validation scale designed here provides an overview of the capabilities and uncertainties
337 of targeted eDNA assays. However, the binary data entry system cannot replace a close check of
338 previous publications as it does not always allow a qualitative assessment. Details for validation
339 variables are often spread across different sections in a publication or ambiguously displayed. Thus, the
340 checklist can be used as standard reporting guidelines for targeted eDNA assays. It should be
341 emphasised that for specific research questions and associated publications, minimal validation efforts
342 may be sufficient. Nevertheless, thorough validation is needed to reduce uncertainties and overcome
343 the limitations associated with eDNA-based species monitoring. Furthermore, it is important that
344 practitioners consider how an assay can be modified (e.g. using different PCR reagents) and whether this
345 changes its validation level.

346 The successful application of targeted eDNA assays for routine species detection and monitoring
347 largely depends on the scientific community and the industry providing eDNA services. Laboratories
348 participating in ring tests such as that proposed for metabarcoding (Blackman et al., 2019) can facilitate

consensus on analysis standards. For now, assay developers must respond to queries and help troubleshoot reproducibility issues. Such engagement will facilitate the application of targeted eDNA assays by other users and outside their original geographic scope or academic context. Finally, it is necessary for both the scientific community and the commercial laboratories to communicate realistic applications and limitations to end-users, as often an assay is not bad *per se*, but simply unsuitable.

Acknowledgements

This article is based upon work from COST Action DNAqua-Net (CA15219), supported by the COST (European Cooperation in Science and Technology) programme. We thank the workshop participants who convened on 26-27 March 2018 at the University of Innsbruck, Austria, where initial brainstorming for this manuscript occurred. We acknowledge additional guidance and input from Micaela Hellström, and are grateful to Cathryn Abbott for the improved design of Fig.1. We thank Nature Metrics Ltd. for support in hosting the independent website formed to operationalise the validation scale for future use.

Author's contributions

KB conceived the original idea which was developed by all co-authors. BT compiled the literature database and BT, KD, LRH, HCR, RCB, DS, MT, and KB collected the data. BT analysed the data and led the writing of the manuscript for which first drafts of sections were provided by LH, RCB and KD. All authors contributed critically to the drafts and gave final approval for publication.

373

374 **Data availability**

375

376 The literature database, the validation checklist compiled from it, the R code, and seven datasets
377 derived from the checklist and used for analyses are available on Figshare at
378 https://doi.org/10.6084/m9.figshare.12184860.v1 The literature database, the validation scale, and the
379 full checklist are also available at https://edna-validation.com/.

380

381

382 **Conflict of interest**

383

384 KB is the co-founder and CEO of Nature Metrics Ltd. a for profit company dedicated to the analysis of
385 environmental DNA. HCR manages environmental DNA services for RSK ADAS Ltd. a for profit
386 environmental consultancy. DS and MT are co-founders of Sinsoma GmbH. a for profit company
387 dedicated to the analysis of environmental DNA.

388

389 **References**

390

391 Agersnap, S., Larsen, W. B., Knudsen, S. W., Strand, D., Thomsen, P. F., Hesselsøe, M., … Møller, P. R.
392 (2017). Monitoring of noble, signal and narrow-clawed crayfish using environmental DNA from
393 freshwater samples. *PLOS ONE*, *12*(6), e0179261. doi:10.1371/journal.pone.0179261

394 Barnes, M. A., & Turner, C. R. (2016). The ecology of environmental DNA and implications for
395 conservation genetics. *Conservation Genetics*, *17*(1), 1–17. doi:10.1007/s10592-015-0775-4

396 Biggs, J., Ewald, N., Valentini, A., Gaboriaud, C., Dejean, T., Griffiths, R. A., … Dunn, F. (2015). Using eDNA
397 to develop a national citizen science-based monitoring programme for the great crested newt
398 (*Triturus cristatus*). *Biological Conservation*, *183*, 19–28. doi:10.1016/j.biocon.2014.11.029

399 Blackman, R., Mächler, E., Altermatt, F., Arnold, A., Beja, P., Boets, P., … Deiner, K. (2019). Advancing the
400 use of molecular methods for routine freshwater macroinvertebrate biomonitoring – the need for
401 calibration experiments. *Metabarcoding and Metagenomics*, *3*, e34735.
402 doi:10.3897/mbmg.3.34735

403 Boyer, F., Mercier, C., Bonin, A., Le Bras, Y., Taberlet, P., & Coissac, E. (2016). obitools: A unix-inspired
404 software package for DNA metabarcoding. *Molecular Ecology Resources*, *16*(1), 176–182.
405 doi:10.1111/1755-0998.12428

406 Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., … Wittwer, C. T. (2009). The
407 MIQE guidelines: Minimum information for publication of quantitative real-time PCR experiments.
408 *Clinical Chemistry*, *55*(4), 611–622. doi:10.1373/clinchem.2008.112797

409 Buxton, A. S., Groombridge, J., & Griffiths, R. (2018a). Comparison of Two Citizen Scientist Methods for
410 Collecting Pond Water Samples for Environmental DNA Studies. *Citizen Science: Theory and*
411 *Practice*, *3*(2). doi:10.5334/cstp.151

412 Buxton, A. S., Groombridge, J. J., & Griffiths, R. A. (2017). Is the detection of aquatic environmental DNA

influenced by substrate type? *PLOS ONE*, *12*(8), e0183371. doi:10.1371/journal.pone.0183371

Buxton, A. S., Groombridge, J. J., & Griffiths, R. A. (2018b). Seasonal variation in environmental DNA detection in sediment and water samples. *PLOS ONE*, *13*(1). doi:10.1371/journal.pone.0191737

Buxton, A. S., Groombridge, J. J., Zakaria, N. B., & Griffiths, R. A. (2017, April 10). Seasonal variation in environmental DNA in relation to population size and environmental factors. *Scientific Reports*. Nature Publishing Group. doi:10.1038/srep46294

Cannon, M. V., Hester, J., Shalkhauser, A., Chan, E. R., Logue, K., Small, S. T., & Serre, D. (2016). In silico assessment of primers for eDNA studies using PrimerTree and application to characterize the biodiversity surrounding the Cuyahoga River. *Scientific Reports*, *6*. doi:10.1038/srep22908

De'Ath, G., & Fabricius, K. E. (2000). Classification and regression trees: A powerful yet simple technique for ecological data analysis. *Ecology*, *81*(11), 3178–3192. doi:10.1890/0012-9658(2000)081[3178:CARTAP]2.0.CO;2

de Souza, L. S., Godwin, J. C., Renshaw, M. A., & Larson, E. (2016). Environmental DNA (eDNA) detection probability is influenced by seasonal activity of organisms. *PLOS ONE*, *11*(10), e0165273. doi:10.1371/journal.pone.0165273

Deiner, K., & Altermatt, F. (2014). Transport distance of invertebrate environmental DNA in a natural river. *PLOS ONE*, *9*(2). doi:10.1371/journal.pone.0088786

Deiner, K., Bik, H. M., Mächler, E., Seymour, M., Lacoursière-Roussel, A., Altermatt, F., … Bernatchez, L. (2017). Environmental DNA metabarcoding: Transforming how we survey animal and plant communities. *Molecular Ecology*, *26*(21), 5872–5895. doi:10.1111/mec.14350

Doi, H., Katano, I., Sakata, Y., Souma, R., Kosuge, T., Nagano, M., … Tojo, K. (2017). Detection of an endangered aquatic heteropteran using environmental DNA in a wetland ecosystem. *Royal Society Open Science*, *4*(7). doi:10.1098/rsos.170568

Doi, H., Takahara, T., Minamoto, T., Matsuhashi, S., Uchii, K., & Yamanaka, H. (2015). Droplet digital

polymerase chain reaction (PCR) outperforms real-time PCR in the detection of environmental DNA from an invasive fish species. *Environmental Science and Technology*, *49*(9), 5601–5608. doi:10.1021/acs.est.5b00253

Doi, H., Uchii, K., Takahara, T., Matsuhashi, S., Yamanaka, H., & Minamoto, T. (2015). Use of droplet digital PCR for estimation of fish abundance and biomass in environmental DNA surveys. *PLOS ONE*, *10*(3), 1–11. doi:10.1371/journal.pone.0122763

Dorazio, R. M., & Erickson, R. A. (2018). ednaoccupancy: An r package for multiscale occupancy modelling of environmental DNA data. *Molecular Ecology Resources*, *18*(2), 368–380. doi:10.1111/1755-0998.12735

Dunn, N., Priestley, V., Herraiz, A., Arnold, R., & Savolainen, V. (2017). Behavior and season affect crayfish detection and density inference using environmental DNA. *Ecology and Evolution*, *7*(19), 7777–7785. doi:10.1002/ece3.3316

Elbrecht, V., & Leese, F. (2017). PrimerMiner: an R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, *8*(5), 622–626. doi:10.1111/2041-210X.12687

Furlan, E. M., & Gleeson, D. (2017). Improving reliability in environmental DNA detection surveys through enhanced quality control. *Marine and Freshwater Research*, *68*(2), 388. doi:10.1071/MF15349

Goldberg, C. S., Strickler, K. M., & Fremier, A. K. (2018). Degradation and dispersion limit environmental DNA detection of rare amphibians in wetlands: Increasing efficacy of sampling designs. *Science of the Total Environment*, *633*, 695–703. doi:10.1016/j.scitotenv.2018.02.295

Goldberg, C. S., Turner, C. R., Deiner, K., Klymus, K. E., Thomsen, P. F., Murphy, M. A., … Taberlet, P. (2016). Critical considerations for the application of environmental DNA methods to detect aquatic species. *Methods in Ecology and Evolution*, *7*(11), 1299–1307. doi:10.1111/2041-210X.12595

461 Halling, K. C., Schrijver, I., & Persons, D. L. (2012). Test verification and validation for molecular
462 diagnostic assays. *Archives of Pathology and Laboratory Medicine*, *136*(1), 11–13.
463 doi:10.5858/arpa.2011-0212-ED
464 Harper, K. J., Anucha, N. P., Turnbull, J. F., Bean, C. W., & Leaver, M. J. (2018). Searching for a signal:
465 Environmental DNA (eDNA) for the detection of invasive signal crayfish, *Pacifastacus leniusculus*
466 (Dana, 1852). *Management of Biological Invasions*, *9*(2), 137–148. doi:10.3391/mbi.2018.9.2.07
467 Harper, L. R., Buxton, A. S., Rees, H. C., Bruce, K., Brys, R., Halfmaerten, D., … Hänfling, B. (2019).
468 Prospects and challenges of environmental DNA (eDNA) monitoring in freshwater ponds.
469 *Hydrobiologia*, *826*(1), 25–41. doi:10.1007/s10750-018-3750-5
470 Harper, L. R., Griffiths, N. P., Lawson Handley, L., Sayer, C. D., Read, D. S., Harper, K. J., … Hänfling, B.
471 (2019). Development and application of environmental DNA surveillance for the threatened
472 crucian carp (*Carassius carassius*). *Freshwater Biology*, *64*(1), 93–107. doi:10.1111/fwb.13197
473 Hunter, M. E., Dorazio, R. M., Butterfield, J. S. S., Meigs-Friend, G., Nico, L. G., & Ferrante, J. A. (2017).
474 Detection limits of quantitative and digital PCR assays and their influence in presence–absence
475 surveys of environmental DNA. *Molecular Ecology Resources*, *17*(2), 221–229. doi:10.1111/1755-
476 0998.12619
477 Hunter, M. E., Oyler-McCance, S. J., Dorazio, R. M., Fike, J. A., Smith, B. J., Hunter, C. T., … Hart, K. M.
478 (2015). Environmental DNA (eDNA) sampling improves occurrence and detection estimates of
479 invasive Burmese pythons. *PLOS ONE*, *10*(4), e0121655. doi:10.1371/journal.pone.0121655
480 Jerde, C. L. (2019). Can we manage fisheries with the inherent uncertainty from eDNA? *Journal of Fish*
481 *Biology*, jfb.14218. doi:10.1111/jfb.14218
482 Klymus, K. E., Merkes, C. M., Allison, M. J., Goldberg, C. S., Helbing, C. C., Hunter, M. E., … Richter, C. A.
483 (2019). Reporting the limits of detection and quantification for environmental DNA assays.
484 *Environmental DNA*, edn3.29. doi:10.1002/edn3.29

Lacoursière-Roussel, A., & Deiner, K. (2019). Environmental DNA is not the tool by itself. *Journal of Fish Biology*, jfb.14177. doi:10.1111/jfb.14177

Larson, E. R., Renshaw, M. A., Gantz, C. A., Umek, J., Chandra, S., Lodge, D. M., & Egan, S. P. (2017). Environmental DNA (eDNA) detects the invasive crayfishes *Orconectes rusticus* and *Pacifastacus leniusculus* in large lakes of North America. *Hydrobiologia*, *800*(1), 173–185. doi:10.1007/s10750-017-3210-7

Lawson Handley, L., Read, D. S., Winfield, I. J., Kimbell, H., Johnson, H., Li, J., … Hänfling, B. (2019). Temporal and spatial variation in distribution of fish environmental DNA in England's largest lake. *Environmental DNA*, *1*(1), 26–39. doi:10.1002/edn3.5

Leese, F., Altermatt, F., Bouchez, A., Ekrem, T., Hering, D., Meissner, K., … Zimmermann, J. (2016). DNAqua-Net: Developing new genetic tools for bioassessment and monitoring of aquatic ecosystems in Europe. *Research Ideas and Outcomes*, *2*, e11321. doi:10.3897/rio.2.e11321

MacDonald, A. J., & Sarre, S. D. (2017). A framework for developing and validating taxon-specific primers for specimen identification from environmental DNA. *Molecular Ecology Resources*, *17*(4), 708–720. doi:10.1111/1755-0998.12618

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, A., & Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology*, *83*(8), 2248–2255.

Mauvisseau, Q., Coignet, A., Delaunay, C., Pinet, F., Bouchon, D., & Souty-Grosset, C. (2018). Environmental DNA as an efficient tool for detecting invasive crayfishes in freshwater ponds. *Hydrobiologia*, *805*(1), 163–175. doi:10.1007/s10750-017-3288-y

Meusnier, I., Singer, G. A. C., Landry, J. F., Hickey, D. A., Hebert, P. D. N., & Hajibabaei, M. (2008). A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics*, *9*. doi:10.1186/1471-2164-9-214

Natural England. (2015). Great crested newts: protection and licences. Retrieved 7 April 2020, from https://www.gov.uk/guidance/great-crested-newts-protection-surveys-and-licences

Nicholson, A., McIsaac, D., MacDonald, C., Gec, P., Mason, B. E., Rein, W., … Hanner, R. H. (2020). An analysis of metadata reporting in freshwater environmental DNA research calls for the development of best practice guidelines. *Environmental DNA*. doi:10.1002/edn3.81

R Core Team. (2019). R: A Language and Environment for Statistical Computing. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.r-project.org/

Rees, H. C., Bishop, K., Middleditch, D. J., Patmore, J. R. M., Maddison, B. C., & Gough, K. C. (2014). The application of eDNA for monitoring of the Great Crested Newt in the UK. *Ecology and Evolution*, *4*(21), 4023–4032. doi:10.1002/ece3.1272

Rees, H. C., Maddison, B. C., Middleditch, D. J., Patmore, J. R. M., & Gough, K. C. (2014). The detection of aquatic animal species using environmental DNA - a review of eDNA as a survey tool in ecology. *Journal of Applied Ecology*, *51*(5), 1450–1459. doi:10.1111/1365-2664.12306

Robinson, C. V., Uren Webster, T. M., Cable, J., James, J., & Consuegra, S. (2018). Simultaneous detection of invasive signal crayfish, endangered white-clawed crayfish and the crayfish plague pathogen using environmental DNA. *Biological Conservation*, *222*, 241–252. doi:10.1016/j.biocon.2018.04.009

Thalinger, B., Oehm, J., Mayr, H., Obwexer, A., Zeisler, C., & Traugott, M. (2016). Molecular prey identification in Central European piscivores. *Molecular Ecology Resources*, *16*(1), 123–137. doi:10.1111/1755-0998.12436

Thalinger, B., Wolf, E., Traugott, M., & Wanzenböck, J. (2019). Monitoring spawning migrations of potamodromous fish species via eDNA. *Scientific Reports*, *9*(1). doi:10.1038/s41598-019-51398-0

Thomsen, P. F., Kielgast, J., Iversen, L. L., Wiuf, C., Rasmussen, M., Gilbert, M. T. P., … Willerslev, E. (2012). Monitoring endangered freshwater biodiversity using environmental DNA. *Molecular*

533 *Ecology*, *21*(11), 2565–2573. doi:10.1111/j.1365-294X.2011.05418.x

534 Wilcox, T. M., Carim, K. J., McKelvey, K. S., Young, M. K., & Schwartz, M. K. (2015). The dual challenges of
535 generality and specificity when developing environmental DNA markers for species and subspecies
536 of *Oncorhynchus*. *PLOS ONE*, *10*(11), e0142008. doi:10.1371/journal.pone.0142008

537 Wilcox, T. M., Young, M. K., McKelvey, K. S., Isaak, D. J., Horan, D. L., & Schwartz, M. K. (2018). Fine-scale
538 environmental DNA sampling reveals climate-mediated interactions between native and invasive
539 trout species. *Ecosphere*, *9*(11), e02500. doi:10.1002/ecs2.2500

540 Williams, M. A., O'Grady, J., Ball, B., Carlsson, J., de Eyto, E., McGinnity, P., … Parle-McDermott, A.
541 (2019). The application of CRISPR-Cas for single species identification from environmental DNA.
542 *Molecular Ecology Resources*, *19*(5), 1106–1114. doi:10.1111/1755-0998.13045

543 Williams, M. R., Stedtfeld, R. D., Engle, C., Salach, P., Fakher, U., Stedtfeld, T., … Hashsham, S. A. (2017).
544 Isothermal amplification of environmental DNA (eDNA) for direct field-based monitoring and
545 laboratory confirmation of *Dreissena* sp. *PLOS ONE*, *12*(10). doi:10.1371/journal.pone.0186462

546 World Health Organization. (2011). *Laboratory quality standards and their implementation.* New Delhi:
547 WHO Regional Office for the Western Pacific.

548 Yamanaka, H., & Minamoto, T. (2016). The use of environmental DNA of fishes as an efficient method of
549 determining habitat connectivity. *Ecological Indicators*, *62*, 147–153.
550 doi:10.1016/j.ecolind.2015.11.022

551 Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., & Madden, T. L. (2012). Primer-BLAST: a tool
552 to design target-specific primers for polymerase chain reaction. *BMC Bioinformatics*, *13*, 134.
553 doi:10.1186/1471-2105-13-134

554

555 **Tables and Figures**

556

557 **Table 1:** A guideline for troubleshooting at different stages of the validation process.

558

559 **Table 2:** Detection methods and their trade-offs used for targeted eDNA assays. Abbreviations are as
560 follows: polymerase chain reaction (PCR), capillary electrophoresis (CE), quantitative (q), intercalating
561 dye (ID), loop-mediated isothermal amplification (LAMP).

562

563 **Table 3:** The thematic variable blocks of the 5-level validation scale and their respective minimum
564 criteria.

565

566 **Figure 1:** An overview of the 5-level validation scale. For each of the levels (incomplete to operational),
567 the main accomplishments in the validation process and an appropriate interpretation of results are
568 provided.

569

570 **Figure 2:** The main outcomes of the meta-analysis based on 546 assays from 327 publications are
571 presented in panels A to E. Assay classification is based on the minimum criteria presented in Figure 1.
572 Level 0 codes for assays that did not reach Level 1 on the validation scale. The colour coding is consistent
573 for all panels: Level 0 (grey), Level 1 (dark purple), Level 2 (blue), Level 3 (turquoise), Level 4 (green), and
574 Level 5 (yellow). Panel A shows the distribution of assays across levels of the validation scale. Panel B
575 displays the percentage of assays (N = 546) rated Level 0 to 4 that have been published each year since
576 2003. Panel C summarises variable reporting per assay level. Panel D shows the percentage of assays
577 reporting a specific variable (colour-coded according to level). Panel E shows the minimum criteria

578 necessary to reach each level of the validation scale, and the percentage of Level 0 to 4 assays that did
579 not report these. All variable abbreviations are listed in SI1.
580
581 **Figure 3:** Classification tree analysis identifying the criteria distinguishing assays at different levels of the
582 validation scale. The conditions along the branches show the criteria on which the dataset is split.
583 Numbers in coloured leaves show the validation level of the assays in the respective leaf. Numbers
584 below the leaves represent the number of assays per validation level, summarised inside an individual
585 leaf. The displayed percentage is the proportion of assays summarised per leaf.
586
587 **Box 1:** Examples of assays rated at Levels 0 to 4. The vertical tile plot shows which of the minimum
588 criteria the assay fulfils (yellow tiles), and the bar chart gives the scoring percentage (i.e. the proportion
589 of variables that were tested or reported) for each of the variable blocks. Bars are coloured according to
590 the score obtained from a block with dark purple coding for "no validation" and yellow coding for
591 "comprehensive validation".

Table 1:

General issues	*In silico* validation	*In vitro* validation	*In situ* validation
Applying assay to a new geographic location	Test non-target sequences available on public/custom databases from this location	Test tissue from non-target species present at this location, including species present at original location if populations are genetically diverse	
Sparse reference data	Where possible, Sanger sequence tissue DNA from target and non-target species to generate reference sequences	Extensively test assay on tissue available for target and non-target species, i.e. test extracts from multiple individuals	Extensively test assay at sites where target species co-occurs with non-target species lacking reference sequences, and at sites where only non-target species lacking reference sequences occur

In vitro validation issues	Troubleshooting guidelines	*In situ* validation issues	Troubleshooting guidelines
Poor amplification efficiency	Optimise reaction volume	**Expected species presence/absence not confirmed**	Survey sites with conventional tools and eDNA metabarcoding if possible
	Optimise cycle number	**False positives**	Discard samples corresponding to contaminated controls; consider lack of specificity; Sanger sequence results
	Optimise primer-probe concentration		Resample with more stringent decontamination measures in place
	Optimise annealing temperature using gradient PCR	**False negatives**	Check for inhibition and treat samples if inhibited; consider timing and spacing of sampling
	Optimise technical replication		Increase technical replication
	Check pipetting accuracy		Resample and increase biological replication and volume of water collected
Poor specificity	Increase annealing temperature	**Poor quality Sanger sequencing**	Purify amplicons prior to sequencing
	Add hydrolysis probe		Concentrate amplicons prior to sequencing
	Perform melting curve analysis	**Inhibition (determined by failed/skewed amplification of IPC or non-target assay)**	Use a DNA extraction kit that includes an inhibitor removal step (e.g. mu-DNA, Qiagen QIAamp DNA Stool Mini Kit, Qiagen DNeasy PowerWater Kit, Qiagen DNeasy PowerSoil Kit)
	Consider redesign		Apply an inhibitor removal kit (e.g. Zymo) to samples
Low sensitivity (see also Klymus et al., 2019)	Ensure enough replication of fresh standards used when establishing the Limit of Detection		Use PCR reagents designed to handle inhibition (e.g. TaqMan Environmental Master Mix, Bovine Serum Albumin)
	Use TE buffer and tRNA to make standard dilutions, not molecular grade water		
	Use low retention tubes and pipette tips when preparing standards to prevent adsorption to plastic		
	If possible, switch to PCR platform with greater sensitivity		
	Consider redesign		

Table 2:

Trade-offs	Endpoint PCR-gel	Endpoint PCR-CE	qPCR-ID	qPCR-probe	dPCR- ID	dPCR-probe	LAMP
Quantification	no	yes	yes	yes	yes	yes	yes, with real-time monitoring
LOD (sensitivity)	medium - high DNA concentration (depending on which intercalating dye)	low - medium DNA concentration (with a well designed assay)	low DNA quantities	low DNA quantities	absolute DNA quantities	absolute DNA quantities	medium DNA concentration
Specificity	medium	medium	medium	high	medium	high	high
Cost reagents	low	medium	high	high	high	very high	medium
Cost equipment	low	medium	medium	medium	high	high	low - medium (depending on LAMP cycler)
Multiplex possible	yes, if different sizes are used	yes, if different sizes or dyes are used	difficult	yes, if different dyes are used	difficult	yes, if different dyes are used	yes
Time (PCR to data)	slow	slow	medium	medium	slow	slow	fast (15-30 min)
Assay type transferable	yes	yes	yes	yes	yes to qPCR	yes to qPCR	not easily
Effort to design assay	low - medium	low - medium	high	high	high	high	high
Specialised software	advisable	advisable	advisable	required	advisable	required	required
Sequencing confirmation	yes	yes	yes	yes	limited	limited	limited

Table 3:

Validation level	Variable blocks	Minimum criteria
Level 1	*in silico* analysis	target species
	target tissue testing	target tissue
	target tissue PCR	primer (and probe) sequence
Level 2	comprehensive reporting of PCR conditions	DNA extract volume in PCR
	in vitro testing on closely related non-target species	any in vitro non-target testing
Level 3	extraction method performed on eDNA samples	method of extraction
	concentration of eDNA from environmental sample	filter type or precipitation chemicals
	detection obtained from environmental samples	detection from an environmental sample (artificial or natural habitat)
Level 4	Limit of Detection (LOD)	LOD determined
	extensive field testing of environmental samples	multiple locations or multiple samples
	in vitro testing on co-occurring non-target species	any advanced *in vitro* testing
Level 5	comprehensive specificity testing	non-co-occurring/closely related species checked from *in silico*
	detection probability estimation from statistical modelling	any effort made towards detection probability estimation
	understanding ecological and physical factors influencing eDNA in the environment	any factor influencing eDNA in the environment tested

Figure 1

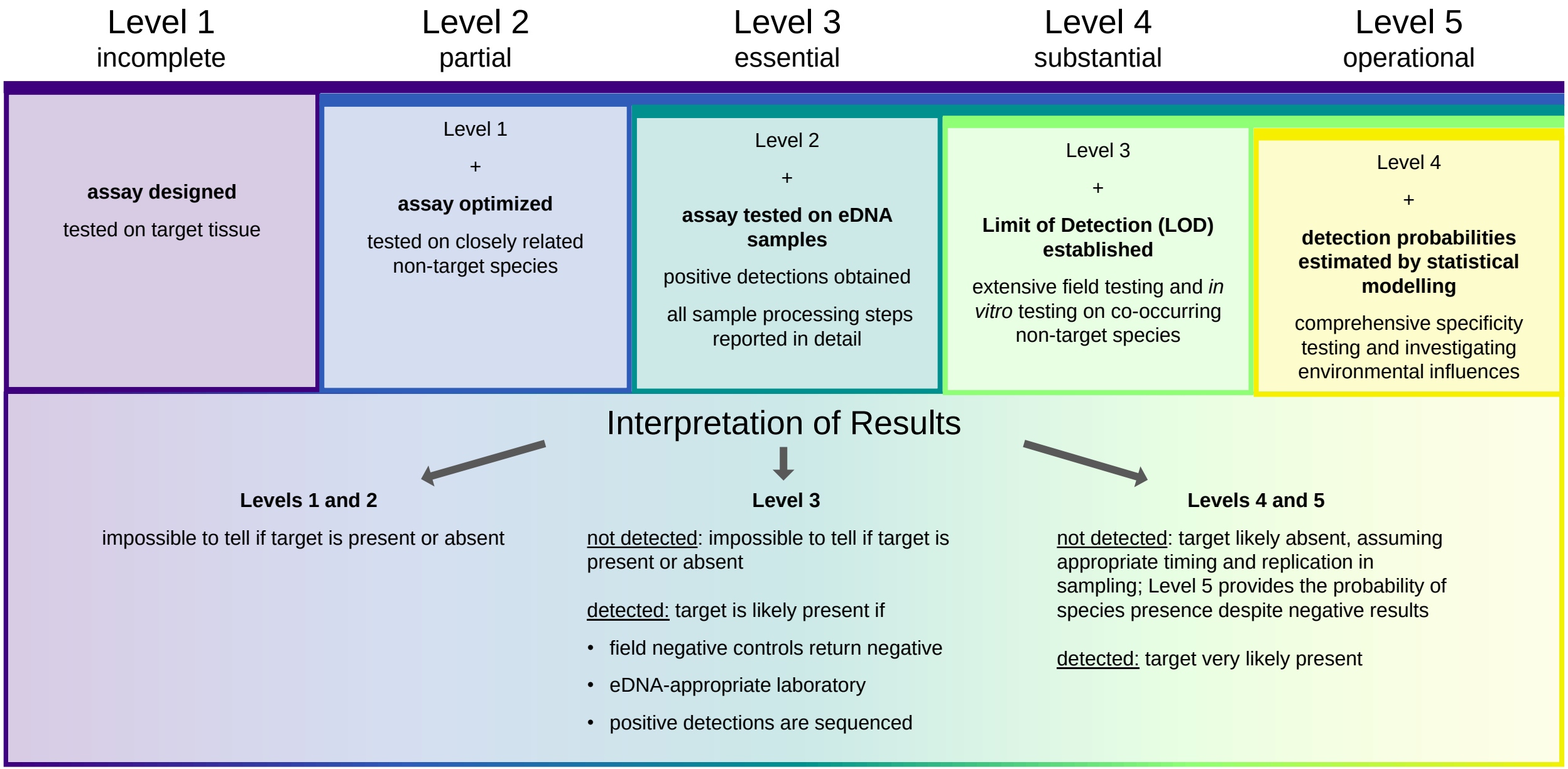

Level 1
incomplete
Level 2
partial
Level 3
essential
Level 4
substantial
Level 5
operational
assay designed
tested on target tissue
Level 1
+
assay optimized
tested on closely related non-target species
Level 2
+
assay tested on eDNA samples
positive detections obtained
all sample processing steps reported in detail
Level 3
+
Limit of Detection (LOD) established
extensive field testing and in vitro testing on co-occurring non-target species
Level 4
+
detection probabilities estimated by statistical modelling
comprehensive specificity testing and investigating environmental influences
Interpretation of Results
Levels 1 and 2
impossible to tell if target is present or absent
Level 3
not detected: impossible to tell if target is present or absent
detected: target is likely present if
• field negative controls return negative
• eDNA-appropriate laboratory
• positive detections are sequenced
Levels 4 and 5
not detected: target likely absent, assuming appropriate timing and replication in sampling; Level 5 provides the probability of species presence despite negative results
detected: target very likely present

Figure 2:

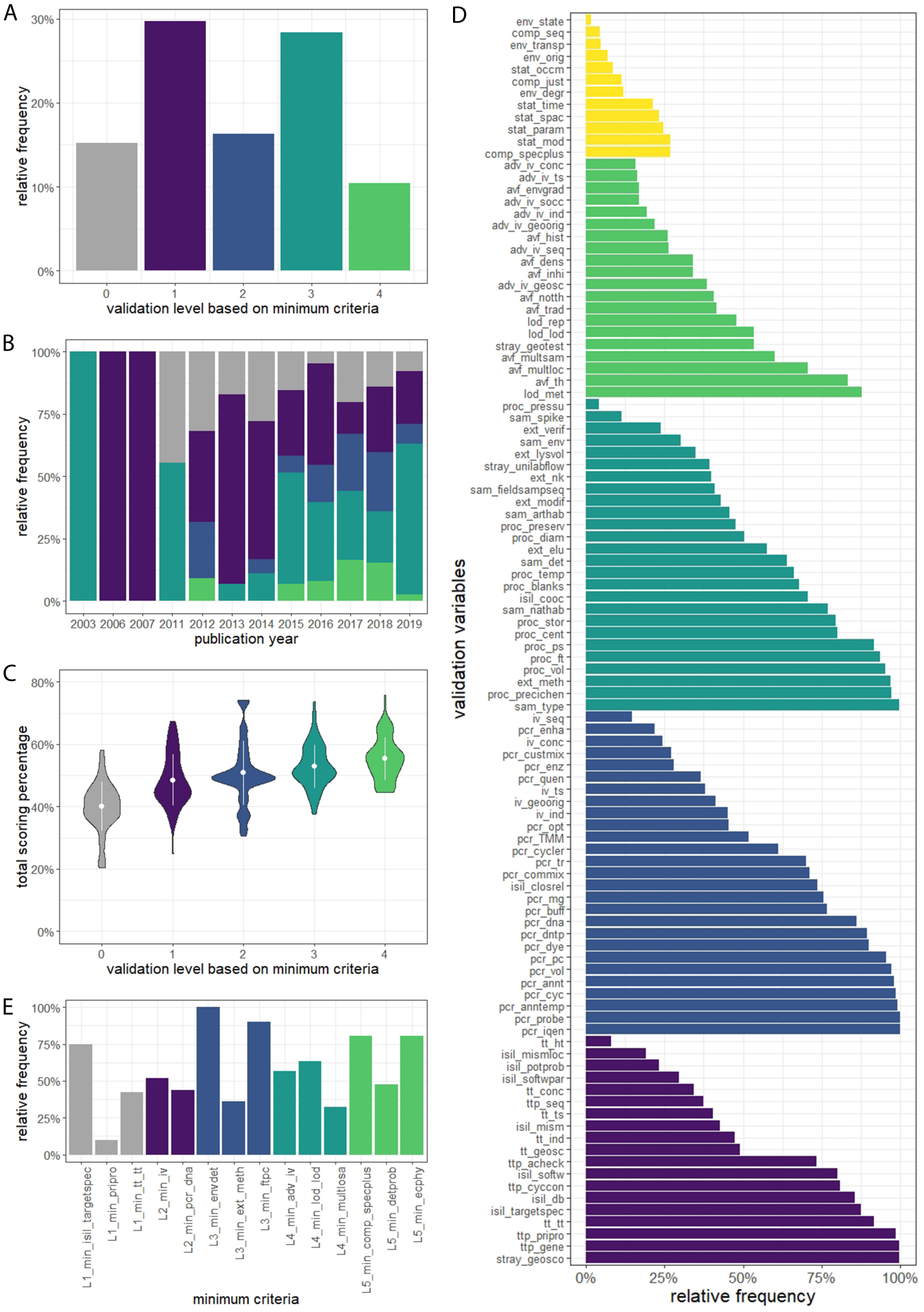
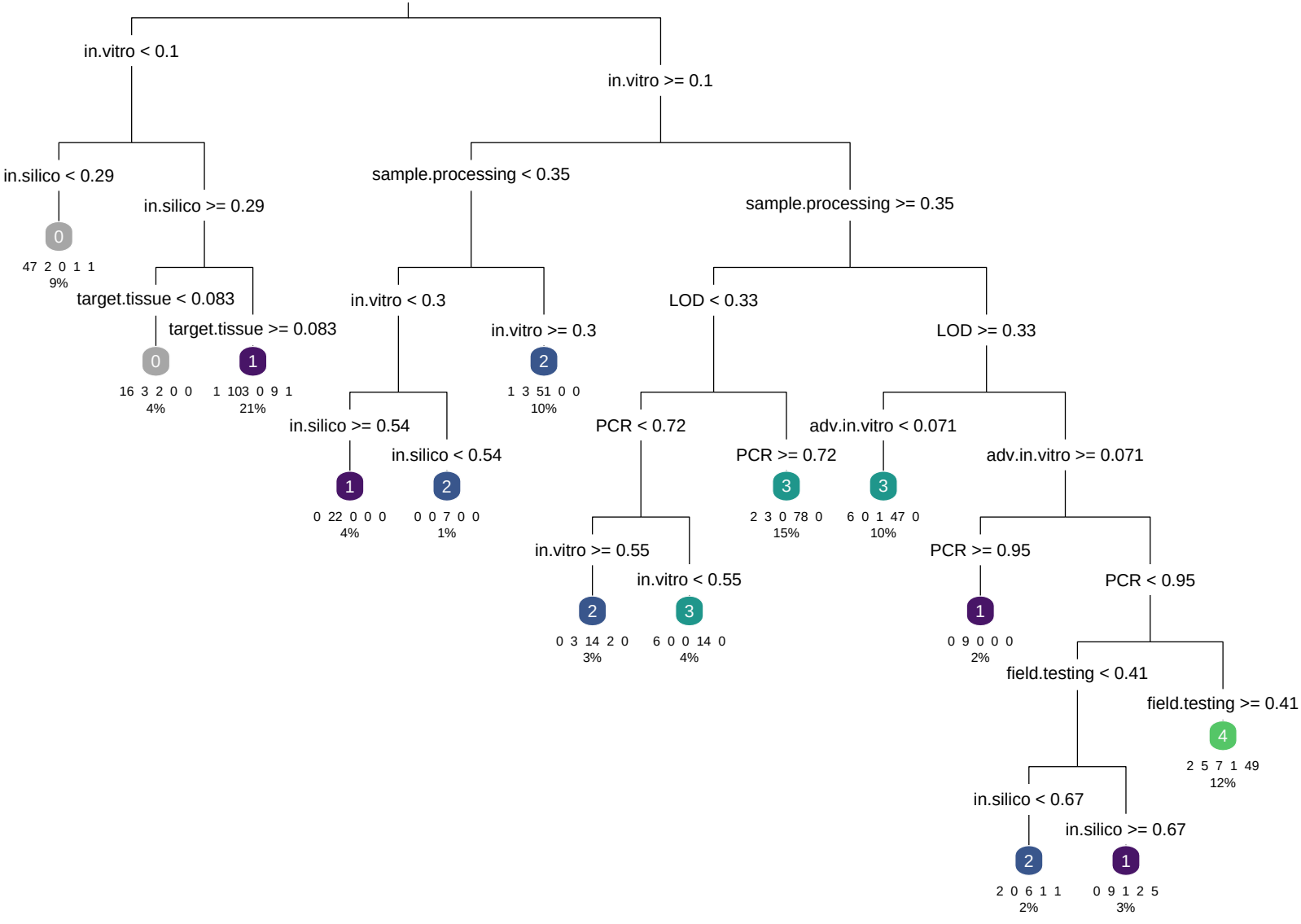
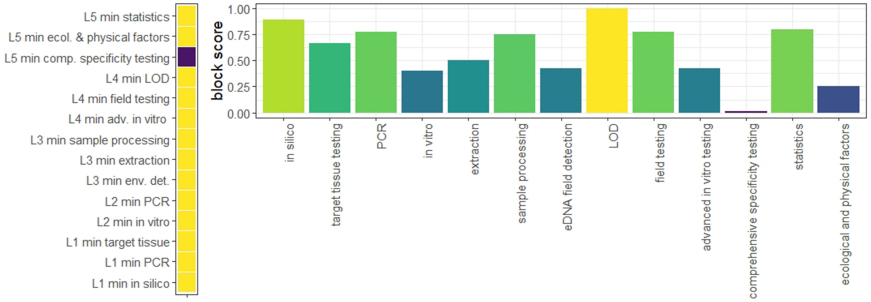

A
relative frequency
30%
20%
10%
0%
0
1
2
3
4
validation level based on minimum criteria
B
relative frequency
100%
75%
50%
25%
0%
2003 2006 2007 2011 2012 2013 2014 2015 2016 2017 2018 2019
publication year
C
total scoring percentage
80%
60%
40%
20%
0%
0
1
2
3
4
validation level based on minimum criteria
E
relative frequency
100%
75%
50%
25%
0%
L1_min_isil_targetspec
L1_min_pripro
L1_min_tt_tt
L2_min_iv
L2_min_pcr_dna
L3_min_envdet
L3_min_ext_meth
L3_min_ftpc
L4_min_adv_iv
L4_min_lod_lod
L4_min_multiosa
L5_min_comp_specplus
L5_min_detprob
L5_min_ecphy
minimum criteria
D
validation variables
env_state
comp_seq
env_transp
env_orig
stat_occm
comp_just
env_degr
stat_time
stat_spac
stat_param
stat_mod
comp_specplus
adv_iv_conc
adv_iv_ts
avf_envgrad
adv_iv_socc
adv_iv_ind
adv_iv_geoorig
avf_hist
adv_iv_seq
avf_dens
avf_inhi
adv_iv_geosc
avf_notth
avf_trad
lod_rep
lod_lod
stray_geotest
avf_multsam
avf_multloc
avf_th
lod_met
proc_pressu
sam_spike
ext_verif
sam_env
ext_lysvol
stray_unilabflow
ext_nk
sam_fieldsampseq
ext_modif
sam_arthab
proc_preserv
proc_diam
ext_elu
sam_det
proc_temp
proc_blanks
isil_cooc
sam_nathab
proc_stor
proc_cent
proc_ps
proc_ft
proc_vol
ext_meth
proc_precichen
sam_type
iv_seq
pcr_enha
iv_conc
pcr_custmix
pcr_enz
pcr_quen
iv_ts
iv_geoorig
iv_ind
pcr_opt
pcr_TMM
pcr_cycler
pcr_tr
pcr_commix
isil_closrel
pcr_mg
pcr_buff
pcr_dna
pcr_dntp
pcr_dye
pcr_pc
pcr_vol
pcr_annt
pcr_cyc
pcr_anntemp
pcr_probe
pcr_iqen
tt_ht
isil_mismloc
isil_potprob
isil_softwpar
tt_conc
ttp_seq
tt_ts
isil_mism
tt_ind
tt_geosc
ttp_acheck
isil_softw
ttp_cyccon
isil_db
isil_targetspec
tt_tt
ttp_pripro
ttp_gene
stray_geosco
0%
25%
50%
75%
100%
relative frequency

Figure 3:

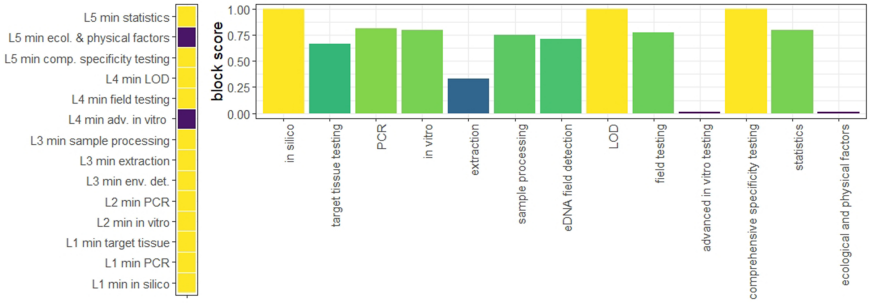

in.vitro < 0.1
in.vitro >= 0.1
in.silico < 0.29
in.silico >= 0.29
0
47 2 0 1 1
9%
target.tissue < 0.083
target.tissue >= 0.083
0
16 3 2 0 0
4%
1
1 103 0 9 1
21%
sample.processing < 0.35
sample.processing >= 0.35
in.vitro < 0.3
in.vitro >= 0.3
2
1 3 51 0 0
10%
in.silico >= 0.54
in.silico < 0.54
1
0 22 0 0 0
4%
2
0 0 7 0 0
1%
LOD < 0.33
LOD >= 0.33
PCR < 0.72
PCR >= 0.72
3
2 3 0 78 0
15%
adv.in.vitro < 0.071
adv.in.vitro >= 0.071
3
6 0 1 47 0
10%
in.vitro >= 0.55
in.vitro < 0.55
2
0 3 14 2 0
3%
3
6 0 0 14 0
4%
PCR >= 0.95
PCR < 0.95
1
0 9 0 0 0
2%
field.testing < 0.41
field.testing >= 0.41
4
2 5 7 1 49
12%
in.silico < 0.67
in.silico >= 0.67
2
2 0 6 1 1
2%
1
0 9 1 2 5
3%

Box 1:

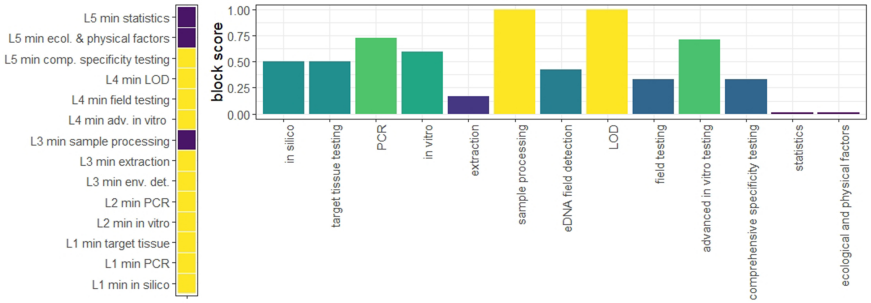
Level 4: Assay no. 165 detecting brook trout, *Salvelinus fontinalis*. Assay information was published in paper nos. 291, 12, 145, 295, 144, with latest publication in 2018. Despite low validation efforts for *in vitro* testing and field detection, the only minimum criterion that prevented this assay attaining Level 5 was "comprehensive specificity testing".



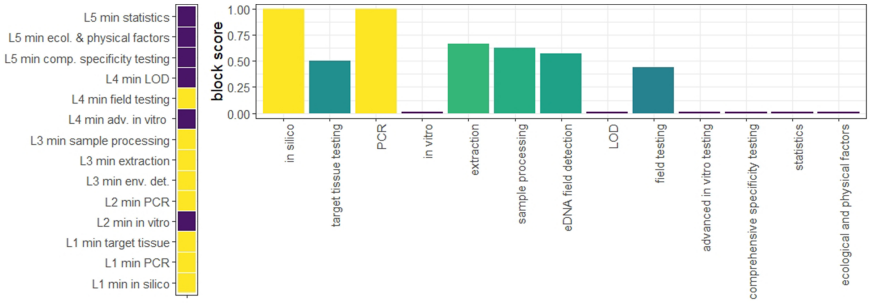
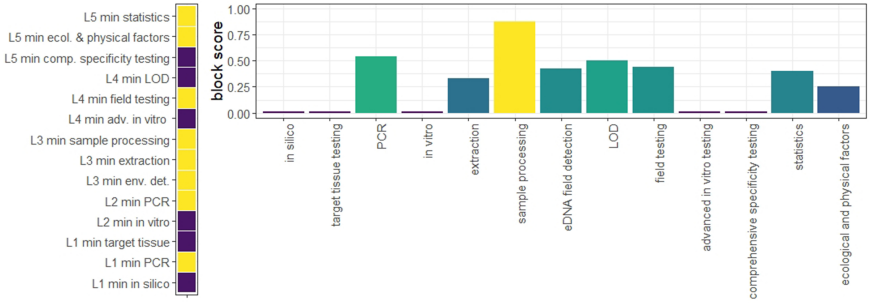
Level 3: Assay no. 473 detecting the crucian carp, *Carassius carassius*. Assay information was published in paper no. 119 in 2019. No advanced *in vitro* testing was conducted for this assay, but validation efforts were substantial for most thematic blocks, especially "*in silico* testing", "LOD determination", and "comprehensive specificity testing".



Level 2: Assay no. 270 detecting common eelgrass, *Zostera marina*. Assay information was published in paper no. 117 in 2018. Despite low validation efforts in many thematic blocks, only the minimum criterion of "sample processing" prevented this assay from reaching Level 4 on the validation scale



Level 1: Assay no. 315 detecting the waterlouse, *Asellus aquaticus*. Assay information was published in paper no. 163 in 2015. With additional *in vitro* testing, this assay would have attained Level 3. A clear decrease in validation efforts is visible for variable blocks associated with higher levels of the validation scale.



Level 0: Assay no. 343 detecting *Daphnia magna*. Assay information was published in paper no. 243 in 2018. Despite some advanced validation, this assay does not fulfill the minimum criterion of "*in silico* testing" and "target tissue testing". Therefore, it could not be classed as Level 1 on the validation scale.