



HAL
open science

Comment faciliter l'identification de l'entomofaune ? : Construction, évaluation et amélioration de clés d'identification numériques

Mathilde Delaunay

► To cite this version:

Mathilde Delaunay. Comment faciliter l'identification de l'entomofaune ? : Construction, évaluation et amélioration de clés d'identification numériques. Biodiversité et Ecologie. Museum national d'histoire naturelle - MNHN PARIS, 2019. Français. NNT : 2019MNHN0009 . tel-02868459

HAL Id: tel-02868459

<https://theses.hal.science/tel-02868459v1>

Submitted on 15 Jun 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

MUSÉUM NATIONAL



D'HISTOIRE NATURELLE

École Doctorale Sciences de la

Nature et de l'Homme – ED 227

Année 2019

N° attribué par la bibliothèque



THÈSE

Pour obtenir le grade de

DOCTEUR DU MUSÉUM NATIONAL D'HISTOIRE NATURELLE

Spécialité : Écologie et Biodiversité

Présentée et soutenue publiquement par

Mathilde DELAUNAY

Le vendredi 10 mai 2019

Comment faciliter l'identification de l'entomofaune ? Construction, évaluation et amélioration de clés d'identification numériques

Sous la direction de :

Régine VIGNES-LEBBE, Professeure, SU

Romain NATTIER, Maître de Conférences, MNHN

JURY :

M. BEISEL Jean-Nicolas	Professeur, Université de Strasbourg, France	Rapporteur
M. CASTELLA Emmanuel	Maître d'Enseignement et de Recherche, Université de Genève, Suisse	Rapporteur
Mme CHARPIN Marie	Maître de Conférences, Université Clermont Auvergne, Clermont-Ferrand, France	Examinatrice
Mme DAJOZ Isabelle	Professeure, Université Paris Diderot, Paris, France	Examinatrice
M. DUBUISSON Jean-Yves	Professeur, Sorbonne Université, Paris, France	Examineur
Mme TURCATI Laure	Ingénieure de Recherche, Sorbonne Université, Paris, France	Examinatrice
M. NATTIER Romain	Maître de Conférences, Muséum national d'Histoire naturelle, Paris, France	Directeur de thèse
Mme VIGNES-LEBBE Régine	Professeure, Sorbonne Université, Paris, France	Directrice de thèse

Remerciements

Mes plus sincères remerciements vont en premier lieu à mes directeurs, Romain et Régine, sans qui je n'aurais jamais eu l'opportunité de réaliser un doctorat, ce travail varié, prenant, passionnant, formateur. Je salue vos immenses qualités d'encadrants, que les circonstances particulières de cette thèse ont malheureusement mises à l'épreuve. Merci de m'avoir à tout moment soutenue, motivée, accompagnée, donné d'excellents conseils scientifiques et personnels, d'avoir supporté les retards dans le travail, de n'avoir jamais lâché. Vous êtes super.

Merci à l'OPIE d'avoir été un partenaire de choix pour la réalisation de ce travail. Merci à la Région Ile-de-France et à son dispositif PICRI pour le financement de ce doctorat.

Un grand merci aux membres de mon comité de thèse : Michel Brulin, Claire David, Samuel Jolivet, Laurent Pelozuelo et Philippe Usseglio-Polatera. Vos avis éclairés et vos suggestions toujours pertinentes ont apporté beaucoup à ce travail de recherche.

Un grand merci également aux membres de mon jury, pour vous être investis dans l'évaluation de ma thèse. Beaucoup de spécialités scientifiques sont représentées dans ce jury et j'espère satisfaire les attentes de chacun.

J'espère garder des liens avec le Muséum après ma thèse, cet établissement qui m'a accueillie depuis mon tout premier stage en 2009, et qui a une place toute particulière dans mon cœur. Ce lieu de connaissances et d'histoire s'intéresse à ce qu'il y a de plus fascinant sur terre : la biodiversité et l'évolution. Le Jardin des Plantes, le bâtiment de l'Herbier et la cantine sont ma seconde maison. Je suis très reconnaissante d'avoir pu y travailler pendant toutes ces années et d'avoir eu l'occasion d'y côtoyer les personnes formidables qui y travaillent.

Je remercie l'UMR ISYEB et les équipes BIPEM et LIS pour m'avoir si bien accueillie et accompagnée dans mon travail. Un grand merci à l'école doctorale 227, en particulier Nathalie Machon et Jérôme Sueur, qui ont été particulièrement compréhensifs et d'une immense gentillesse.

Je garderai un excellent souvenir de mes quelques passages à la DRIEE. Merci Claire et Jean-Luc pour les sessions de terrain formatrices et parfois épiques ! Ça peut être sacrément

dangereux les suivis de biodiversité ! Un grand merci au laboratoire d'hydrobiologie pour la mise à disposition de ces centaines de spécimens d'éphémères.

Je remercie vivement Sylvain Bouquin, expert en Xper (haha). Sans toi aucune analyse n'aurait été possible ! Merci pour ton efficacité et ta réactivité.

Merci à Colin Fontaine, Nicolas Deguines et Mathieu de Flores pour les données et les informations relatives au programme Spipoll. Je lui souhaite encore des décennies de succès !

Merci à Anne-Caroline Prévot-Julliard et à Marine Levé pour m'avoir ouverte au monde des sciences humaines et sociales et conseillé pour une partie de mon travail.

Merci à Armelle Caron pour son accompagnement jusque pendant mon doctorat. Le mastère PPSE a été une révélation personnelle et orientera ma carrière professionnelle.

Un immense merci à mes stagiaires : Aurélia Dubois, Mathieu Faure-Brac et Paola Campos. J'ai eu la chance de co-encadrer des étudiants impliqués, appliqués, et passionnés. Trois personnalités riches à qui je souhaite une très belle réussite dans la recherche, les sciences naturelles et la protection de la biodiversité (c'est déjà très bien parti !).

Merci à Guillaume Doucet, Jacques Le Doaré et Alexandre Ruffoni pour leur expertise sur les odonates et plécoptères.

Un très grand merci à Michel Brulin, puits de science sur les éphémères. Merci pour tout le temps que vous m'avez accordé, votre envoi d'une collection de référence (quel luxe !), vos anecdotes passionnantes et nos discussions interminables à base de "touffe de filaments trachéens" et de branchies "vraiment très très *slightly* asymétriques" ! Je vous souhaite encore de nombreuses années de passion et de pouvoir mener à bien tous vos projets.

Je remercie le groupe Perla, ainsi que tous les photographes amateurs et professionnels (notamment sur Flickr) pour m'avoir autorisée à utiliser leurs images dans les clés d'identification. Merci également à Christophe Brochard pour ses superbes clichés de larves d'odonates.

Un grand merci aux testeuses et testeurs de la clé des odonates. Il en faut du monde pour mener à bien un projet de thèse ! Merci pour votre temps, votre enthousiasme et votre patience.

Les amis du bureau. Il est en passé du monde en quatre ans dans cette grande salle Cosson ! Vous avez tous été une compagnie fantastique, chaque départ fut un déchirement. Pour ceux qui étaient là sur la fin de ma thèse : vous avez été mon ancre, ma bouée ! Vous êtes

fantastiques ! À mon tour de quitter le bureau, pleine de nostalgie et de gratitude pour des collègues aussi adorables qui sont également pour la plupart devenus de vrais amis. Les séjours à Barcelone et Trouville resteront de formidables souvenirs. En espérant n'oublier personne : Anaëlle, Camille, Émeline, Felipe, Gladys, Jérôme, Julie, Karen, Kévin, Lauriane, Laurie, Lucie, Lucile, Mathieu, Michèle, Oudomphone, Patricia, Simon, Tian, Wafa.

Un merci tout particulier à Catherine G., Micaela N. et Valérie A.

The Oxford people: you rock! Gemma, James, John, Marie, and everybody else! Love you!

Une tendre pensée naturaliste pour Timarcha, toutes les Timarchiennes et tous les Timarchiens ! En particulier Catherine, Isaure, Paul, Romain, et tous les autres membres du bureau pendant toutes ces années. En espérant fêter les 50 ans de l'association et même encore plus !

Attention, on passe aux copains. Je remercie du fond du cœur :

- ♥ Laetitia et Nico : Honnêtement, sans vous je ne serais pas arrivée au bout. Vous êtes des amis très très précieux. Tipoulpe a un air encore plus satisfait quand il pense à vous.
- ♥ Lucie : Partenaire de bureau, de Timarcha, de Mada, je t'ai connue avec les cheveux jusqu'aux fesses, c'est dire si ça date ! Je te souhaite plein de bonheur, en Chine et ailleurs.
- ♥ Charlichou d'amour : Unique, irremplaçable, je suis tellement contente que tu sois rentrée !
- ♥ Rémi : Je te kiffe espèce d'INTJ :-)
- ♥ Ninon : THE Ninon. Merci pour tous ces déjeuners à la cantine et pour tout le reste depuis si longtemps. Heureusement tu vas re-traverser l'Atlantique très bientôt !
- ♥ Morgane et Perrine : Je vous kiffe, cantine-buddies et bien plus ! Loukiiii !!! Perrine, j'espère suivre encore longtemps tes aventures sur la route, sur la mer, dans les airs !
- ♥ François : Je pense à toi mon Clarac qui te pèles les miches là-haut. Hâte de te revoir et de faire du navion.
- ♥ Houda : Ton soutien et ton écoute ont été irremplaçables. L'Afrique du Sud c'était INCROYABLE.
- ♥ Lucas : Avoue Legender je n'ai (presque) que des bons goûts musicaux. C'était stylé ce concert à deux. Et l'Afrique du Sud !!!!

- ♥ Boris : Merci pour ton accueil en Québécoisie. Elle est cool ta ville.
- ♥ Mich-Mich : La légende absolue. Ces remerciements sont tellement longs qu'il y a probablement de quoi faire plein de stats débiles.
- ♥ Malcolm : Une autre légende. Vivent les montages ultra cool, les gifs intempestifs et les réponses en 2 syllabes.
- ♥ Paul : Coloc' d'ENFP !! Basilic, *Ragnagna* & muuuuultipass !
- ♥ Raphi, mon grand frère de cœur. Comment tu fais pour me supporter depuis si longtemps ?! Merci pour TOUT. Je te love + + +
- ♥ Agnès, Alix, Aurélie, Ben, Camille, Chaaaaaarles, Dim "comme les slips", NéotÉli, Florian Guibert, Jérémy, Kévin, Laura, Lou, Louma, Ludwik, Marc, Marta, Rémi L., Séverin, Tifenn, Thomas, Rineau *et al.*, Victorche, Wahid...
- ♥ Et tous mes autres groupes-sœurs et groupes-frères ! Un lien pareil c'est pour la vie. VAVE !!!

Je remercie Antoine, mes parents et mes grands-parents, ainsi que le reste de ma famille, pour leur amour et leur soutien inconditionnels depuis toujours. Je vous aime.

Do, tu mérites un livre entier de remerciements. Ta présence depuis toutes ces années est la chose la plus précieuse dans ma vie. Je t'aime grand comme ça, et même encore beaucoup plus.

À Clément

Pour toujours dans mon cœur et dans ma vie, frérot ♥

Table des matières

Préambule : contexte de la thèse	2
Chapitre 1 – Introduction	3
1.1. Décrire, identifier et étudier la biodiversité.....	4
1.1.1. État et enjeux de la biodiversité.....	4
1.1.2. Comment décrire et identifier la biodiversité ?.....	5
1.1.3. Les sciences participatives : l'implication des citoyens pour l'étude de la biodiversité.....	7
1.2. Le cas particulier de l'entomofaune.....	13
1.2.1. Les insectes : des organismes clés très diversifiés mais en fort déclin.....	13
1.2.2. Les difficultés rencontrées dans l'identification des insectes	15
1.3. L'entomofaune étudiée dans ce travail de thèse : des organismes d'intérêts variés.....	17
1.3.1. Les insectes aquatiques de la région Ile-de-France	17
1.3.2. Les coccinelles de France métropolitaine.....	23
1.3.3. Les insectes pollinisateurs de France métropolitaine	26
1.4. Objectifs et problématiques de la thèse	28
Chapitre 2 – Caractéristiques et historique des clés d'identification.....	31
2.1 Les clés d'identification à accès simple	32
2.1.1 Fonctionnement et particularités des clés d'identification à accès simple.....	32
2.1.2 A l'origine des clés d'identification à accès simple.....	34
2.1.3 Types de caractères utilisés dans les clés	36
2.1.4 Clés numériques à accès simple.....	38
2.2 Les clés d'identification numériques à accès libre.....	39
2.2.1 Principes et caractéristiques.....	39
2.2.2 Pourquoi ce type de clés ? Avantages et inconvénients des clés d'identification numériques à accès libre.....	41
2.2.3 L'apport des clés numériques à accès multiples pour les sciences participatives : l'exemple du Spipoll	43
Chapitre 3 – Comment construire des clés d'identification numériques ?.....	47
3.1 Taxons étudiés et logiciels utilisés	48
3.1.1 Échantillonnage taxonomique pour l'entomofaune	48
3.1.2 Xper, une suite d'outils pour la construction et l'analyse de bases de données et de clés d'identification.....	57
3.2 Méthodologie pour la construction de clés d'identification numériques sous Xper ³	62
3.2.1 Modèle descriptif.....	62
3.2.2 Caractères morfo-anatomiques.....	66
3.3 Construction de clés Xper ³ pour les insectes dans le cadre de la thèse	67

Chapitre 4 – Comment évaluer des clés d’identification numériques ?.....	71
4.1 Évaluation du contenu des clés d’identification	72
4.1.1 Contenu général des clés d’identification	72
4.1.2 Pouvoir discriminant des descripteurs.....	73
4.2 Évaluation de l’utilisation des clés d’identification	77
4.2.1 Le système Xperience.....	77
4.2.2 Analyse des démarches d’identification pour la clé Spipoll	80
4.2.3 Test de l’utilisation de la clé des larves d’odonates d’Ile-de-France	102
Chapitre 5 – Comment améliorer des clés d’identification numériques ?.....	117
5.1 Intégration de critères écologiques dans les clés d’identification.....	118
5.1.1 Le cas des larves d’insectes aquatiques : odonates et éphémères.....	118
5.1.2 L’apport des critères écologiques	123
5.1.3 Une étude moléculaire pour affiner les critères écologiques	133
5.2 Prise en compte de l’expérience des utilisateurs des clés d’identification.....	142
5.2.1 Pour la clé “Insectes” du Spipoll.....	142
5.2.2 Pour la clé des larves d’odonates d’Ile-de-France.....	145
5.3 Questionnaire et expérience des volontaires au Spipoll	147
5.3.1 Contexte	147
5.3.2 Matériel et méthodes.....	150
5.3.3 Résultats et discussion.....	151
5.3.4 Pistes d’amélioration de la clé d’identification du Spipoll	167
Chapitre 6 – Discussion générale et conclusion	171
6.1 Bilan du travail de thèse	172
6.2 De nouvelles méthodes d’identification.....	177
6.2.1 Les logiciels d’identification automatique	177
6.2.2 L’identification par séquençage rapide d’ADN	181
6.3 L’utilisation de critères non morphologiques dans les clés d’identification.....	182
6.4 Conclusion	185
RÉFÉRENCES	188
ANNEXES	202
Annexe 1 : Article publié dans <i>ZooKeys</i>	202
Annexe 2 : Descripteurs dans la clé des larves d’odonates (T1)	222
Annexe 3 : Descripteurs dans la clé des larves d’éphémères (T2)	224
Annexe 4 : Descripteurs dans la clé des éphémères imagos (T3).....	226
Annexe 5 : Descripteurs dans la clé des larves de plécoptères (T4).....	228
Annexe 6 : Descripteurs dans la clé des plécoptères imagos (T5).....	229

Annexe 7 : Descripteurs dans la clé des coléoptères aquatiques (T6).....	230
Annexe 8 : Descripteurs dans la clé des coccinelles de France (T7).....	231
Annexe 9 : Article (à soumettre dans <i>PLoS ONE</i> ou <i>PeerJ</i>) sur les clés des larves aquatiques et le test de la clé des larves d'odonates.....	232
Annexe 10 : Détails des 85 descripteurs de la base de données Xper ³ du Spipoll	255
Annexe 11 : Liste des descripteurs de la base classés selon leur pouvoir discriminant d'après l'indice Xper (le plus discriminant en premier)	258
Annexe 12 : Poster sur les caractères écologiques dans la clé des odonates larves (tous stades larvaires).....	261
Annexe 13 : Questionnaire	262
Annexe 14 : Poster sur l'expérience des volontaires aux sciences participatives.....	268

Préambule : contexte de la thèse

Ce doctorat a été financé par la Région Ile-de-France, dans le cadre du dispositif PICRI (Partenariats institutions-citoyens pour la recherche et l'innovation). Les projets PICRI sont des études scientifiques réalisées en collaboration entre des laboratoires de recherche et des acteurs de la société civile, faisant appel aux expertises propres aux différentes structures.

Ce projet était donc porté par le Muséum national d'Histoire naturelle (MNHN) et l'Office pour les insectes et leur environnement (OPIE), une association française ayant pour mission de favoriser la connaissance des insectes dans leurs milieux naturels (inventaires, études scientifiques, animation scientifique pour les professionnels et le grand public...). Au Muséum, j'ai été hébergée par l'Unité Mixte de Recherche ISYEB (Institut de SYstématique, Évolution, Biodiversité – UMR 7205 CNRS, MNHN, SU, EPHE) ; je faisais partie des équipes BIPEM (Biologie intégrative des populations, évolution moléculaire) et LIS (Laboratoire informatique et systématique).

Une partie du travail de thèse a été réalisée en "co-construction" entre différents acteurs – étudiante, chercheurs, experts naturalistes, entomologistes amateurs – dans l'esprit des projets PICRI ; c'est le cas par exemple de la construction des clés d'identification des insectes aquatiques et l'utilisation de bases de données issues de programmes développés par des associations et agences pour la biodiversité.

Au cours de mon doctorat, j'ai également poursuivi et achevé une formation que je suivais depuis plusieurs mois avant de débiter ce travail de thèse : le Mastère Spécialisé "Politiques Publiques et Stratégies pour l'Environnement" (PPSE), formation de niveau Bac+6 à AgroParisTech. J'ai effectué la "mission professionnelle" du Mastère au cours de mon doctorat, co-encadrée par Armelle Caron, Ingénieure de Recherche à AgroParisTech. Le travail effectué a donné lieu à la rédaction d'une "thèse professionnelle" intitulée "Citoyens et sciences participatives : motivations et perception de la biodiversité". Cette étude sera présentée et détaillée dans les chapitres 4 et 5 du présent document.

Chapitre 1 – Introduction

1.1. Décrire, identifier et étudier la biodiversité

1.1.1. État et enjeux de la biodiversité

La biodiversité (ou diversité biologique) est la diversité de la vie sur Terre. Le terme "biodiversité" est utilisé depuis les années 1980 et l'intérêt pour sa conservation ne cesse de croître, notamment depuis la Convention sur la diversité biologique qui fut adoptée lors du sommet de la Terre à Rio de Janeiro en juin 1992¹. La biodiversité s'appréhende à plusieurs échelles : les gènes, les espèces, les écosystèmes. Son importance numérique est très conséquente – **entre 1,2 et 1,9 millions d'espèces décrites** à ce jour (Chapman 2009, Mora *et al.* 2011, Fontaine *et al.* 2012a, Larsen *et al.* 2017) – et elle est essentielle pour l'homme. D'une part, il s'agit d'un **patrimoine** naturel que nous devons maintenir et protéger pour des raisons éthiques (Yamin 1995, Tilman 2000). D'autre part, la biodiversité apporte des services aux sociétés humaines sans lesquels ces dernières ne pourraient se maintenir ; c'est ce qu'on appelle les **services écosystémiques** (nourriture, matériaux, régulation du climat, pollinisation, santé notamment grâce aux plantes médicinales, cycles globaux de l'eau et du carbone – pour ne citer qu'eux) (Costanza *et al.* 1997, Daily 1997, Worm *et al.* 2006).

Son importance fondamentale mérite et nécessite que la biodiversité soit étudiée et bien connue. Du fait du nombre colossal d'êtres vivants sur Terre, l'appréhension de leur fonctionnement à de multiples échelles (génétique, physiologie, anatomie et morphologie d'un organisme, population, espèce, écosystème, biome...) et de leurs particularités représente un travail soutenu et continu. Encore aujourd'hui, **l'étendue des connaissances restant à acquérir est incommensurable** (Jetz *et al.* 2012, Tyler *et al.* 2012, Chambers *et al.* 2017). En ce qui concerne la diversité spécifique, les recherches sont toujours en cours pour mieux évaluer l'importance numérique et la connaissance des millions d'espèces, et de nouveaux taxons sont découverts chaque jour dans le monde – environ 15 000 nouvelles espèces par an (Chapman 2009, Mora *et al.* 2011, Fontaine *et al.* 2012a, Larsen *et al.* 2017).

¹ Convention on Biological Diversity (1992) *International Legal Materials*, 21, p. 848.

L'étude des organismes vivants débute par une étape essentielle : savoir sans ambiguïté à quelle espèce (ou groupe d'êtres vivants) appartient le spécimen que l'on souhaite étudier, ou du moins être en mesure de séparer ce spécimen des autres taxons connus (Ruse 1969). C'est ce qu'on appelle **l'identification ou la détermination**. Une identification incorrecte peut fortement réduire la qualité des résultats d'une étude, ce qui entraîne de sérieuses conséquences lorsque cette dernière nécessite une précision et une fiabilité sans faille, par exemple pour des suivis de biodiversité et études d'impact, des contrôles de populations invasives ou ravageuses de cultures, des diagnostics de santé et soins médicaux (Marshall 2000). Il est donc fondamental d'avoir à notre disposition des outils de qualité pour la détermination des êtres vivants.

1.1.2. Comment décrire et identifier la biodiversité ?

La **taxonomie** – la science qui définit, décrit et classe les organismes vivants (Simpson 1961, Mayr 1969) – a grandement bénéficié des avancées technologiques récentes dans le domaine de la biologie moléculaire, et de nouvelles techniques qui ont ouvert d'immenses perspectives pour la description et la classification des organismes. De nombreux taxons peuvent maintenant être identifiés à partir de leur ADN (séquençage, barcoding, ADN environnemental...) (Hebert *et al.* 2003, Blaxter 2004, Shendure & Ji 2008, Bohmann *et al.* 2014). Malgré ces nouveaux outils, **la méthode d'identification "traditionnelle" basée sur des caractères morpho-anatomiques reste incontournable** (Hillis 1987, Will & Rubinoff 2004, Dayrat 2005, Page *et al.* 2005). En effet, les bases de données génétiques sont actuellement inexistantes ou très incomplètes pour de nombreux groupes (Padial *et al.* 2010, Ko *et al.* 2011, Osmundson *et al.* 2013, Grattepanche *et al.* 2018) ; la majeure partie des espèces reste décrite uniquement à partir de caractères morpho-anatomiques. D'autre part, **les techniques de séquençage et de barcoding sont encore coûteuses** (même si les coûts se réduisent rapidement), **et difficiles à utiliser sur le terrain**. Les connaissances en morpho-anatomie possédées par les scientifiques et experts amateurs restent irremplaçables pour la description et la détermination de la plupart des taxons (Will & Rubinoff 2004, Cowart *et al.* 2015). Les deux grandes approches – morpho-anatomique et

moléculaire – doivent se compléter pour permettre un travail d'identification le plus fiable et rapide possible.

Parmi les outils facilitant la détermination des spécimens à partir de critères morpho-anatomiques, les **collections scientifiques** de référence sont un ensemble de spécimens conservés, entre autres, au sein des musées d'histoire naturelle partout dans le monde (Shaffer *et al.* 1998, Suarez & Tsutsui 2004, Lister *et al.* 2011). Il s'agit historiquement de spécimens matériels – organismes ou parties d'organismes – récoltés et mis en collection depuis des centaines d'années au cours des missions de terrain et expéditions scientifiques, qui peuvent être vivants (graines, microorganismes...) ; mais il existe aussi des collections moléculaires (fragments d'ADN) et immatérielles, comme la Sonothèque du MNHN². Un nom (normalement, avec une nomenclature binomiale d'après le modèle de Linné) est associé à chaque spécimen mis en collection qui a pu être formellement identifié. Un souci actuel est que de nombreux spécimens restent encore à déterminer et identifier dans les collections (Fontaine *et al.* 2012b). D'autre part, l'utilisation de méthodes moléculaires telles que le barcoding est souvent difficile voire impossible sur des spécimens anciens ou des fossiles.

D'autre part des **publications scientifiques** de référence en taxonomie (description de nouvelles espèces ou monographies taxonomiques portant sur un taxon de rang supérieur à l'espèce), mais également des ouvrages, souvent richement illustrés, décrivent précisément la morphologie et l'anatomie des spécimens.

Parmi ces ouvrages, les **guides naturalistes et de terrain** sont conçus pour permettre l'identification d'un spécimen (à des niveaux de précision variables qui peuvent aller jusqu'aux sous-espèces et aux variations phénotypiques observées au sein d'une population), le plus rapidement et/ou aisément possible et en utilisant des caractères observables sur le terrain ou en laboratoire (Law & Lynch 1988). La difficulté d'utilisation, le rang taxonomique, l'étendue de l'aire géographique concernée ainsi que d'autres paramètres,

² sonotheque.mnhn.fr

peuvent varier pour chaque guide en fonction du public auquel il s'adresse (professionnels, naturalistes, grand public, scolaires...).

Enfin, il existe les **clés d'identification** (ou clés de détermination), qui proposent de suivre un chemin d'identification, c'est-à-dire une suite d'étapes au cours de laquelle l'utilisateur de l'outil renseigne les attributs du spécimen observé, caractère par caractère. Les clés d'identification sont souvent intégrées aux publications scientifiques ou aux guides de détermination. Leur objectif est d'aboutir à une détermination la plus précise et correcte possible d'un spécimen en s'attachant à la description de caractères morpho-anatomiques particuliers, qui permettent de distinguer efficacement plusieurs spécimens entre eux.

Le contexte actuel de changements climatiques et de crise de la biodiversité au niveau mondial renforce la nécessité de mettre des moyens financiers et humains suffisants pour l'étude approfondie des êtres vivants présents sur Terre (Fontaine *et al.* 2012a). Scientifiques et experts naturalistes combinent leurs efforts dans cette optique et depuis une trentaine d'années, le grand public est lui-même sollicité pour apporter sa contribution à l'effort général, notamment dans le cadre des sciences participatives.

1.1.3. Les sciences participatives : l'implication des citoyens pour l'étude de la biodiversité

Le grand public (citoyens de tout milieu et tout âge dès l'enfance) fait partie intégrante de la prise de conscience générale par rapport aux défis écologiques auxquels le monde fait face (Maloney & Ward 1973, Thompson & Barton 1994, Kollmuss & Agyeman 2002). Une grande proportion de la population est sensibilisée aux problématiques liées à la perte de biodiversité, et s'intéresse aux sciences de manière plus large³. Les sciences participatives consistent en la **participation d'un public non scientifique à la recherche, via notamment la collecte de données scientifiques, dans le cadre de protocoles adaptés et standardisés** (Irwin 1995, Couvet *et al.* 2008, Bonney *et al.* 2009).

³ Enquête Ipsos Sopra-Steria pour *La Recherche et Le Monde* (2016) par Etienne Mercier et Vincent Dusseaux – "Les Français et les sciences participatives"

Dans le cadre d'un programme de sciences participatives, les citoyens observateurs volontaires sont acteurs de la production de données qui sont ensuite vérifiées, validées, analysées et utilisées pour répondre à des questions scientifiques. Les observateurs sont en contact avec les gestionnaires du programme – scientifiques et associations naturalistes – pour demander de l'aide si besoin, transmettre des retours d'expérience et parfois participer à l'amélioration du programme ; de leur côté, les gestionnaires s'engagent à fournir aux volontaires des retours réguliers sur l'avancée de la recherche liée au programme et les premiers résultats obtenus (**Figure 1**). À terme, les résultats issus des données des programmes de sciences participatives sont diffusés *via* des **publications scientifiques** dans des revues à comité de lecture, mais aussi des journaux de vulgarisation scientifique et tout type de média grand public.

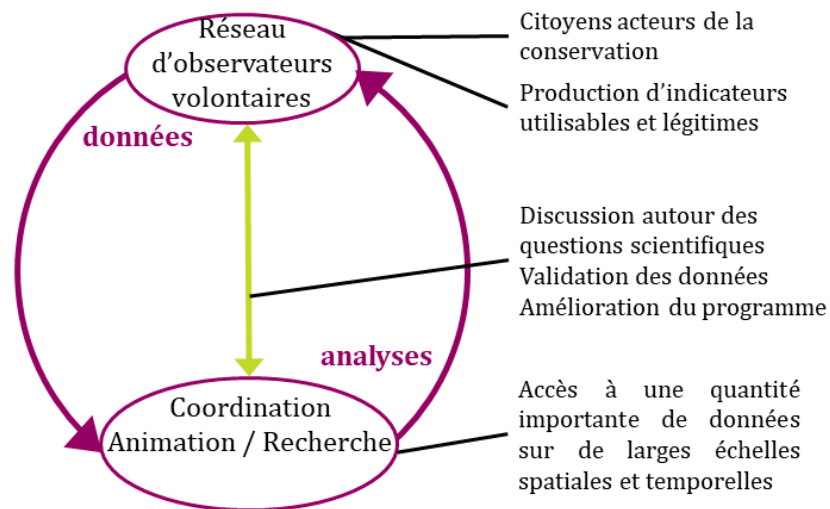


Figure 1. Fonctionnement d'un programme de sciences participatives et nature des interactions entre les observateurs volontaires et les organisateurs du programme.

Ces programmes permettent de récolter un grand nombre de données, sur de larges échelles de temps et d'espace – ce qui est particulièrement utile dans le contexte actuel où l'urgence de la crise de la biodiversité nécessite une acquisition rapide d'informations, et où les moyens alloués à la recherche et aux suivis professionnels de biodiversité sont insuffisants (Fontaine *et al.* 2012a). Dans les domaines de l'écologie et de la biologie de la

conservation, faire appel à l'aide des citoyens est une pratique devenue courante (Dickinson *et al.* 2012).

Le terme "sciences participatives" est désormais **bien repris dans les médias**, quel que soit leur support de diffusion et le type d'audience qu'ils visent. Sur leurs sites web par exemple, on peut trouver à ce sujet des articles récents de journaux généralistes tels que *Le Monde*⁴, *La Croix*^{5,6} et *L'Express*⁷, qui s'ajoutent aux articles de magazines et blogs spécialisés dans l'environnement (*Environnement Magazine*⁸, *ConsoGlobe*⁹...). À la radio et à la télévision (France Inter, France 5, Arte, M6...), des émissions et documentaires sont consacrés à ce nouveau mode d'acquisition de données scientifiques qui fait la part belle au travail collectif et à l'implication du grand public. En France, environ **deux tiers des citoyens trouvent la démarche intéressante et sont prêts à s'impliquer** dans des programmes qui feraient appel à eux pour la collecte de données scientifiques¹⁰. Les sciences participatives se font également plus présentes au sein des systèmes **politiques** et des actions en faveur de l'environnement ; en témoigne par exemple un récent rapport synthétique sur les sciences participatives en France (Houllier & Merilhou-Goudard 2016) qui avait été remis à la Ministre de l'Éducation nationale, de l'Enseignement supérieur et de la Recherche.

⁴ *Sciences participatives : les Français prêts à participer à la recherche*. 23 mai 2016. Site du journal *Le Monde*, consulté le 09/10/18. Disponible à l'adresse : www.lemonde.fr/sciences/article/2016/05/23/les-francais-prets-a-participer-a-la-recherche_4924875_1650684.html

⁵ *Quelles sont les vertus des sciences participatives ?* 8 septembre 2014. Site du journal *La Croix*, consulté le 09/10/18. Disponible à l'adresse : www.la-croix.com/Ethique/Sciences-Ethique/Sciences/Quelles-sont-les-vertus-des-sciences-participatives-2014-09-08-1202809

⁶ *Quand la science s'ouvre au grand public*. 4 octobre 2016. Site du journal *La Croix*, consulté le 09/10/18. Disponible à l'adresse : www.la-croix.com/Sciences-et-ethique/Sciences-et-ethique/Quand-la-science-s-ouvre-au-grand-public-2016-10-04-1200793625

⁷ *Les scientifiques du dimanche font progresser la recherche*. 17 juillet 2015. Site du journal *L'Express*, consulté le 09/10/18. Disponible à l'adresse : www.lexpress.fr/actualite/sciences/les-scientifiques-du-dimanche-ont-progresser-la-recherche_1697973.html

⁸ *Sciences participatives : le nouveau terrain de jeu des applications mobiles pour les naturalistes*. 25 mai 2016. Site de *Environnement Magazine*, consulté le 07/10/16. Disponible à l'adresse : www.environnement-magazine.fr/article/47502-sciences-participatives-le-nouveau-terrain-de-jeu-des-applications-mobiles/

⁹ *Sciences participatives : 4 programmes pour l'observation de la nature*. 12 avril 2016. Site *ConsoGlobe*, consulté le 09/10/18. Disponible à l'adresse :

www.consoglobe.com/science-participative-un-programme-complet-pour-curieux-de-toutes-natures-cg

¹⁰ Enquête Ipsos Sopra-Steria pour *La Recherche* et *Le Monde* (2016) par Etienne Mercier et Vincent Dusseaux – "Les Français et les sciences participatives"

Depuis une trentaine d'années, surtout en Europe et en Amérique du Nord, le nombre de programmes de sciences participatives est en **forte augmentation** (Follett & Strezov 2015, Kullenberg & Kasperowski 2016). Il existe aujourd'hui de tels programmes dans tous les grands domaines scientifiques : astronomie, climatologie, sociologie, éthologie, linguistique, documentation, histoire, archéologie, biologie moléculaire, médecine, écologie, biodiversité...

Les modes d'acquisition des données sont multiples et varient selon le programme et les méthodologies couramment employées dans les différentes disciplines. Il peut s'agir d'un questionnaire soumis régulièrement aux participants, comme dans le programme Grippenet¹¹ qui suit depuis 2012 l'état de santé des Français pendant la saison grippale. Les observateurs peuvent également être invités à effectuer des inventaires d'objets sur le terrain, comme c'est le cas pour le programme Vigie-Flore¹², ou à prendre des photos d'animaux, comme dans le Spipoll¹³ qui s'intéresse aux insectes pollinisateurs. Certaines approches relèvent de la documentation, à l'instar des Herbonautes¹⁴, programme en ligne qui invite les participants à déchiffrer des étiquettes associées à des spécimens d'herbier et à rentrer ces informations dans une base de données numérique.

Chaque programme de sciences participatives s'adresse à un **type de public plus ou moins spécialisé et restreint** : grand public pour des programmes tels que Grippenet et le Spipoll, naturalistes spécialistes de certains groupes biologiques pour Vigie-Flore, professionnels et gestionnaires de terrain pour des programmes comme le Suivi Temporel des Libellules (Steli¹⁵), public scolaire et enseignants pour les programmes de la plateforme Vigie-Nature Ecole¹⁶... Des listes de programmes existants sont mises en place pour aider le public à choisir dans quel(s) domaine(s) il souhaite s'impliquer dans sa région, comme par

¹¹ grippenet.fr

¹² www.vigienature.fr/fr/vigie-flore

¹³ spipoll.org

¹⁴ lesherbonautes.mnhn.fr

¹⁵ steli.mnhn.fr

¹⁶ www.vigienature-ecole.fr

exemple sur NatureFrance, le site d'information sur les données de la biodiversité¹⁷, ou encore OPEN, un portail des sciences participatives récemment créé par le MNHN¹⁸.

Un grand nombre de programmes de sciences participatives concerne les sciences naturelles, par exemple pour des suivis de biodiversité : oiseaux, mollusques, insectes, mammifères, plantes à fleurs, algues... On peut citer en France, le dispositif Vigie-Nature¹⁹ porté par le MNHN qui regroupe une vingtaine de programmes en lien avec la biodiversité, s'intéressant à de multiples groupes d'animaux et de plantes vivant dans des environnements divers (terrestres, d'eau douce, littoraux...) (Legrand 2013, Julliard 2017).

La pertinence des protocoles dans les programmes de sciences participatives dépend des **coûts et bénéfices associés au travail d'observateurs bénévoles** (une main d'œuvre quasiment gratuite mais le plus souvent non spécialiste) par rapport à l'emploi de scientifiques professionnels (à l'expertise fiable mais très coûteuse) (Couvét *et al.* 2008, Dickinson *et al.* 2010). Les programmes de sciences participatives peuvent fournir de **grandes quantités de données sur un temps long**, mais doivent pour cela mettre en place des stratégies pour maintenir l'intérêt des volontaires, les fidéliser, et ainsi obtenir des données sur le long terme (Freitag & Pfeffer 2013).

Les sciences participatives représentent pour les citoyens un moyen ludique et éducatif de participer activement à des programmes scientifiques. Les programmes intègrent des **aspects pédagogiques et de sensibilisation à l'environnement**, et combinent donc les observations en sciences naturelles et une éducation scientifique du public (Dickinson *et al.* 2012). Au cours des différentes étapes requises dans chaque programme de sciences participatives, les créateurs et animateurs du programme prennent soin d'inclure des indications pour une participation efficace, des précisions scientifiques, des détails sur le domaine concerné par le programme et/ou sur les objets d'étude (groupes d'êtres vivants par exemple). Souvent, l'un des partenaires du programme est un établissement d'enseignement ou une structure active dans l'animation et la médiation scientifique. Il est

¹⁷ www.naturefrance.fr/sciences-participatives

¹⁸ www.open-sciences-participatives.org

¹⁹ vigienature.mnhn.fr

également fourni aux participants un moyen de contacter les animateurs des programmes pour demander de l'aide ou poser des questions (courriel, forum...). L'aspect pédagogique peut prendre différentes formes : courriels réguliers donnant des informations complémentaires en rapport avec le programme, questionnaires envoyés aux participants afin d'enquêter sur leur expérience de volontaire, les difficultés éventuellement rencontrées, les suggestions qu'ils auraient à faire pour améliorer le programme, leur implication à long terme, etc. D'autres comme le Spipoll proposent régulièrement des "rencontres", au cours desquelles animateurs du programme et volontaires se retrouvent pour une session participative. Spipoll est également un bon exemple de programme promouvant l'entraide entre participants, puisqu'il permet à chacun de poster des commentaires sur les photos mises en ligne sur le site web, et qu'il met également à disposition un forum pour permettre à tous les acteurs de communiquer.

Afin de maximiser la fiabilité et l'exactitude de leurs données et résultats, les programmes de sciences participatives reposant sur des objets naturels nécessitent, comme pour tout travail en histoire naturelle, de pouvoir **correctement identifier** et classer ces objets. Dans ce cadre, fournir aux participants des outils pratiques et fiables pour identifier les êtres vivants, tels que nous l'avons vu précédemment, est crucial.

1.2. Le cas particulier de l'entomofaune

1.2.1. Les insectes : des organismes clés très diversifiés mais en fort déclin

Les insectes constituent la plus grande classe de métazoaires (animaux) : environ un million d'espèces ont été décrites ; presque trois animaux sur quatre sont des insectes (Figure 2).

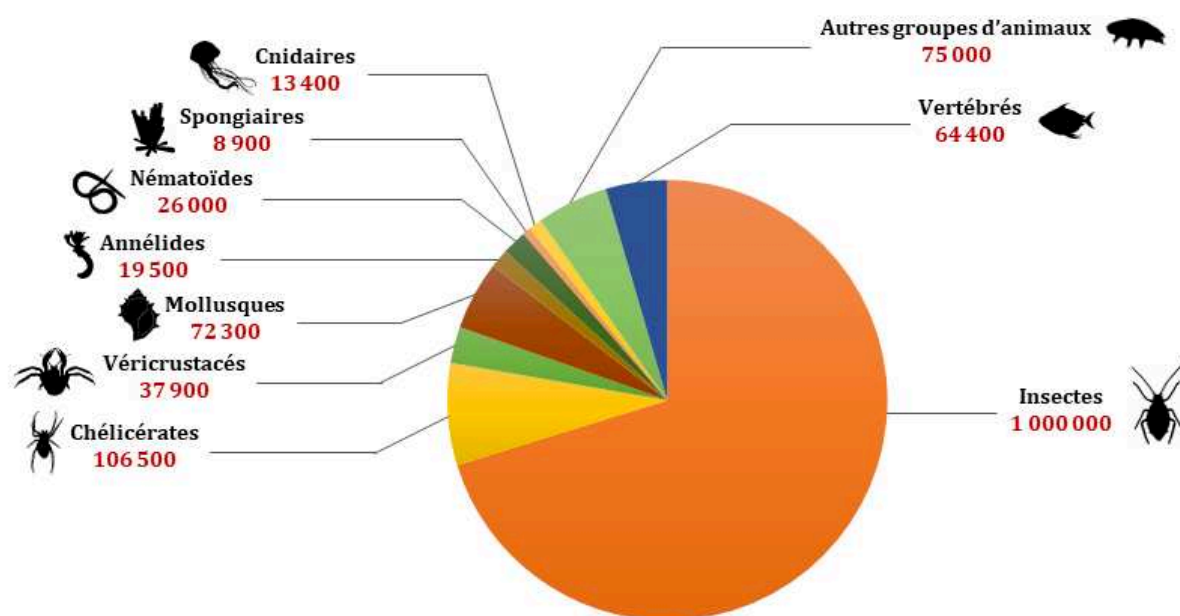


Figure 2. Principaux groupes d'animaux et nombre approximatif d'espèces décrites dans le monde pour chaque groupe (en rouge). D'après les données de l'OPIE et Lecointre & Le Guyader 2017. Silhouettes : phylopic.org et Malcolm T. Sanders.

En France, environ 35 200 espèces d'insectes sont décrites. Près de la moitié appartient aux ordres des coléoptères (scarabées, coccinelles...) et des hyménoptères (abeilles, bourdons, guêpes, fourmis...) – respectivement 9600 et 8000 espèces. Les trois

autres grands ordres sont les diptères (mouches et moustiques), les lépidoptères (papillons) et les hémiptères (punaises, pucerons, cochenilles, cigales...) ²⁰ (Figure 3).

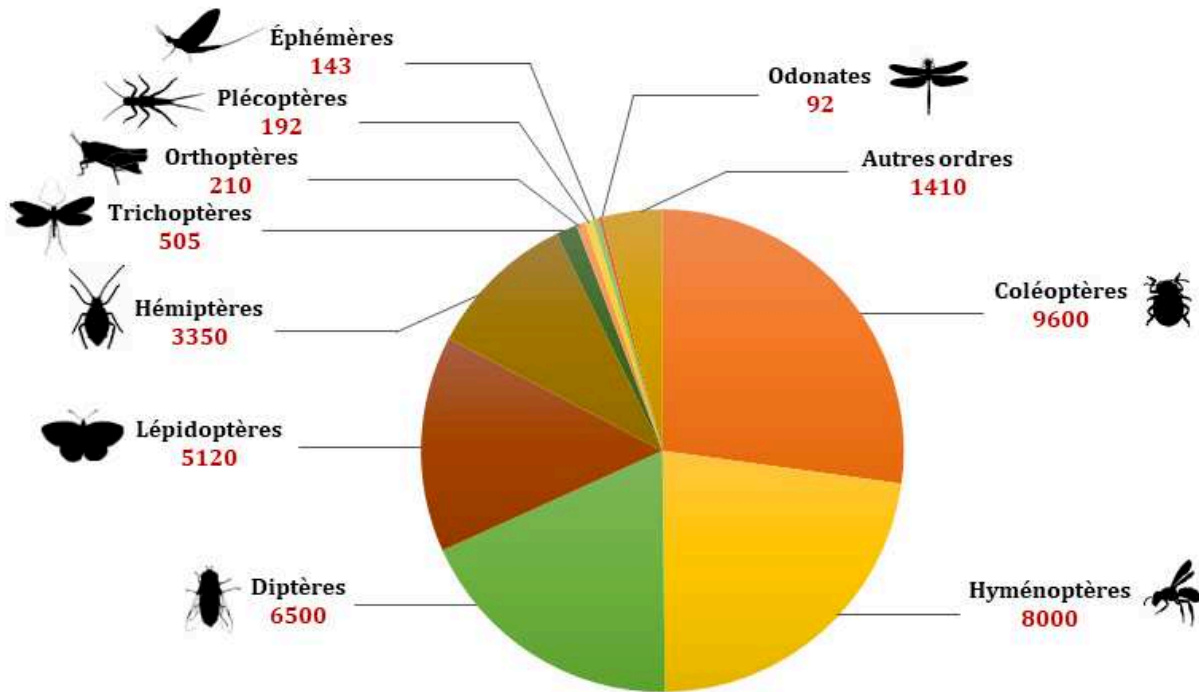


Figure 3. Principaux ordres d'insectes et nombre approximatif d'espèces décrites en France métropolitaine pour chaque ordre (en rouge). D'après les données de l'OPIE, OPIE-Benthos, programme INVFMF et SFO. Silhouettes : phylopic.org.

Les insectes sont présents dans tous les écosystèmes de la planète, à tous les types de climat, et toutes les latitudes. Ils rendent des **services écosystémiques** très variés pour les sociétés humaines, parmi lesquels la pollinisation, l'équilibre d'écosystèmes complexes (Losey & Vaughan 2006, Power 2010), la régulation des ravageurs de culture (Malausa *et al.* 2016)... Les insectes constituent la **source de nourriture** principale de nombreux organismes, dont beaucoup d'espèces d'oiseaux (Holland *et al.* 2006). Ils sont consommés traditionnellement par l'homme dans le monde entier, et représentent une source nutritive particulièrement intéressante pour le futur des sociétés humaines. De nombreuses recherches sont en cours en rapport avec la consommation de cette source riche en protéines,

²⁰ www.insectes.org/opie/nombre-insectes-monde.html

et son acceptation par les sociétés occidentales comme source valable de nourriture (Bukkens 1997, Tan *et al.* 2015, Zielińska *et al.* 2019).

Les insectes sont des organismes vulnérables et comme tant d'autres groupes d'êtres vivants, **leurs populations sont fortement menacées et impactées par les activités humaines**. En Europe, sur les 27 dernières années, plus de 75% des individus auraient disparu (Hebert *et al.* 2004, Hallmann *et al.* 2017). Les principales causes de ce déclin sont d'origine anthropique : destruction des habitats, modification de l'usage des terres et développement de l'urbanisation, surexploitation des populations, pratiques agricoles intensives, monoculture, pollution de l'air et des sols, pollution lumineuse, utilisation massive de pesticides, changements climatiques globaux... (Pyle *et al.* 1981, Fox 2013)

Face à ce déclin rapide et massif, il est urgent de pouvoir **identifier correctement les taxons d'insectes afin de mieux les connaître et de les protéger plus efficacement**. Les outils, ressources et méthodes présentés précédemment sont largement utilisés pour la détermination des insectes. Cependant, plusieurs particularités de ces organismes font qu'un certain nombre de groupes d'insectes restent délicats à identifier, et leur cas représente un enjeu particulièrement important en taxonomie.

1.2.2. Les difficultés rencontrées dans l'identification des insectes

Les insectes sont des organismes vivants peu connus au vu de leur diversité (un million d'espèces décrites sur potentiellement 100 millions d'espèces présentes sur Terre – Chapman 2009). En plus de leur nombre extrêmement élevé, beaucoup sont de petite taille et certains taxons se ressemblent beaucoup morphologiquement.

Les insectes ont la particularité de présenter **une phase larvaire et une phase adulte**, qui ne sont souvent pas morphologiquement identiques. Les différences morpho-anatomiques entre la larve et l'adulte sont particulièrement marquées chez les insectes hémimétaboles et holométaboles. Chez certains groupes tels que les coléoptères aquatiques, le lien entre la forme larvaire et la forme adulte n'est pas connu pour toutes les espèces

(Richoux 1982, Tachet *et al.* 2010). Parfois, les experts sont spécialistes d'une des deux phases seulement.

Le nombre de taxonomistes spécialistes d'un ou plusieurs groupe(s) d'insectes, ainsi que la quantité d'études scientifiques en entomologie, sont faibles relativement au nombre total d'espèces, en particulier si l'on compare à d'autres groupes d'animaux plus "populaires" tels que les mammifères ou les oiseaux (Troudet *et al.* 2017). En entomologie, les experts de l'identification sont souvent des amateurs (Fontaine *et al.* 2012a). Nous faisons face à un manque criant de chercheurs et autres scientifiques professionnels pour la connaissance pointue et la détermination précise des spécimens d'insectes (Fontaine *et al.* 2012a, Troudet *et al.* 2017), alors même que les besoins sont très importants pour l'étude de la biodiversité et la biologie de la conservation. Dans ce contexte, il est nécessaire de **multiplier l'effort pour le développement d'outils fiables d'identification de l'entomofaune**. C'est dans ce cadre que s'inscrit ce travail de thèse qui s'attache à explorer les possibilités et les perspectives d'un type particulier d'outils d'identification : les clés d'identification numériques.

1.3. L'entomofaune étudiée dans ce travail de thèse : des organismes d'intérêts variés

Les insectes présentent une très grande diversité taxonomique et écologique. J'ai souhaité mettre l'accent sur des groupes d'insectes variés, qui possèdent des particularités en termes d'écologie et d'enjeux de conservation, une importance patrimoniale et culturelle, ou encore qui soulèvent d'importantes difficultés pour leur identification.

- Certains ordres d'**insectes aquatiques** présents dans la région Ile-de-France : les odonates (libellules et demoiselles), les éphémères, les plécoptères et les coléoptères
- Les **coccinelles** de France métropolitaine
- Les **insectes pollinisateurs** de France métropolitaine.

1.3.1. Les insectes aquatiques de la région Ile-de-France

Les milieux aquatiques d'eau douce – incluant les zones humides et les cours d'eau – remplissent des **fonctions écologiques majeures** (Barnaud & Coïc 2011) : ils sont au cœur de la réception, du stockage et de la restitution de l'eau, ils participent à la formation des sols et à la régulation des crues, ce sont des filtres naturels pour la matière minérale et organique (Barbier 2011), une source de nourriture, des réseaux de transports, des lieux de loisirs (Costanza *et al.* 1997, Postel & Carpenter 1997)... Ces milieux accueillent une **biodiversité riche et particulière**, dont de nombreuses espèces vivant exclusivement dans ce type d'habitats. On peut citer des batraciens, urodèles, téléostéens uniques à ce milieu, arachnides d'eau douce, ainsi qu'une diversité exceptionnelle de plantes, et de micro-organismes animaux et végétaux (Dudgeon *et al.* 2006). En Europe, plus de 10% des espèces d'insectes dépendent strictement de ces environnements pendant une partie ou la totalité de leur cycle de vie (Nilsson 1996a). Certains écosystèmes comme les tourbières sont peu communs et très sensibles aux modifications environnementales. En France, les zones humides couvrent 6% du territoire.

L'intégrité de ces écosystèmes est gravement menacée par les activités humaines depuis le début de l'ère industrielle. Parmi elles, l'intensification de l'agriculture et de l'aquaculture, les aménagements à outrance des cours d'eau (digues, barrages, bétonnage des berges...), le développement de l'urbanisation et des infrastructures

(Trombulak & Frissell 2000), la pollution industrielle massive, et le dérèglement climatique global (Meyer *et al.* 1999). Ces nuisances sont encore plus importantes dans les régions fortement sujettes aux activités humaines (Dudgeon *et al.* 2006) telles que la région parisienne, où l'urbanisation et l'agriculture sont particulièrement intensives. Des projets tels que le "Grand Paris" nécessitent la création d'un grand nombre de nouvelles infrastructures qui perturbent des écosystèmes déjà fragilisés²¹. En Ile-de-France, on estime qu'environ 90% des mares ont disparu depuis 1900 (données "Inventaire des mares" SNPN)²².

Depuis la fin du XX^{ème} siècle, la prise de conscience du déclin de ces habitats n'a cessé de croître. Dans ce cadre, des actions ont été mises en place dans l'objectif de les préserver. La conservation des milieux aquatiques est encadrée par des textes de loi, dont la **Directive Cadre sur l'Eau**²³ au niveau européen, et la Loi sur l'eau et les milieux aquatiques (LEMA)²⁴ en France. Le Grenelle de l'Environnement²⁵ a développé à partir de 2007 le projet "Trame bleue" pour favoriser les corridors écologiques aquatiques. La Société nationale de protection de la nature (SNPN) anime depuis 2010 un programme participatif appelé "Inventaire des mares d'Ile-de-France", pour l'amélioration des connaissances et une meilleure prise en compte de ces milieux dans les politiques d'aménagement²⁶. Des programmes s'attachent particulièrement à la connaissance et la protection des insectes aquatiques en France métropolitaine ; parmi eux : les Inventaires Éphémères (INVFMR), Plécoptères (INVP), Hétéroptères (INVHA), Mégaloptères et Névroptères (INVMN), Coléoptères Aquatiques (INVCA) et Trichoptères (INVT) de l'OPIE-Benthos²⁷, les programmes "Inventaire cartographique des odonates de France" (Invod) et "Complément à

²¹ *Soutenabilité du Grand Paris - Note sur les enjeux dans le domaine de l'eau* (2012)

²² www.essonne.fr/fileadmin/patrimoine_naturel/essonne/BilanMareIDF_2011-2012-1.pdf

²³ Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy

²⁴ Loi n° 2006-1772 du 30 décembre 2006 sur l'eau et les milieux aquatiques

²⁵ www.vie-publique.fr/actualite/dossier/grenelle-an/grenelle-environnement-an-apres.html

²⁶ www.snpn.mares-idf.fr/

²⁷ www.opie-benthos.fr

l'inventaire des libellules de France" (Cilif)²⁸ portés par la Société Française d'Odonatologie (SFO)²⁹, et les "listes rouges" de l'UICN pour plusieurs ordres d'insectes aquatiques.

Pour que les mesures de protection des milieux aquatiques soient les plus efficaces possible, il est nécessaire d'avoir une connaissance avancée des organismes vivant dans ces habitats, en particulier parce que les inventaires de biodiversité développés dans le cadre de certains textes de loi se basent sur la présence ou l'absence de certains taxons à un endroit donné. L'IBG-DCE, "l'Indice biologique global" pour la Directive européenne Cadre sur l'Eau (adapté de l'Indice biologique global normalisé – IBGN), repose sur les occurrences de "macro-invertébrés" vivant dans les eaux courantes, qui servent d'indicateurs de la qualité des cours d'eau³⁰. Depuis 2016 l'indice I2M2 (Indice invertébrés multimétrique), plus complet et affiné par rapport à l'IBG-DCE, est utilisé, mais il repose toujours largement sur les occurrences de "macro-invertébrés aquatiques" (Mondry *et al.* 2012).

Cependant les insectes aquatiques subissent ce qu'on appelle le "*taxonomic impediment*" (Dubois 2010). Du fait de leur petite taille et des difficultés associées à leur identification sur des critères morpho-anatomiques (Williams *et al.* 2006), une expertise taxonomique est souvent requise pour les identifier correctement, et globalement nous manquons d'experts des insectes aquatiques par rapport au travail nécessaire. De plus, on observe un biais taxonomique dans les données disponibles pour ces groupes. En effet de façon générale, un relatif petit nombre d'espèces attire la majorité de l'attention du grand public et des milieux scientifiques, alors que la plupart des taxons – dont les insectes – sont relativement peu étudiés et connus, comparé à leur grande diversité spécifique (Troudet *et al.* 2017). De ce fait, les bases de données pour les insectes aquatiques sont très incomplètes, que ce soit concernant les paramètres écologiques (distribution, traits d'histoire de vie...) ou l'identification morpho-anatomique fine – avec des variations selon les ordres (Vieira *et al.* 2006, Jackson *et al.* 2014).

²⁸ www.libellules.org/protocole/cilif.html

²⁹ www.libellules.org

³⁰ www.forum-zones-humides.org/iso_album/indice_biologique_global.pdf

Les espèces inféodées au milieu aquatique sont moins connues que les autres groupes d'insectes. L'un des obstacles à leur étude est l'accès relativement difficile aux habitats aquatiques – la récolte des insectes nécessite un matériel plus spécialisé et encombrant. D'autre part, les insectes aquatiques sont souvent de petite taille, en particulier les plus jeunes stades larvaires. Ils sont très **fragiles** et nécessitent du matériel d'observation adapté, notamment des loupes binoculaires voire des microscopes.

Dans ce contexte – comme pour la majorité des études écologiques (gestion des écosystèmes, établissement et amélioration des politiques publiques environnementales...) – une détermination juste et fiable des espèces est essentielle (Vander Zanden *et al.* 2010, Jackson *et al.* 2014, Leal *et al.* 2016). Divers types de publics – naturalistes, amateurs et professionnels – doivent avoir la possibilité d'identifier correctement les organismes qu'ils étudient (Farnsworth *et al.* 2013).

La première nécessité pour les personnes travaillant sur la biodiversité est la possibilité d'accéder à des guides d'identification. Cependant, les guides facilement utilisables par les non-taxonomistes et le grand public sont rares, et disponibles pour relativement peu de groupes taxonomiques (SCBD 2010). La plupart des nouvelles espèces d'insectes aquatiques sont décrites par des taxonomistes non-professionnels (Fontaine *et al.* 2012b) et la distribution, la biologie et l'écologie d'une majorité d'organismes restent peu connus (Costello *et al.* 2006). Pour les ordres des odonates, éphémères, plécoptères et coléoptères aquatiques, des programmes de sciences participatives ont été créés dans le but de réduire les lacunes, notamment aux USA et au Royaume-Uni – par exemple le Inletkeeper program³¹, le Dragonfly Pond Watch Project³² et le Anglers' Riverfly Monitoring Initiative³³. En France, quelques programmes de sciences participatives s'intéressent aux insectes aquatiques ; c'est le cas du Le Suivi temporel des libellules (Steli)³⁴ qui fait partie de la plateforme Vigie-Nature.

³¹ inletkeeper.org

³² xerces.org/dragonfly-migration/pondwatch/

³³ www.riverflies.org/rp-riverfly-monitoring-initiative

³⁴ steli.mnhn.fr

Dans cette perspective et pour les futurs programmes qui seront créés sur ces organismes, **les outils d'identification interactifs, visuels et faciles à utiliser, offrent d'énormes potentiels pour le grand public et les non-spécialistes** (Riccamboni *et al.* 2010).

La plupart des clés d'identification pour les insectes aquatiques sont difficiles à utiliser par les non-spécialistes, à cause de la terminologie spécifique à l'entomologie et la petite taille des individus, par exemple : Tachet *et al.* 2010 (pour tous les "invertébrés" aquatiques), Nilsson 1996b, Dijkstra & Lewington 2006 et Grand *et al.* 2014 (pour les odonates), Hynes 1977, Nilsson 1996a et Pryce *et al.* 2007 (pour les plécoptères), Nilsson 1996a, Macadam & Bennett 2010, Bauernfeind & Soldán 2012 (pour les éphémères).

Les outils modernes qui ont accompagné le développement des technologies numériques rendent l'identification plus intuitive et facile pour l'utilisateur. Plusieurs clés d'identification interactives sont disponibles en ligne pour les "invertébrés" aquatiques, par exemple celle du Groupe d'éducation et d'écosurveillance de l'eau du Québec³⁵ et celle de la base de données Perla pour la France métropolitaine³⁶. Ces clés peuvent être abondamment illustrées à l'aide d'images et de dessins, et inclure des ressources additionnelles telles que des liens web et des fichiers sonores. Cependant, la plupart de ces outils ne sont que des versions numérisées de clés à accès simple et maintiennent les mêmes restrictions d'utilisation, en particulier le fait de devoir suivre une série d'étapes dans un ordre imposé (voir la partie 2.1 pour plus de détails sur les clés d'identification à accès simple).

Il n'existe pas pour le moment de clé d'identification interactive numérique à accès multiple pour les insectes aquatiques de la région Ile-de-France, et ce malgré un net besoin pour ce type d'outils qui peuvent également être utilisés directement sur le terrain pour les inventaires et suivis de biodiversité, qui sont cruciaux dans les zones fortement dégradées.

³⁵ www.g3e-ewag.ca/documents/jeux/cle_macro/cle.html

³⁶ www.perla.developpement-durable.gouv.fr/index.php/arbre

L'un des objectifs majeurs de ce travail de thèse était donc de construire des bases de données et des clés d'identification numériques permettant d'améliorer les connaissances sur les insectes aquatiques de cette région, que ce soit au niveau taxonomique (facilitation des identifications morpho-anatomiques), pour pallier quelques lacunes dans les bases de données incomplètes (au niveau de l'écologie, du mode de vie...), pour sensibiliser les Franciliens à la richesse de la biodiversité et des milieux qui les entourent, ou encore pour stimuler les efforts de conservation des insectes aquatiques de la région.

Les quatre ordres d'insectes aquatiques traités dans cette thèse sont les suivants :

- Les **odonates**, qui sont au nombre de 92 espèces en France métropolitaine, dont 64% en Ile-de-France (59 espèces)
- Les **éphémères** – 143 espèces en France métropolitaine dont 27% en Ile-de-France (39 espèces)
- Les **plécoptères** – 192 espèces en France métropolitaine dont 5% en Ile-de-France (10 espèces)
- Les **coléoptères aquatiques** – environ 650 espèces en France métropolitaine dont 38% en Ile-de-France (245 espèces).

Les clés d'identification numériques ont été construites pour les larves (nymphe), les imagos (adultes), ou les deux, selon les ordres (**Figure 4**).

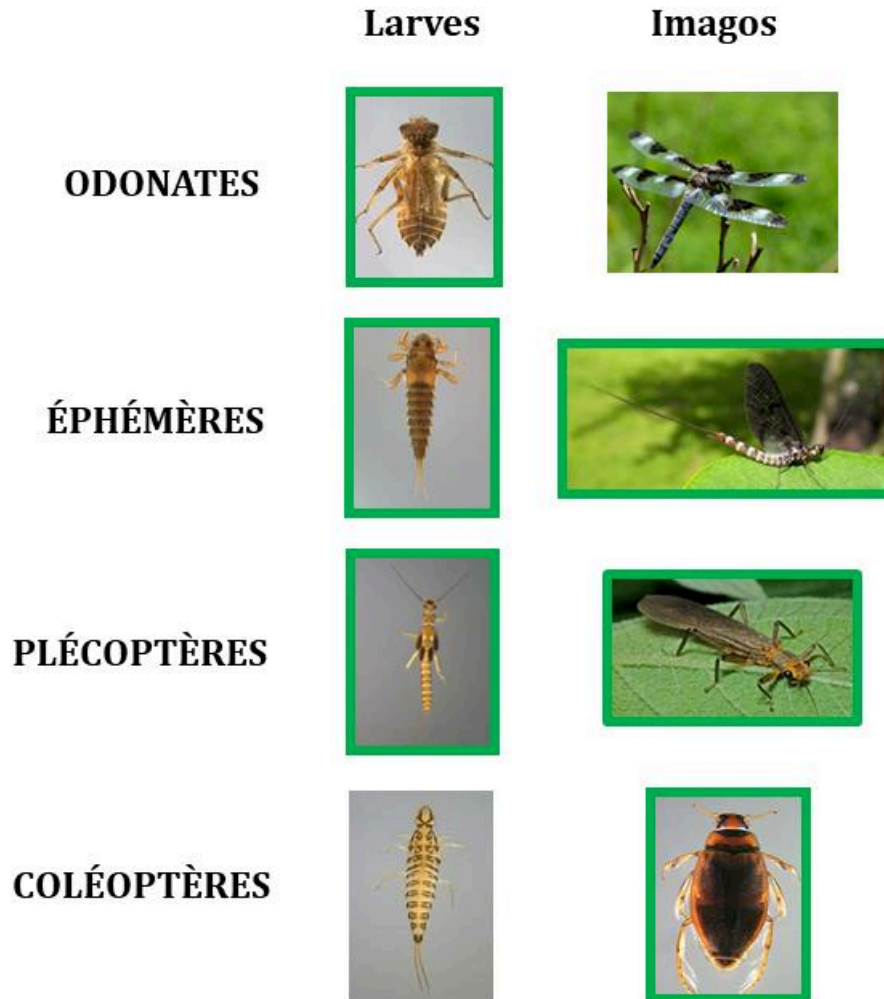


Figure 4. Représentation type d'une forme larvaire et d'une forme imago (adulte) pour les quatre ordres d'insectes aquatiques ayant fait l'objet d'une clé d'identification au cours de la thèse. Les stades traités dans les clés d'identification pour l'Ile-de-France sont encadrés en vert. Photographies : Lilian Bonnafoux pour Perla, Nigel Jones, Jo Naylor, Mick Talbot.

Le choix de l'échantillonnage taxonomique pour les bases de données et les clés d'identification est détaillé dans la partie 3.1.1.

1.3.2. Les coccinelles de France métropolitaine

Les coccinelles (Coccinellidae) sont une famille de coléoptères populaires et appréciés des naturalistes comme du grand public. Leur rôle écologique et économique en tant que prédateurs d'insectes ravageurs de cultures (pucerons, cochenilles...) fait que leur identification peut être utile pour les naturalistes, les amateurs et les professionnels

(Hemptinne *et al.* 2005, Hodek & Honěk 2009, Ali *et al.* 2014). Plusieurs programmes de sciences participatives ont été développés dans le but de décrire les patrons de distribution des populations de coccinelles ; nous pouvons mentionner le Harlequin Ladybird Survey³⁷ et le Ladybird Survey³⁸ au Royaume-Uni, le Lost Ladybug Project³⁹ et le Buckeye Lady Beetle Blitz⁴⁰ aux USA, et le Coccinula Recording Scheme en Belgique (Bagnée *et al.* 2011).

Les populations indigènes de coccinelles sont en forte régression depuis plusieurs décennies, non seulement à cause de l'impact direct des activités humaines (pertes d'habitats, utilisation massive de pesticides...), mais aussi depuis l'introduction (à partir des années 1980 en Europe) d'une coccinelle invasive venue d'Asie, *Harmonia axyridis*, très efficace comme régulatrice des insectes ravageurs de culture (Majerus *et al.* 2006, Comont *et al.* 2014). Cependant cette espèce de grande taille, prédatrice vorace qui se reproduit rapidement, supprime les populations natives d'autres espèces de coccinelles, par prédation et compétition pour les ressources (Roy *et al.* 2012). Des outils d'identification fiables pour les espèces de coccinelles permettent entre autres, de pouvoir suivre et étudier finement l'évolution des populations.

La majorité des clés d'identification des coccinelles d'Europe et d'Amérique du Nord sont des clés à accès simple et souvent difficiles à utiliser par les non-spécialistes (Dauguet 1949, Iablokoff-Khnzorian 1982, Gordon 1985, Chapin & Brou 1991). D'autres clés se basent principalement sur les formes et les colorations, mais la plupart des caractères nécessitent un vocabulaire spécifique, ce qui rend ces clés trop compliquées pour le grand public : Bagnée & Branquart (2000) pour la Belgique, Le Monnier & Livory (2003) pour l'ouest de la France, Roy *et al.* (2013) pour les îles britanniques.

L'un des principaux avantages que présentent les coccinelles, en particulier par rapport à beaucoup d'autres groupes d'insectes, est que leur identification peut dans beaucoup de cas être réalisée uniquement à partir de caractères morpho-anatomiques

³⁷ www.harlequin-survey.org

³⁸ www.ladybird-survey.org

³⁹ www.lostladybug.org

⁴⁰ entomology.osu.edu/about-us/multimedia/buckeye-lady-beetle-blitz

aisément visibles, et de patrons de coloration sur les élytres et le pronotum – du moins pour les espèces de grande taille.

Au sein des Coccinellidae, les variations de couleurs inter- et intra-spécifiques sont nombreuses (**Figure 5**). Nous avons fait le choix d'intégrer ces types de caractères dans la clé d'identification numérique des coccinelles présentée dans ce document, comme détaillé aux chapitres 3 et 4 de ce document, et dans l'article publié dans la revue *ZooKeys* (**Annexe 1**).



Chilocorus renipustulatus



Hippodamia variegata



Tytthaspis sedecimpunctata



Coccinula quatuordecimpustulata



Propylea quatuordecimpunctata



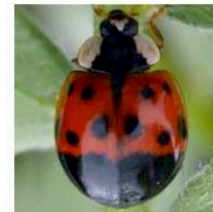
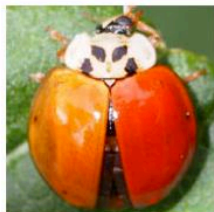
Anatis ocellata



Halyzia sedecimguttata



Myzia oblongoguttata

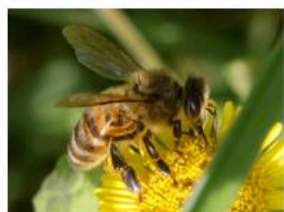


Harmonia axyridis

Figure 5. Exemples de variabilité morphologique intra- et inter-spécifiques chez les Coccinellidae. Photographies : Christoph Benisch, Kirsten Eta, Robert Keul, René Köhler, C. Schmitt-Riegraf.

1.3.3. Les insectes pollinisateurs de France métropolitaine

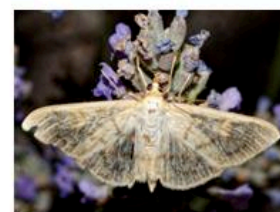
La majeure partie du service de pollinisation (en termes de productivité) est assuré par des insectes, en particulier des hyménoptères (abeille mellifère domestique, abeilles sauvages, bourdons...) et des diptères (syrphes, mouches, mites...) (Thapa 2006, Rader *et al.* 2016). Mais les coléoptères, lépidoptères, hémiptères, et d'autres ordres dans une moindre mesure, participent également aux systèmes de pollinisation (**Figure 6**).



Abeille mellifère (*Apis mellifera*)



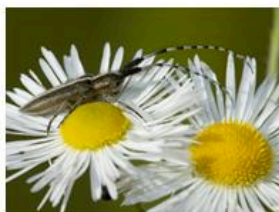
Pentatome rayé (*Graphosomalineatum*)



Pyrale du Houblon (*Pleuroptyaruralis*)



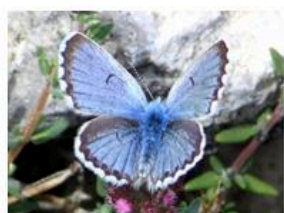
Abeille Hylaeus à taches blanches (*Hylaeus sp.*)



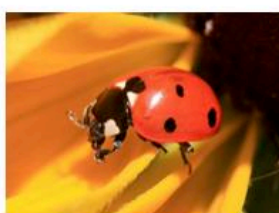
Agapanthie des Chardons (*Agapanthiacardui*)



Éristale taeniops (*Eristalinus taeniops*)



Azuré du Thym (*Pseudophilotes baton*)



Coccinelle à sept points (*Coccinella septempunctata*)



Mouche du genre *Platystoma*

Figure 6. Exemples d'insectes pollinisateurs (ordres Coleoptera, Diptera, Hemiptera, Hymenoptera et Lepidoptera). Photographies © Spipoll.

Dans le monde, au moins 75% des espèces de végétaux cultivés nécessitent l'intervention d'insectes pollinisateurs pour leur reproduction (Williams 1994, Roubik 1995, Klein *et al.* 2007). Les systèmes de pollinisation et les pollinisateurs sont nécessaires à l'équilibre des écosystèmes à grande échelle, au maintien de millions d'espèces de plantes et

à la production agricole. Leur valeur économique est estimée à plusieurs dizaines de milliards de dollars à l'échelle mondiale (Allsopp *et al.* 2008).

Or depuis quelques dizaines d'années, on assiste à un déclin des populations d'insectes pollinisateurs, déclin qui ne fait que s'accroître avec l'intensification des activités anthropiques au cours du temps. Les causes de la perte massive des pollinisateurs sont diverses, et les interactions qu'elles peuvent avoir entre elles aggravent leurs effets. Parmi ces causes nous pouvons citer : la perte et la fragmentation des habitats, l'urbanisation, l'homogénéisation des cultures (agriculture mono-spécifique), l'utilisation abusive de pesticides, les pathogènes et maladies, les espèces animales et végétales introduites et invasives, les changements climatiques globaux (Cane & Tepedino 2001, Kluser & Peduzzi 2007, Goulson *et al.* 2008, Potts *et al.* 2010, Fox 2013, Goulson *et al.* 2015, Geslin *et al.* 2017).

Actuellement, les problématiques liées à la pollinisation font l'objet d'une attention particulière de la part des décideurs politiques. En France en 2016, L'OPIE et le Ministère de l'Écologie, du Développement durable et de l'Énergie, ont lancé un Plan national d'actions (PNA) pour la période 2016-2020, appelé "France, terre de pollinisateurs"⁴¹. Début 2019, le gouvernement a annoncé la mise en place d'un groupe de travail dont le but est de mieux protéger les insectes pollinisateurs de l'utilisation des produits phytopharmaceutiques⁴². Au niveau européen, début 2018 le Parlement de l'Union Européenne a voté une résolution demandant à la Commission et aux États membres de mettre en place un plan de lutte à grande échelle contre la mortalité des abeilles⁴³.

Le grand public est également sensibilisé, en particulier via des initiatives spécifiques sur le sujet. Le programme de sciences participatives Spipoll – présenté en détails dans la partie 2.2.4 – lancé en 2010, en est un exemple. Il fait appel au grand public dans l'objectif de mieux caractériser et comprendre les réseaux de pollinisation en France métropolitaine.

⁴¹ http://www.insectes.org/opie/pdf/3993_pagesdynadocs570e1d6156925.pdf

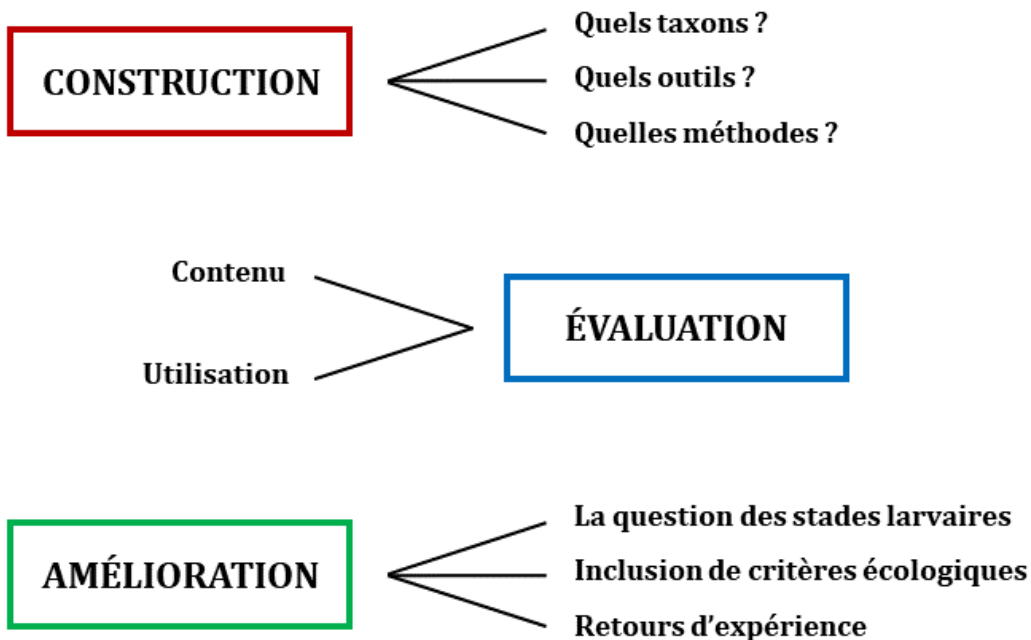
⁴² <https://www.gouvernement.fr/pour-une-meilleure-protection-des-abeilles-et-des-insectes-pollinisateurs>

⁴³ European Parliament resolution of 1 March 2018 on prospects and challenges for the EU apiculture sector (2017/2115(INI))

1.4. Objectifs et problématiques de la thèse

L'objectif général de ce travail de thèse est d'explorer des méthodes innovantes pour faciliter l'identification de l'entomofaune. Un accent particulier a été mis sur l'utilisation de la suite de logiciels Xper et d'outils leur étant associés, pour la création et l'étude de clés d'identification numériques. La démarche globale réalisée consistait en trois étapes :

- **La construction de clés d'identification numériques** pour différents groupes d'insectes de France métropolitaine
- **L'analyse et l'évaluation de clés** d'identification numériques, du point de vue de leur **contenu**, de la façon dont ces outils sont **utilisés**, et de la **perception** qu'ont les utilisateurs des clés et de l'entomofaune
- **L'amélioration de clés** d'identification numériques, pour certains groupes dont la détermination peut être problématique, grâce à l'intégration dans les clés de **critères d'identification** autres que les caractères morpho-anatomiques, et les **retours d'expérience** des utilisateurs des clés d'identification.



Cette démarche a été illustrée avec plusieurs clés numériques, concernant différents ensembles de taxons au sein des grands groupes d'insectes d'intérêt présentés dans le sous-chapitre précédent.

Ce document s'attachera à répondre aux questions suivantes :

- **Comment construire des outils pour l'identification des insectes de France, qui soient à la fois fiables, intuitifs, collaboratifs, évolutifs et accessibles au plus grand nombre ?**
- **Que nous apprend l'analyse et l'évaluation de ces clés d'identification numériques sur la façon dont elles sont utilisées, et sur la perception qu'ont les utilisateurs de ce type d'outils et de l'entomofaune de manière plus générale ?**
- **Comment améliorer de telles clés d'identification, de manière à les rendre plus fiables et précises pour la détermination des organismes, plus faciles à appréhender par les utilisateurs, et adaptables à différents types de publics, du grand public novice à l'expert entomologiste ?**

Chapitre 2 – Caractéristiques et historique des clés d'identification

Dans les sciences naturelles, une clé d'identification est un **outil dont l'objectif est de permettre et/ou faciliter la détermination d'un spécimen** ou d'une partie de spécimen. Il existe deux grands types de clés d'identification : les clés à accès simple (ou accès unique) et les clés à accès multiple (ou accès libre).

2.1 Les clés d'identification à accès simple

2.1.1 *Fonctionnement et particularités des clés d'identification à accès simple*

Une clé d'identification à accès simple peut prendre plusieurs formes, mais elle est toujours constituée d'une série d'étapes que l'utilisateur doit suivre pour aboutir à une identification de son spécimen. On la dit "à accès simple" car cette suite d'étapes est fixée pour une même clé d'identification. L'ordre des étapes successives est imposé (**Figure 7**). À chaque étape de la clé, plusieurs alternatives sont proposées à l'utilisateur, qui doit sélectionner celle qui correspond au spécimen qu'il souhaite identifier. Le choix d'une alternative détermine l'étape suivante, et ainsi de suite, jusqu'à aboutir à une identification finale. Si deux alternatives exactement sont proposées à chaque étape – par exemple, "présence d'ailes" *versus* "absence d'ailes", ou encore "pattes de couleur rouge" *versus* "pattes de couleur jaune" – la clé est dite dichotomique. Si trois alternatives ou plus sont proposées à chaque étape – par exemple, feuille "lancéolée" ou "ovale" ou "palmée" ou "pennée" ou "réniforme" – la clé est alors dite polytomique.

Les clés d'identification à accès simple telles qu'elles sont représentées dans les ouvrages ne suivent pas toujours la systématique des taxons d'intérêt ni leur classification phylogénétique au sein d'un groupe plus large d'être vivants. Elles ne renseignent pas forcément sur les relations évolutives entre organismes. La discrimination progressive entre les différents objets inclus dans une clé donnée se fait à chaque étape du chemin d'identification, selon des critères discriminants arbitraires propres aux objets décrits et à cette clé d'identification en particulier. L'auteur de la clé fait le choix des critères distinctifs de sorte à réduire le nombre d'étapes nécessaires à l'obtention d'une identification pour

l'objet étudié, et privilégie les caractères les moins ambigus possibles à chaque étape du chemin.

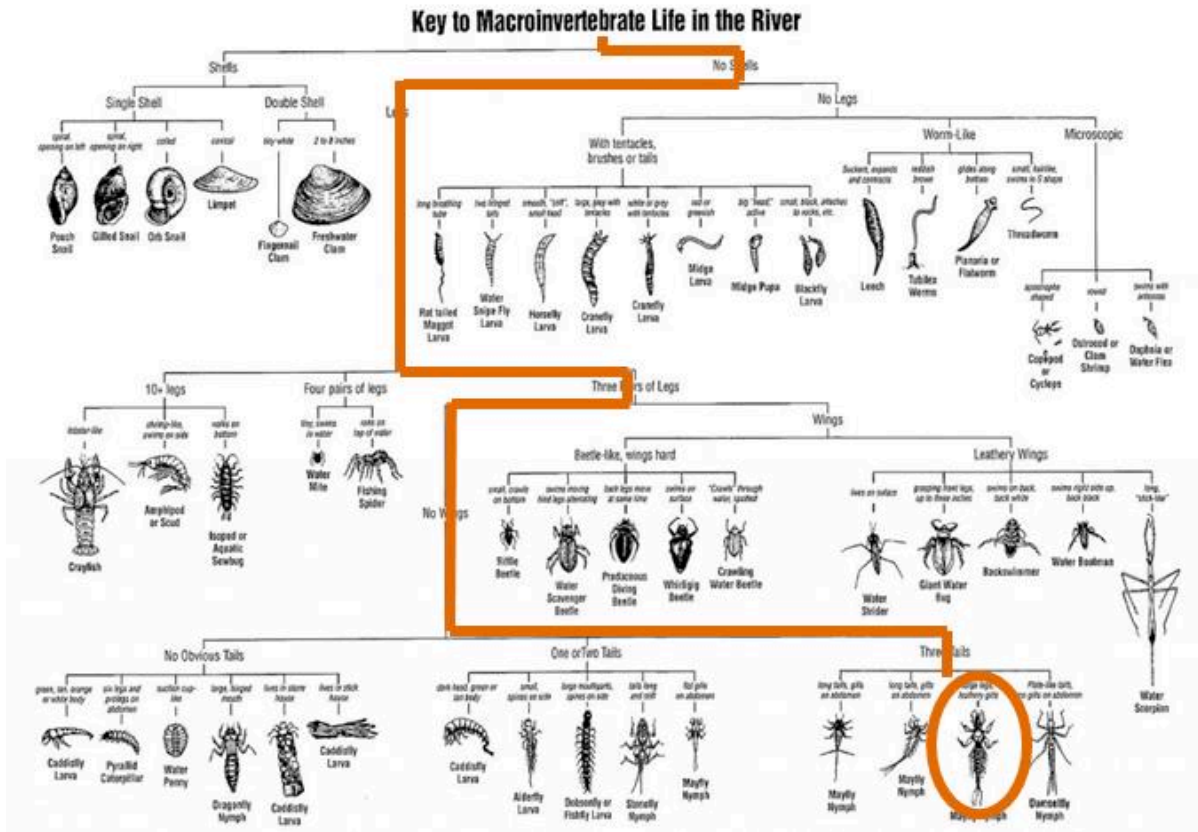


Figure 7. Clé d'identification des grands groupes de "macro-invertébrés" vivant dans les eaux douces courantes. Pour obtenir une détermination finale (tout en bas), l'utilisateur doit suivre un chemin d'identification imposé en commençant par la première étape (tout en haut). En orange est représenté le chemin que l'utilisateur doit suivre pour obtenir la détermination d'une nymphe d'éphémère. Modifié depuis une figure réalisée par l'université du Wisconsin en collaboration avec le Department of Natural Resources du Wisconsin.

Par exemple, on pourrait imaginer une clé d'identification pour les grands groupes de vertébrés, incluant une étape distinguant les "animaux ailés" et les "animaux sans ailes". Dans ce cas, les chauves-souris et les oiseaux se retrouveraient dans une branche distincte du chemin d'identification, à l'opposé des animaux sans ailes. Cette discrimination se fait sur un critère facile à visualiser et pertinent pour la distinction de groupes que l'auteur souhaite faire dans ce cas particulier. Cependant, la présence d'ailes représente ici une convergence évolutive et pas une homologie. L'ensemble "chauves-souris + oiseaux" n'est pas un clade.

Les clés d'identification qui cherchent à traduire la classification sont dites "naturelles", les autres sont nommées "clés artificielles" et se veulent essentiellement pratiques.

2.1.2 *A l'origine des clés d'identification à accès simple*

Depuis le début de leur existence, les clés de détermination peuvent traiter de toutes sortes d'éléments – y compris en-dehors des sciences naturelles – à partir du moment où ces derniers sont reconnaissables et caractérisables grâce à différents critères. Parmi les premières clés d'identification connues, celles de Petrus Ramus (1515–1572), logicien et philosophe français, étaient dichotomiques et portaient sur la dialectique (Ramus 1574) (**Figure 8a**) ou encore sur les opérations arithmétiques (Ramus 1592, Ong 1958). La volonté de Ramus était de mettre à disposition des informations de façon claire et compréhensible, afin qu'elles puissent être enseignées et soumises au débat (Griffing 2011). Dès leurs origines, les clés dichotomiques se voulaient spécifiquement une aide à l'apprentissage, basée sur la découverte et sur une simplification des idées pour les mettre à portée du plus grand nombre.

John Ray (1627–1705), naturaliste anglais membre de la Royal Society, publie en 1686 la première partie de son *Historia plantarum generalis* (Ray 1686), qui peut être considérée comme la première flore mondiale. L'ouvrage possède des clés d'identification dichotomiques (**Figure 8b**). Cependant, l'absence de schémas ou de dessins au sein de ces clés met rapidement en lumière l'importance des illustrations pour l'étude des objets de sciences naturelles. Ray inclut donc ensuite de nombreux dessins à l'aquarelle dans son ouvrage intitulé *Synopsis methodica stirpium Britannicarum* (Ray 1690). De même Richard Waller (1650–1725), un autre naturaliste anglais membre de la Royal Society, publie ses *Tables of the English herbs* (Waller 1689), une clé d'identification dichotomique richement illustrée au crayon et à l'aquarelle (**Figure 9**). Comme Ramus, la volonté de Waller est de s'adresser à tous (y compris aux "novices") grâce à la simplification de certaines classifications botaniques, et surtout à ses nombreuses illustrations, au lieu de se reposer uniquement sur du texte pour ses descriptions (Griffing 2011).

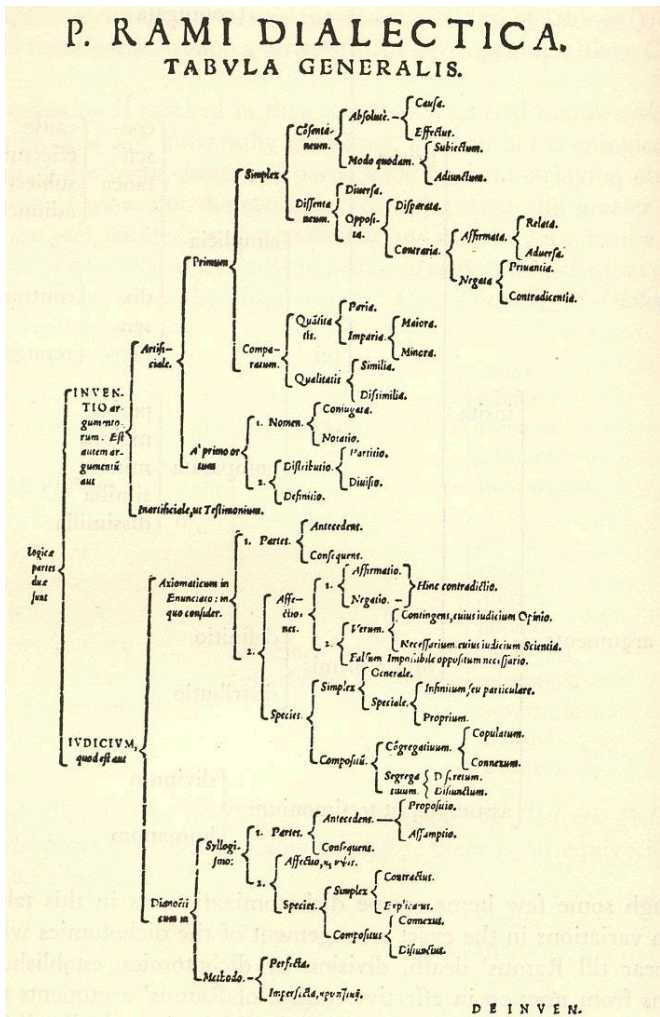
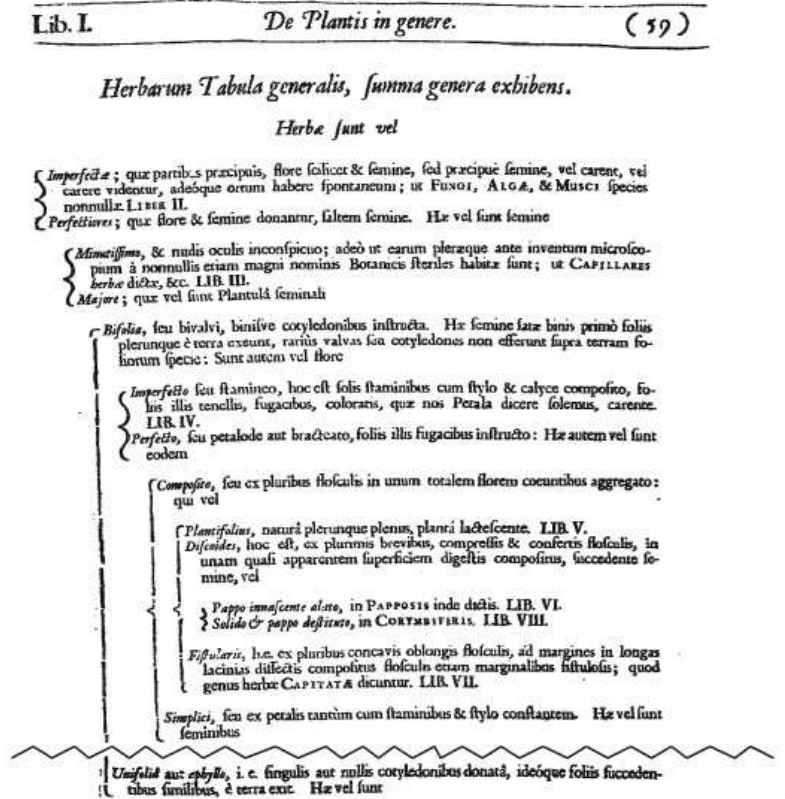


Figure 8. 8a-Extrait d'une clé d'identification

dichotomique de Petrus Ramus portant sur la dialectique (Ramus 1574). **8b**-Extrait d'une clé d'identification dichotomique présente dans *Historia plantarum generalis* (Ray 1686) ; figure extraite de Griffing 2011.



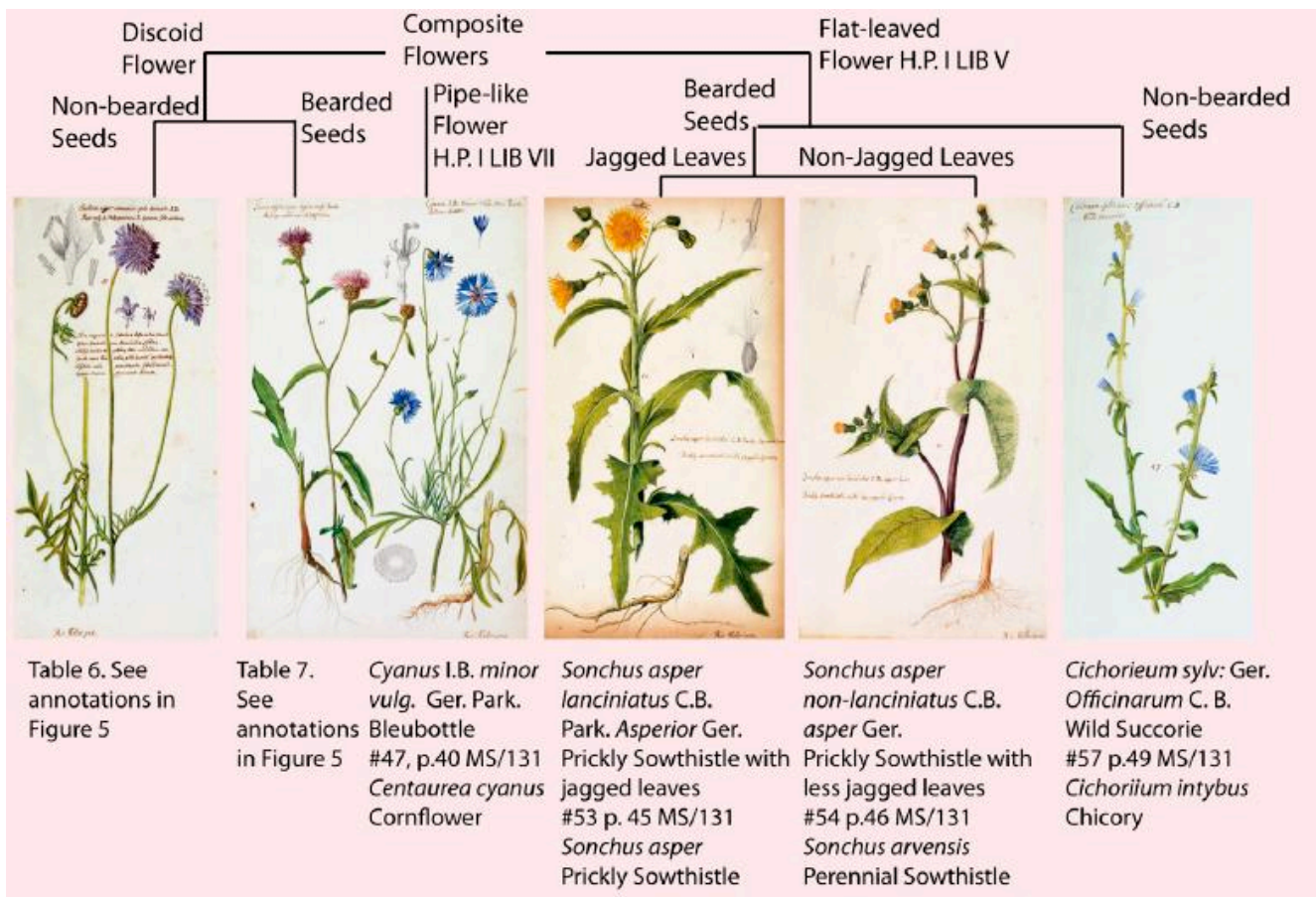


Figure 9. Extrait d'une des *Tables of the English herbs* (Waller 1689). Figure extraite de Griffing 2011.

L'ouvrage *Species plantarum* de Carl Linnæus (1707–1778) (Linnæus 1753) constitue lui, une clé d'identification botanique polytomique, forme que prenaient couramment les clés d'identification jusqu'au XIX^e siècle. Jean-Baptiste de Lamarck (1744–1829) publie en 1778 sa *Flore française*, dont la clé (non illustrée) est souvent citée comme étant la première clé d'identification dichotomique (Lamarck 1778).

2.1.3 Types de caractères utilisés dans les clés

La grande majorité des clés d'identification repose sur l'utilisation de **caractères morphologiques ou anatomiques** pour décrire les spécimens. Ces types de descripteurs sont relativement simples à coder en tant que descripteurs discrets, par exemple "présence" vs "absence" ou "pattes noires" vs "pattes jaunes" vs "pattes orange".

Cependant, d'autres types de critères peuvent être utilisés, seuls ou en complément des structures morpho-anatomiques. Si l'utilisateur de la clé a accès à ces informations, il est possible d'inclure dans les clés des **traits relatifs à l'écologie ou à l'histoire de vie** des organismes. Il peut s'agir par exemple de caractéristiques de leur milieu de vie (localisation géographique, pH et température de l'eau pour les organismes aquatiques, habitat ombragé ou non...) ou de particularités biologiques et comportementales (période de ponte ou de floraison, mode de déplacement, régime alimentaire...). La clé réalisée par Zervakis *et al.* (2014) a l'originalité de prendre en compte les associations entre des champignons et des plantes. Certaines clés numériques réalisées sous Xper³ (la plateforme que j'ai utilisée pour réaliser des clés d'identification) intègrent des critères écologiques. Parmi elles, nous pouvons noter la clé "Insectes" du Spipoll⁴⁴ (présentée en détails au 4.2.2a), une clé des orchidées de France⁴⁵ (Dubuisson *et al.* 2018), une clé des guêpes du genre *Pteromalus* (Klimmek & Baur 2018) et une clé des chouettes et hiboux de France⁴⁶.

Certaines clés d'identification sont basées sur des critères sonores, par exemple les **chants, cris ou sons** émis par des animaux. Dans certains cas, il est possible de coder des caractères simples tels que la durée, la hauteur, la fréquence ou les patrons de répétition d'un son. Cela a déjà été réalisé, par exemple pour certains oiseaux, chauve-souris et criquets (Reinhold *et al.* 2001, Warren 2003, Szövényi *et al.* 2012, Iorgu *et al.* 2017). Les clés de détermination en ligne permettent souvent aussi d'intégrer des fichiers audio, qui peuvent servir à la description et/ou à vérifier la justesse de l'identification. C'est le cas de la clé "Key to garden and village birds"⁴⁷ qui combine une description écrite des sons produits par les oiseaux et des fichiers audio permettant d'écouter un court extrait des chants, cris et sons des oiseaux (Trilar 2010).

Des clés d'identification moléculaires basées sur des **séquences d'ADN** (produits de PCR notamment) existent également, pour des organismes aussi divers que l'olivier, des levures, des insectes thysanoptères, des annélides polychètes et des moules (Gerke &

⁴⁴ spipoll.snv.jussieu.fr/mkey/mkey-spipoll.html

⁴⁵ orchidees-france.identificationkey.org/mkey.html

⁴⁶ chouettes-hiboux-de-france.identificationkey.org/mkey.html

⁴⁷ www.keytonature.eu/wiki/Key_to_Garden_and_Village_Birds




Tiedemann 2001, Cadez *et al.* 2002, Rugman-Jones *et al.* 2006, Blank *et al.* 2008, Alba *et al.* 2009).

2.1.4 Clés numériques à accès simple

Depuis la généralisation de l'informatique, les clés à accès simple se retrouvent également ailleurs que sur des supports papier. Ce type de clés peut ainsi être numérisé, et accompagné si besoin de supports tels que des illustrations, des fichiers audio et vidéo, des liens URL renvoyant vers des pages web d'intérêt. C'est le cas de la clé d'identification des coccinelles disponibles sur le site web "Discover Life"⁴⁸, et de la clé générée à partir de la base de données Perla pour les "invertébrés d'eau douce"⁴⁹. Pour faire un choix à chaque étape du chemin d'identification, l'utilisateur doit cliquer sur le lien web correspondant à la proposition qu'il estime juste (**Figure 10**), ce qui l'amène à l'étape suivante de la discrimination.

Accueil > Arbre de décision

Macroinvertébrés

<p>Cl. Insectes</p>  <p>3 paires de pattes</p>	<p>Hydracariens</p>  <p>4 paires de pattes</p>	<p>Cl. Crustacés</p>  <p>5 paires de pattes ou plus</p>
---	---	---

Arborescence des questions Précédent

Nombre de paires de pattes

- 3 paires
- 4 paires
- 5 paires ou plus

[111]

⁴⁸ www.discoverlife.org/20/q?guide=Ladybug

⁴⁹ www.perla.developpement-durable.gouv.fr/index.php/arbre

Figure 10. Interface de la clé d'identification de la base de données Perla. Pour faire un choix à chaque étape du chemin d'identification, l'utilisateur clique sur l'une des propositions (ici : "3 paires" ou "4 paires" ou "5 paires ou plus").

Cependant, malgré des avancées technologiques intéressantes – liens web, fichiers associés, accès facilité pour tous par la mise en ligne de l'outil – ces clés restent à accès simple. Un chemin d'identification est toujours imposé et l'utilisateur a l'obligation de passer chaque étape du chemin dans l'ordre afin d'obtenir une discrimination progressive.

Il existe d'autres méthodes de détermination que les clés d'identification à accès simple. Dans ce travail, nous nous sommes plus particulièrement intéressés aux clés d'identification numériques à accès libre, qui permettent de palier un certain nombre des inconvénients et limitations des clés d'identification à accès simple.

2.2 Les clés d'identification numériques à accès libre

2.2.1 Principes et caractéristiques

Depuis les années 1960, l'informatique a grandement aidé les biologistes dans leurs travaux pour la construction de clés d'identification. Des méthodes d'identification assistée par ordinateur (IAO) ont vu le jour pour permettre la réalisation de clés plus souples, dites interactives, et pour les diffuser sur le web. Parmi ces IAO, nous pouvons citer les logiciels "Online" (Pankhurst & Aitchinson 1975), Xper (Lebbe 1984), DELTA et IntKey (Dallwitz 1993).

Les clés d'identification interactives à accès libre sont des **outils numériques d'aide à l'identification. Elles ne suivent pas un cheminement imposé** comme c'est le cas pour les clés à accès simple (arbre de décision). L'utilisateur choisit lui-même les critères qu'il va utiliser pour la description de son spécimen, et l'ordre dans lequel il les utilise (**Figure 11**). Il n'y a pas de limite au nombre de taxons, de caractères et d'états de caractères qui peuvent être inclus dans la clé de détermination. Les caractères numériques peuvent être codés directement, sans être divisés en catégories (l'utilisateur peut indiquer un nombre) (Dallwitz

et al. 2013). Les données sont codées sous la forme d'une matrice taxons-caractères. Il est possible de mettre en place des dépendances, c'est-à-dire des caractères qui ne sont applicables que si certaines valeurs précises ont été sélectionnées pour un autre caractère donné – ils sont appelés "descripteurs-fils" dans Xper³.

L'utilisateur peut également revenir sur ses choix pour les modifier, ainsi que sélectionner plusieurs réponses possibles s'il doute à une étape. Il est possible de configurer l'algorithme de telle sorte qu'une identification correcte peut être réalisée malgré une ou plusieurs "erreurs" de la part de l'utilisateur au cours de la détermination. De manière générale, l'un des objectifs principaux des clés à accès libre est de réduire le risque d'obtenir une mauvaise détermination à la fin de la démarche d'identification.

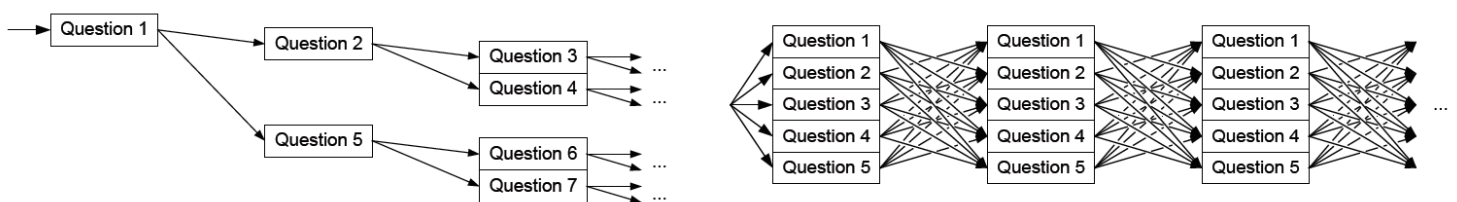


Figure 11. Représentation des chemins et étapes d'identification de l'utilisateur. Dans une clé à accès simple (à gauche), les chemins d'identification possibles sont imposés. Dans une clé à accès libre (à droite), le chemin d'identification est déterminé par les choix de l'utilisateur. Figure reprise de Hagedorn (2007).

A chaque étape du chemin d'identification, le logiciel élimine les taxons dont les attributs ne correspondent pas à ceux du spécimen à identifier. Ce processus continue jusqu'à ce qu'un (ou un petit nombre de) taxon(s) reste(nt) (Dallwitz *et al.* 2013).

Grâce à leur caractère numérique, le contenu de ce type de clés d'identification peut être modifié à n'importe quel moment – ajout, retrait ou modification des taxons, caractères et états de caractères inclus dans la base de données – et ainsi suivre l'évolution des connaissances, ou être adaptées pour différents types de publics (van Spronsen *et al.* 2010). Lorsque la plateforme d'édition de la base de données est en ligne (ce qui est le cas de Xper³), les bases de données et clés d'identification associées peuvent être construites, gérées et modifiées de façon collaborative, par simple connexion en ligne des différents constructeurs des clés. Les clés d'identification peuvent être accessibles à tous *via* un lien URL.

Le prochain sous-chapitre fera une comparaison plus approfondie entre les clés à accès simple et les clés numériques à accès libre.

2.2.2 Pourquoi ce type de clés ? Avantages et inconvénients des clés d'identification numériques à accès libre

Les clés à accès simple présentent un certain nombre d'inconvénients et limites. Certains d'entre eux peuvent être contournés ou réduits par l'utilisation d'outils d'identification numériques à accès libre. Le **Tableau 1** résume les principales limites des clés à accès simple et les caractéristiques des clés à accès libres qui permettent de pallier ces inconvénients.

De manière générale, les clés d'identification informatisées interactives apportent une facilité d'utilisation et une flexibilité supérieures à celles des clés à accès unique (Edwards & Morse 1995). Cependant, elles apportent leur lot de difficultés :

- L'abondance de critères de détermination proposés et la liberté de choix de l'utilisateur rend le nombre d'étapes d'identification, et donc le temps passé pour effectuer la détermination, très variables selon l'utilisateur et l'état du spécimen à décrire – et souvent supérieurs à l'optimal
- L'absence de chemin d'identification imposé et l'abondance de possibilités peut rendre plus difficile sa tâche à l'utilisateur, qui peut se sentir perdu ou submergé
- La possibilité de sélectionner plusieurs choix aux étapes d'identification fait que plusieurs propositions finales de taxons peuvent être conservées ; l'utilisateur n'obtient pas alors d'identification unique.

Tableau 1. Tableau récapitulatif des avantages et inconvénients des clés d'identification à accès simple et des clés d'identification à accès libre. Informations partiellement reprises et complétées à partir de Edwards & Morse (1995), Hagedorn *et al.* (2010) Ung *et al.* (2010) Varese *et al.* (2010) et Dallwitz *et al.* (2013).

Clé d'identification à accès simple	Clé d'identification (numérique) à accès libre
Fixité qui peut la rendre rapidement obsolète (découverte d'espèces, changements dans la nomenclature et la taxonomie...)	Peut être corrigée et mise à jour en temps réel pour suivre les nouvelles connaissances
Contenu défini par l'auteur et restreint aux choix et à la logique de ce dernier	Multiplication possible des critères discriminants et de leur type (morphologie, écologie, comportement...)
La quantité et la nature des informations contenues dans la clé sont limitées	La clé peut contenir une quantité très importante d'informations et de ressources : illustrations, fichiers audio et vidéo, texte...
Chaque clé est créée pour une situation donnée : un ensemble d'organismes, une aire géographique ou un type d'écosystème, un stade de développement (ex : larve ou adulte), un type d'utilisateurs	Possibilité d'adapter facilement la clé à différents types d'utilisateurs (grand public, scolaires, naturalistes, spécialistes...)
À chaque étape de l'identification, les choix possibles sont discrets et fixés – par ex : "longueur inférieure à 10 mm" vs "longueur supérieure ou égale à 10 mm"	L'utilisateur peut proposer des états de caractère continus – par ex, indiquer "32" dans un encart demandant de donner une longueur en mm
Chemin d'identification imposé et unique pour un taxon donné : l'utilisateur doit suivre les étapes dans l'ordre. S'il se trompe à une étape, l'identification sera incorrecte. S'il ne sait pas faire un choix à une étape (il est "bloqué"), l'identification est impossible.	Démarche d'identification libre, l'utilisateur choisit l'ordre dans lequel il renseigne les informations pour chaque étape. Possibilité de sélectionner plusieurs propositions à chaque étape en cas de doute.
Identification correcte conditionnée à un spécimen complet et en bon état, au risque de ne pas pouvoir passer certaines étapes du chemin d'identification	L'absence de certaines structures morfo-anatomiques ou le mauvais état d'un spécimen n'empêchent pas d'obtenir une identification finale ; si certaines parties du spécimen sont absentes ou trop abîmées, l'utilisateur peut ignorer les étapes associées à ces structures
L'identification finale est soit juste, soit incorrecte, ce qui peut avoir des conséquences sur le travail scientifique	Possibilité d'obtenir une identification finale même si l'utilisateur a "mis de côté" certains caractères ou sélectionné plusieurs possibilités pour certaines étapes. L'identification peut être partielle (proposition finale de plusieurs taxons répondant aux critères sélectionnés).
Clés d'identification le plus souvent imprimées sur papier : encombrement si utilisation de livres sur le terrain, accessibilité limitée	Clés d'identification facilement transportables sur des appareils mobiles (smartphones et tablettes), et plus facilement accessibles et diffusables grâce au numérique

2.2.3 *L'apport des clés numériques à accès multiples pour les sciences participatives : l'exemple du Spipoll*

Le programme Spipoll (Suivi Photographique des Insectes POLLinisateurs) est un projet de sciences participatives ayant pour but de collecter sur l'ensemble de la France métropolitaine et selon un protocole défini, des observations d'insectes pollinisateurs et/ou floricoles. Ces données sont ensuite utilisées à des fins scientifiques d'analyse des variations de diversité, de densité et de structure des réseaux de pollinisation sur le territoire. Depuis sa création en 2010, des articles scientifiques utilisant les données récoltées par le programme Spipoll ont été publiés dans des revues internationales à comité de lecture (Deguines *et al.* 2012, Shwartz *et al.* 2013, Deguines *et al.* 2016, Desaegher *et al.* 2018, Levé *et al.* 2019).

Les volontaires effectuent des "sessions d'observation" qui durent au moins 20 minutes (durée minimale requise dans le protocole) mais qui peuvent être prolongées si le volontaire le désire. Ce dernier choisit une espèce de plante à fleurs (plantée par l'homme ou non), prend en photo la plante et son environnement, note des paramètres environnementaux (température, milieux de vie proches, intensité et régularité du vent, ensoleillement...) ainsi que la date et l'heure de la session d'observation. Pendant l'intégralité de la session, le volontaire photographie tous les insectes qu'il voit se poser sur les fleurs de l'espèce qu'il a choisie. Ensuite, il trie ses photos pour garder un cliché par type d'insecte observé, et dépose ses données sur le site web du programme⁵⁰. L'ensemble "photos de la fleur + photo de l'environnement + une photo par type d'insecte qui s'est posé sur une fleur" est appelé "collection" (**Figure 12**).

⁵⁰ spipoll.org

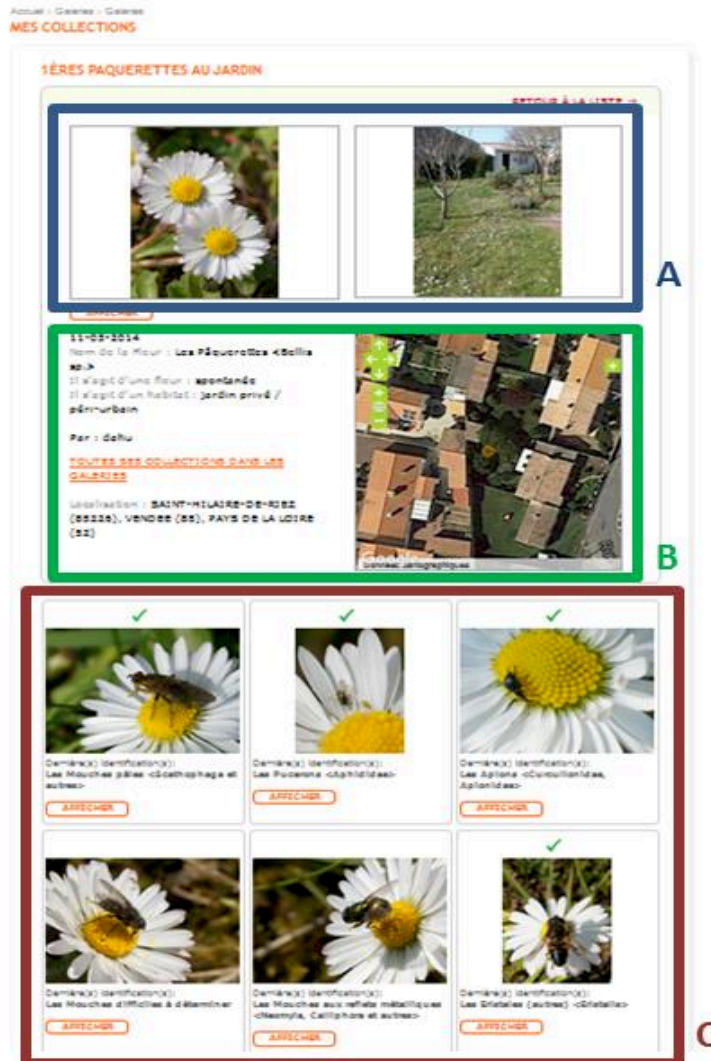


Figure 12. Une "collection" déposée sur le site du programme Spipoll.

A. Photos de l'espèce de fleur choisie et de l'environnement direct des fleurs.

B. Date, nom de l'espèce de fleur, type d'habitat, commune où a été faite la session d'observation, pseudonyme du participant, localisation GPS.

C. Chaque photo représente un type d'insecte observé sur les fleurs de l'espèce choisie. Un signe ✓ de couleur verte au-dessus d'une photo indique que l'identification du taxon a été validée (ou corrigée) par un expert entomologiste de l'OPIE.

Dans un objectif de gain de temps pour les chercheurs mais aussi d'apprentissage de l'entomologie pour le public non scientifique, le participant a ensuite la possibilité d'identifier les insectes présents sur ses photos. Pour aider à cela, le site du Spipoll met à disposition une clé d'identification numérique et interactive des insectes pollinisateurs et floricoles de

France métropolitaine. Cette clé d'identification a été développée sous Xper³, plateforme qui sera présentée au chapitre 3.

La clé interactive "Insectes" du Spipoll est accessible via ce lien : <http://spipoll.snv.jussieu.fr/mkey/mkey-spipoll.html>.

Chaque observation est identifiée, représentée par une photographie et stockée dans une base de données. Cet outil fournit un excellent terrain d'expérimentation pour des études sur l'utilisation des outils du programme Spipoll par le public : les utilisateurs sont très nombreux, certains sont novices, d'autres expérimentés et l'on dispose de quelques informations sur chaque utilisateur (date de première participation, nombre de participations, compétences naturalistes...).

Les déterminations des taxons proposées par les utilisateurs, associées chacune à une photographie, sont ensuite vérifiées, validées ou corrigées par des experts entomologistes de l'OPIE. Les utilisateurs peuvent échanger sur le site web du programme, via des commentaires sur les pages des photos et des collections déposées, ou via le forum du Spipoll⁵¹.

Des rencontres régulières ("Rencontres nationales" bisannuelles) entre les coordinateurs du programme et les "Spipolliens" permettent d'obtenir un retour d'expérience et d'évaluer les besoins futurs du projet et les améliorations à apporter à la clé d'identification.

Ce chapitre a traité du fonctionnement et des caractéristiques des clés d'identification. En particulier, nous avons vu que les clés d'identification numériques interactives à accès libre offrent de multiples possibilités pour la détermination des êtres vivants. Le chapitre 3 aura pour objet la construction de telles clés, à l'aide d'un logiciel particulier nommé Xper³, dans l'objectif de faciliter l'identification de l'entomofaune.

⁵¹ www.spipoll.org/forum

Chapitre 3 – Comment construire des clés d'identification numériques ?

Ce chapitre couvre la description du matériel et des méthodes utilisées pour la construction des sept clés d'identification numériques de ce travail de thèse. Je détaille ici l'échantillonnage taxonomique et les caractéristiques des logiciels utilisés. Les bases de données ont servi de support à la création des clés d'identification. Dans la seconde partie du chapitre, je présente les caractéristiques principales des clés d'identification obtenues.

3.1 Taxons étudiés et logiciels utilisés

3.1.1 Échantillonnage taxonomique pour l'entomofaune

Les clés dites "des larves" (T1, T2 et T4) traitent uniquement de la morphologie des stades larvaires les plus développés (également appelés "nymphe"), précédant la mue imaginale. Contrairement aux stades plus jeunes (pas complètement développés), la nymphe présente toutes les caractéristiques morphologiques observables au niveau larvaire, et son observation permet donc sans ambiguïté de discriminer chaque spécimen à l'espèce.

Les clés dites "des imagos" (T3, T5, T6 et T7) concernent les individus "adultes" ayant effectué leur dernière mue et possédant leurs couleurs et morphologie finales. Les "subimagos" ne sont pas traités.

3.1.1a Insectes aquatiques de la région Ile-de-France

Pour les insectes aquatiques – odonates, éphémères, plécoptères et coléoptères, correspondant aux bases de données T1 à T6 – j'ai inclus toutes les espèces connues (ou supposées) comme étant présentes, au moins à un moment de leur cycle de vie, dans la région Ile-de-France. La liste d'espèces provient principalement de données récentes (depuis 1950) avec des occurrences vérifiées sur le terrain, fournies par les structures et programmes suivants : l'Office Pour les Insectes et leur Environnement (OPIE)⁵², la base de données du

⁵² <http://www.insectes.org>

Cettia-idf (observatoire francilien de la biodiversité)⁵³, et l'Inventaire national du patrimoine naturel (INPN)⁵⁴.

Quelques espèces ont été ajoutées sur les conseils d'experts des taxons, qui assument qu'elles pourraient être rencontrées en Ile-de-France, sachant qu'elles sont présentes dans un ou plusieurs département(s) limitrophe(s) de la région (Michel Brulin, Jacques Le Doaré, comm. pers.).

Les ordres traités ont été choisis sur la base des ressources disponibles pour la création de bases de données, de l'intérêt pour la conservation des milieux aquatiques, et de la faisabilité sur un travail de 3 ans. Il s'agit d'ordres utilisés pour des inventaires de biodiversité dans le cadre de la Directive européenne Cadre sur l'Eau (voir 1.3.1). Le nombre de taxons reste raisonnable pour la faisabilité d'un travail exploratoire de thèse, et une quantité assez conséquente d'informations est disponible pour permettre la construction d'une clé d'identification juste et fiable basée sur la morphologie.

La priorité a été donnée aux larves d'odonates par rapport aux imagos, dû au manque d'outils et de guides d'identification pour les larves, qui sont beaucoup moins connues.

Seuls les coléoptères adultes (imagos) ont été inclus, car les connaissances sont très lacunaires pour les larves, et pour certaines espèces la correspondance taxonomique entre larve et adulte n'est pas connue. La clé d'identification des coléoptères aquatiques est la seule à proposer une détermination au genre et pas à l'espèce, compte tenu du nombre important d'espèces présentes en Ile-de-France (environ 250).

Les bases de données des insectes aquatiques réalisées pendant cette thèse sont donc ainsi composées :

- Odonates : 59 espèces (22 zygoptères et 37 anisoptères) appartenant à 28 genres et 9 familles (Clé T1, **Tableau 2**)
- Éphémères : 39 espèces appartenant à 21 genres et 9 familles (Clés T2 et T3, **Tableau 3**)

⁵³ <http://cettia-idf.fr>

⁵⁴ <https://inpn.mnhn.fr>

- Plécoptères : 10 espèces appartenant à 8 genres et 5 familles (Clés T4 et T5, **Tableau 4**)
- Coléoptères aquatiques : 68 genres appartenant à 15 familles (Clé T6, **Tableau 5**).

3.1.1b Coccinelles (imagos) de France métropolitaine

L'objectif des clés construites au cours de ce doctorat était qu'elles puissent être aisément appréhendées par tous les types de publics adultes, en utilisant des caractères visibles à l'œil nu ou à l'aide d'une loupe binoculaire. De ce fait, nous avons restreint l'échantillonnage taxonomique aux tribus Chilacorini, Coccinellini and Epilachnini. Les membres de ces tribus sont relativement grands (3 à 9 mm) et présentent une importante diversité de colorations, ce qui les rend facilement visibles dans leur environnement et identifiables par des non-spécialistes. Nous avons également inclus dans l'échantillonnage (sous forme d'items dans la clé d'identification) les formes de colorations les plus courantes, afin de couvrir au mieux les variabilités intraspécifiques de ces espèces.

La liste d'espèces est tirée de Tronquet (2014) et la taxonomie suit Seago *et al.* (2011). Pour plus de détails concernant les choix faits pour l'échantillonnage taxonomique, se référer à l'article correspondant (Jouveau & Delaunay *et al.* 2018 – **Annexe 1**). La plupart des spécimens ont été examinés dans les collections du MNHN.

La base de connaissances des coccinelles de France réalisée pendant cette thèse comprend donc 47 espèces appartenant à 24 genres et 3 tribus, tous inclus dans la sous-famille Coccinellinae et la famille Coccinellidae (Clé T7, **Tableau 6**).

Tableau 2. Couverture taxonomique de la base de connaissances T1 : odonates d'Ile-de-France

Sous-ordre	Famille	Genre	Espèce
Zygoptera	Calopterygidae	<i>Calopteryx</i>	<i>Calopteryx splendens</i>
			<i>Calopteryx virgo</i>
	Lestidae	<i>Chalcolestes</i>	<i>Chalcolestes viridis</i>
			<i>Lestes</i>
		<i>Lestes</i>	<i>Lestes barbarus</i>
			<i>Lestes dryas</i>
			<i>Lestes sponsa</i>
			<i>Lestes virens</i>
			<i>Sympecma</i>
	Coenagrionidae	<i>Ceriagrion</i>	<i>Ceriagrion tenellum</i>
			<i>Coenagrion</i>
		<i>Coenagrion</i>	<i>Coenagrion mercuriale</i>
			<i>Coenagrion puella</i>
			<i>Coenagrion pulchellum</i>
			<i>Coenagrion scitulum</i>
		<i>Enallagma</i>	<i>Enallagma cyathigerum</i>
		<i>Erythromma</i>	<i>Erythromma lindenii</i>
	<i>Erythromma najas</i>		
	<i>Erythromma viridulum</i>		
	<i>Ischnura</i>	<i>Ischnura elegans</i>	
<i>Ischnura pumilio</i>			
<i>Pyrrhosoma</i>	<i>Pyrrhosoma nymphula</i>		
Platycnemididae	<i>Platycnemis</i>	<i>Platycnemis acutipennis</i>	
		<i>Platycnemis pennipes</i>	
Anisoptera	Aeshnidae	<i>Aeshna</i>	<i>Aeshna affinis</i>
			<i>Aeshna cyanea</i>
			<i>Aeshna grandis</i>
			<i>Aeshna isoceles</i>
			<i>Aeshna mixta</i>
		<i>Anax</i>	<i>Anax imperator</i>
		<i>Anax</i>	<i>Anax parthenope</i>
		<i>Boyeria</i>	<i>Boyeria irene</i>
		<i>Brachytron</i>	<i>Brachytron pratense</i>
		<i>Hemianax</i>	<i>Hemianax ephippiger</i>
	Gomphidae	<i>Gomphus</i>	<i>Gomphus pulchellus</i>
			<i>Gomphus simillimus</i>
			<i>Gomphus vulgatissimus</i>
	<i>Onychogomphus</i>	<i>Onychogomphus forcipatus</i>	
	Cordulegastridae	<i>Cordulegaster</i>	<i>Cordulegaster boltonii</i>
	Corduliidae	<i>Cordulia</i>	<i>Cordulia aenea</i>
		<i>Epithea</i>	<i>Epithea bimaculata</i>
		<i>Oxygastra</i>	<i>Oxygastra curtisii</i>
		<i>Somatochlora</i>	<i>Somatochlora flavomaculata</i>
	<i>Somatochlora metallica</i>		
	Libellulidae	<i>Crocothemis</i>	<i>Crocothemis erythraea</i>
			<i>Leucorrhinia</i>
		<i>Leucorrhinia</i>	<i>Leucorrhinia caudalis</i>
			<i>Leucorrhinia pectoralis</i>
			<i>Libellula</i>
		<i>Libellula</i>	<i>Libellula depressa</i>
			<i>Libellula fulva</i>
	<i>Libellula quadrimaculata</i>		
	<i>Orthetrum</i>	<i>Orthetrum albistylum</i>	
		<i>Orthetrum brunneum</i>	

			<i>Orthetrum cancellatum</i>
			<i>Orthetrum coerulescens</i>
		<i>Sympetrum</i>	<i>Sympetrum danae</i>
			<i>Sympetrum flaveolum</i>
			<i>Sympetrum fonscolombii</i>
			<i>Sympetrum meridionale</i>
			<i>Sympetrum sanguineum</i>
			<i>Sympetrum striolatum</i>
			<i>Sympetrum vulgatum</i>

Tableau 3. Couverture taxonomique des bases de données T2 et T3 : éphémères d'Ile-de-France

Famille	Genre	Espèce
Ameletidae	<i>Metreletus</i>	<i>Metreletus balcanicus</i>
Baetidae	<i>Baetis</i>	<i>Baetis buceratus</i>
		<i>Baetis fuscatus</i>
		<i>Baetis liebenauae</i>
		<i>Baetis lutheri</i>
		<i>Baetis rhodani</i>
		<i>Baetis scambus</i>
		<i>Baetis vernus</i>
	<i>Baetopus</i>	<i>Baetopus tenellus</i>
	<i>Centroptilum</i>	<i>Centroptilum luteolum</i>
	<i>Cloeon</i>	<i>Cloeon dipterum</i>
<i>Cloeon simile</i>		
<i>Labiobaetis</i>	<i>Labiobaetis atrebatinus</i>	
<i>Procloeon</i>	<i>Procloeon bifidum</i>	
	<i>Procloeon pennulatum</i>	
Caenidae	<i>Brachycercus</i>	<i>Brachycercus harrisellus</i>
	<i>Caenis</i>	<i>Caenis horaria</i>
		<i>Caenis lactea</i>
		<i>Caenis luctuosa</i>
		<i>Caenis macrura</i>
		<i>Caenis pusilla</i>
	<i>Caenis robusta</i>	
<i>Cercobrachys</i>	<i>Cercobrachys minutus</i>	
Ephemeridae	<i>Ephemera</i>	<i>Ephemera danica</i>
		<i>Ephemera glaucops</i>
		<i>Ephemera lineata</i>
		<i>Ephemera vulgata</i>
	<i>Serratella</i>	<i>Serratella ignita</i>
Heptageniidae	<i>Ecdyonurus</i>	<i>Ecdyonurus dispar</i>
		<i>Ecdyonurus venosus</i>
	<i>Electrogena</i>	<i>Electrogena ujhelyii</i>
	<i>Heptagenia</i>	<i>Heptagenia flava</i>
<i>Heptagenia sulphurea</i>		
Leptophlebiidae	<i>Habrophlebia</i>	<i>Habrophlebia lauta</i>
	<i>Leptophlebia</i>	<i>Leptophlebia marginata</i>
	<i>Paraleptophlebia</i>	<i>Paraleptophlebia submarginata</i>
Polymitarcyidae	<i>Ephoron</i>	<i>Ephoron virgo</i>
Potamanthidae	<i>Potamanthus</i>	<i>Potamanthus luteus</i>
Prosopistomatidae	<i>Prosopistoma</i>	<i>Prosopistoma pennigerum</i>

Tableau 4. Couverture taxonomique des bases de données T4 et T5 : plécoptères d'Ile-de-France

Famille	Genre	Espèce
Capniidae	<i>Zwicknia</i>	<i>Zwicknia bifrons</i>
Leuctridae	<i>Leuctra</i>	<i>Leuctra geniculata</i>
		<i>Leuctra nigra</i>
Nemouridae	<i>Nemoura</i>	<i>Nemoura cinerea cinerea</i>
	<i>Nemurella</i>	<i>Nemurella pictetii</i>
Perlodidae	<i>Isogenus</i>	<i>Isogenus nubecula</i>
	<i>Isoperla</i>	<i>Isoperla grammatica</i>
Taeniopterygidae	<i>Brachyptera</i>	<i>Brachyptera risi</i>
	<i>Taeniopteryx</i>	<i>Taeniopteryx nebulosa</i>
		<i>Taeniopteryx schoenemundi</i>

Tableau 5. Couverture taxonomique de la base de connaissances T6 : coléoptères aquatiques imagos d'Ile-de-France

Famille	Genre	Espèce (pour les genres monospécifiques en Ile-de-France)
Dytiscidae	<i>Acilius</i>	
	<i>Agabus</i>	
	<i>Bidessus</i>	
	<i>Colymbetes</i>	<i>Colymbetes fuscus</i>
	<i>Cybister</i>	<i>Cybister lateralimarginalis</i>
	<i>Dytiscus</i>	
	<i>Graphoderus</i>	
	<i>Graptodytes</i>	
	<i>Hydaticus</i>	
	<i>Hydroglyphus</i>	<i>Hydroglyphus geminus</i>
	<i>Hydroporus</i>	
	<i>Hydrovatus</i>	
	<i>Hygrotus</i>	
	<i>Hyphydrus</i>	<i>Hyphydrus ovatus</i>
	<i>Ilybius</i>	
	<i>Laccophilus</i>	
	<i>Laccornis</i>	<i>Laccornis oblongus</i>
	<i>Liopterus</i>	<i>Liopterus haemorrhoidalis</i>
	<i>Nebrioporus</i>	
	<i>Platambus</i>	<i>Platambus maculatus</i>
	<i>Porhydrus</i>	<i>Porhydrus lineatus</i>
	<i>Rhantus</i>	
	<i>Scarodytes</i>	<i>Scarodytes halensis</i>
<i>Stictionectes</i>	<i>Stictionectes lepidus</i>	
<i>Stictotarsus</i>	<i>Stictotarsus duodecimpustulatus</i>	
<i>Yola</i>	<i>Yola bicarinata</i>	
Haliplidae	<i>Brychius</i>	<i>Brychius elevatus</i>
	<i>Haliplus</i>	
	<i>Peltodytes</i>	<i>Peltodytes caesus</i>
Noteridae	<i>Noterus</i>	
Hygrobiiidae	<i>Hygrobia</i>	<i>Hygrobia hermanni</i>
Gyrinidae	<i>Aulonogyrus</i>	<i>Aulonogyrus concinnus</i>
	<i>Gyrinus</i>	
	<i>Orectochilus</i>	<i>Orectochilus villosus</i>
Helophoridae	<i>Helophorus</i>	
Hydrochidae	<i>Hydrochus</i>	
Hydrophilidae	<i>Anacaena</i>	
	<i>Berosus</i>	
	<i>Cercyon</i>	
	<i>Coelostoma</i>	<i>Coelostoma orbiculare</i>
	<i>Cymbiodyta</i>	<i>Cymbiodyta marginella</i>
	<i>Enochrus</i>	
	<i>Helochares</i>	
	<i>Hydrobius</i>	<i>Hydrobius fuscipes</i>
	<i>Hydrochara</i>	<i>Hydrochara caraboides</i>
	<i>Laccobius</i>	
	<i>Limnoxenus</i>	<i>Limnoxenus niger</i>
Hydraenidae	<i>Hydraena</i>	
	<i>Limnebius</i>	

	<i>Ochthebius</i>	
Dryopidae	<i>Dryops</i>	
Heteroceridae	<i>Heterocerus</i>	
Elmidae	<i>Elmis</i>	
	<i>Limnius</i>	<i>Limnius volckmari</i>
	<i>Macronychus</i>	<i>Macronychus quadrituberculatus</i>
	<i>Riolus</i>	
	<i>Stenelmis</i>	<i>Stenelmis canaliculata</i>
Scirtidae	<i>Contacyphon</i>	
	<i>Elodes</i>	
	<i>Microcara</i>	<i>Microcara testacea</i>
	<i>Prionocyphon</i>	<i>Prionocyphon serricornis</i>
	<i>Scirtes</i>	
Chrysomelidae	<i>Donacia</i>	
	<i>Plateumaris</i>	
Cucurlionidae	<i>Amalorrhynchus</i>	<i>Amalorrhynchus melanarius</i>
	<i>Bagous</i>	
	<i>Drupenatus</i>	<i>Drupenatus nasturtii</i>
	<i>Phytobius</i>	<i>Phytobius leucogaster</i>

Tableau 6. Couverture taxonomique de la base de connaissances T7 : coccinelles de France métropolitaine

Tribu	Genre	Espèce ou forme
Chilocorini	<i>Chilocorus</i>	<i>Chilocorus bipustulatus</i>
		<i>Chilocorus renipustulatus</i>
	<i>Exochomus</i>	<i>Exochomus octosignatus</i>
		<i>Exochomus quadripustulatus</i>
<i>Parexochomus</i>	<i>Parexochomus nigromaculatus</i>	
Coccinellini	<i>Halyzia</i>	<i>Halyzia sedecimguttata</i>
	<i>Psyllobora</i>	<i>Psyllobora vigintiduopunctata</i>
	<i>Vibidia</i>	<i>Vibidia duodecimguttata</i>
	<i>Anisosticta</i>	<i>Anisosticta novemdecimpunctata</i>
	<i>Coccinula</i>	<i>Coccinula quatuordecimpustulata</i>
		<i>Coccinula sinuatoinarginata</i>
	<i>Tytthaspis</i>	<i>Tytthaspis sedecimpunctata</i> – f. <i>duodecimpunctata</i>
	<i>Adalia</i>	<i>Adalia bipunctata</i> – f. <i>annulata</i>
		<i>Adalia bipunctata</i> – f. <i>pantherina</i>
		<i>Adalia bipunctata</i> – f. <i>quadrimaculata</i>
		<i>Adalia bipunctata</i> – f. <i>sempustulata</i>
		<i>Adalia bipunctata</i> – f. <i>typica</i>
		<i>Adalia conglomerata</i> – f. <i>decas</i>
		<i>Adalia decempunctata</i> – f. <i>decempustulata</i>
		<i>Adalia decempunctata</i> – f. <i>guttatopunctata</i>
		<i>Adalia decempunctata</i> – f. <i>lutea</i>
<i>Adalia decempunctata</i> – f. <i>quadripunctata</i>		
<i>Adalia decempunctata</i> – f. <i>scribai</i>		
<i>Adalia decempunctata</i> – f. <i>subpunctata</i>		
<i>Adalia decempunctata</i> – f. <i>terna</i>		
<i>Adalia decempunctata</i> – f. <i>typica</i>		

	<i>Anatis</i>	<i>Anatis ocellata</i>
	<i>Aphidecta</i>	<i>Aphidecta obliterata</i> – f. <i>typica</i>
	<i>Calvia</i>	<i>Calvia decemguttata</i>
		<i>Calvia quatuordecimguttata</i>
		<i>Calvia quindecimguttata</i>
	<i>Ceratomegilla</i>	<i>Ceratomegilla alpina</i>
		<i>Ceratomegilla notata</i>
		<i>Ceratomegilla rufocincta</i>
		<i>Ceratomegilla undecimnotata</i>
	<i>Coccinella</i>	<i>Coccinella venusta</i>
		<i>Coccinella hieroglyphica</i>
		<i>Coccinella magnifica</i>
		<i>Coccinella quinquepunctata</i>
		<i>Coccinella septempunctata</i>
		<i>Coccinella undecimpunctata</i>
	<i>Harmonia</i>	<i>Harmonia axyridis</i> – f. <i>conspicua</i>
		<i>Harmonia axyridis</i> – f. <i>novemdecimsignata</i>
		<i>Harmonia axyridis</i> – f. <i>spectabilis</i>
		<i>Harmonia conformis</i>
		<i>Harmonia quadripunctata</i> – f. <i>sedecimpunctata</i>
		<i>Harmonia quadripunctata</i> – f. <i>typica</i>
	<i>Hippodamia</i>	<i>Hippodamia septemmaculata</i> – f. <i>cestiva</i>
		<i>Hippodamia tredecimpunctata</i>
		<i>Hippodamia variegata</i> – f. <i>quinquemaculata</i>
		<i>Hippodamia variegata</i> – f. <i>undecimpunctata</i>
		<i>Hippodamia variegata</i> – f. <i>carpini</i>
		<i>Hippodamia variegata</i> – f. <i>constellata</i> <i>Hippodamia variegata</i> – f. <i>velox</i>
	<i>Myrrha</i>	<i>Myrrha octodecimguttata</i>
	<i>Myzia</i>	<i>Myzia oblongoguttata</i>
	<i>Oenopia</i>	<i>Oenopia conglobata</i>
		<i>Oenopia doublieri</i>
		<i>Oenopia impustulata</i>
		<i>Oenopia lyncea</i>
	<i>Propylea</i>	<i>Propylea quatuordecimpunctata</i> – f. <i>weisei</i>
	<i>Sospita</i>	<i>Sospita vigintiguttata</i>
Epilachnini	<i>Henosepilachna</i>	<i>Henosepilachna argus</i>
		<i>Henosepilachna elaterii</i>
	<i>Subcoccinella</i>	<i>Subcoccinella vigintiquatuorpunctata</i> – f. <i>limbata</i> <i>Subcoccinella vigintiquatuorpunctata</i> – f. <i>typica</i>

3.1.2 Xper, une suite d'outils pour la construction et l'analyse de bases de données et de clés d'identification

Xper³ (xper3.fr), est un **système permettant de stocker et d'éditer des données descriptives en ligne. Il est dédié à la gestion de descriptions taxonomiques structurées et à leur analyse.** Il contient plusieurs outils d'analyse de cohérence des données descriptives. Il est muni d'un module de génération de clés d'identification interactives à accès libre à partir des bases de données stockées dans le système (*Kerner et al.* 2011, Corvez & Grand 2014, Martin *et al.* 2015). Xper³ offre également la possibilité de partager les données avec des utilisateurs tiers, rendant ainsi possible la gestion collaborative de données descriptives (*Vignes-Lebbe et al.* 2016). Un wiki et une documentation pour Xper³ sont disponibles en ligne⁵⁵.

Ce modèle de représentation des connaissances formalise les descriptions des spécimens en les décomposant en plusieurs objets :

- Les **“items”** sont les entités décrites auxquelles les spécimens seront identifiés. Dans le cadre d'études de sciences naturelles, un item peut représenter une espèce ou sous-espèce, ou encore un groupe d'organismes (pas forcément inclus dans un même clade). Les auteurs de la clé définissent la précision taxonomique qui peut être atteinte grâce à l'identification par la clé – précision qui peut être limitée par l'étendue des connaissances pour chaque groupe biologique.
- Les **“descripteurs”** sont des critères (morpho-anatomiques, spatio-temporels, écologiques...) servant à décrire des taxons. Ils peuvent être discrets (composés de plusieurs états délimités) ou continus (associés à une valeur ou un ensemble de valeurs numériques, par exemple “taille en mm”). Des descripteurs-fils peuvent être également créés, qui ne sont accessibles qu'à la condition qu'un ou plusieurs état(s) du descripteur-père ai(en)t été sélectionné(s).
- Un **“état de descripteur”** est une valeur ou un état pris par un descripteur. Par exemple, le descripteur “couleurs des ailes” pourrait avoir comme états possibles

⁵⁵ wiki.xper3.fr/index.php

“blanc”, “rouge” et “jaune”. Les descripteurs dits “catégoriques” (discrets) comportent entre deux états de descripteurs ou plus. Les descripteurs dits “numériques” (continus) acceptent des valeurs numériques, par exemple s’il s’agit d’une taille (ex : “longueur de l’abdomen en mm”). L’utilisateur indique alors manuellement une valeur numérique (ex : 5,3).

Une **base de données (ou base de connaissances) Xper** est donc un ensemble d’items (dans le cas de ce travail de thèse, des taxons), tous décrits selon un même modèle et une même terminologie. Des ressources documentaires (textes, images, hyperliens, fichiers sonores...) peuvent être associées aux objets (items, descripteurs et états de descripteurs). Ces ressources permettent de guider l’utilisateur afin de faciliter sa démarche et d’obtenir une détermination la plus exacte possible pour son spécimen.

Le modèle descriptif est composé d’une hiérarchie de descripteurs (ici, des caractères morpho-anatomiques discrets et continus) et une terminologie choisie pour exprimer les différentes valeurs possibles de ces descripteurs (états de descripteurs).

Lors de l’identification par l’utilisateur, l’algorithme inclus dans le système Xper applique un raisonnement déductif par la négation afin d’éliminer les taxons non compatibles. À chaque étape, l’algorithme de la clé “conseille” l’utilisateur en lui proposant une liste de descripteurs triés selon leur “pouvoir discriminant”, c’est-à-dire la capacité à différencier les taxons restant compatibles avec les informations fournies lors de la démarche d’identification.

Le **pouvoir discriminant** d’un descripteur peut être calculé à l’aide de plusieurs indices, dont “l’indice original Xper” (Lebbe 1991), qui a été implémenté dans Xper², une version précédente de Xper qui n’était pas en ligne mais un logiciel à installer sur le PC (Ung *et al.* 2010). L’indice Xper se base sur les incompatibilités entre les descriptions. Deux items sont incompatibles (ou non-similaires ou discriminés) si pour un descripteur donné, il n’y a aucun état de descripteur en commun pour ces deux items. Pour chaque descripteur et pour tous les items, l’indice Xper a une valeur comprise entre 0 (le pouvoir discriminant du descripteur est nul) et 1 (le descripteur peut discriminer tous les items). Il est possible de comparer les descriptions entre différents items ou groupes d’items grâce aux fonctions

“comparer des items” et “comparer des groupes” dans l’onglet “outil” de l’interface du logiciel (**Figure 13**). Lors de la création de la description, il est possible d’en avoir une vue d’ensemble avec l’outil “voir la matrice de description” (**Figure 14**), et de vérifier si des erreurs ont été faites (ou des descriptions non complétées pour certains items et descripteurs) grâce à la fonction “checkboxbase”. Il est possible d’exprimer une “valeur inconnue” pour certains descripteurs dont on ne connaît pas les états pour un item donné, et le logiciel le prend en compte pour la description lors de l’utilisation de la clé d’identification. Cependant, il est conseillé de limiter le nombre de “valeurs inconnues” car plus la base de connaissances est complète, plus l’identification sera rapide, fiable et juste.

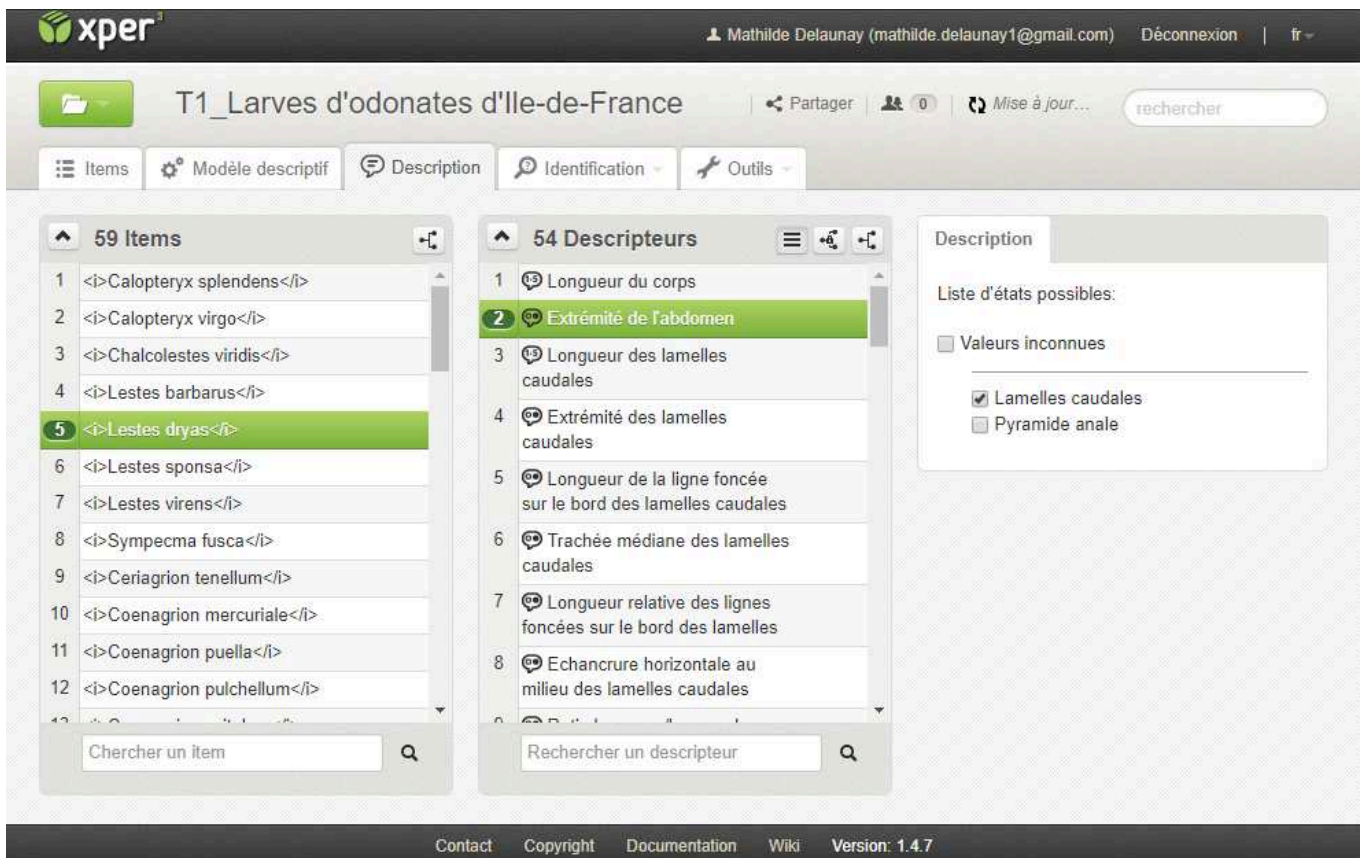


Figure 13. Capture d’écran de l’interface de création d’une base de connaissances sur Xper³. L’utilisateur compose la liste des items et des descripteurs qu’il souhaite utiliser dans sa base de données, ainsi que les états associés à chaque descripteur (onglets “items” et “descripteurs”). L’onglet “description” (montré sur cette figure) permet d’indiquer au logiciel le(s) état(s) de descripteurs qui correspondent à la description de chaque item pour chaque descripteur.

T1_Larves d'odonates d'Ile-de-France								
xper ³	Longueur du corps	Extrémité de l'abdomen	Longueur des lamelles caudales	Extrémité des lamelles caudales	Longueur de la ligne foncée sur le bord des lamelles caudales	Trachée médiane des lamelles caudales	Longueur relative des lignes foncées sur le bord des lamelles	Echancrure horizontale milieu des lamelles caudales
Calopteryx splendens	[min,max] = [27,38] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Lamelles caudales	[min,max] = [2,14] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Pointue	Unknown	• Saillante	Unknown	• Absente
Calopteryx virgo	[min,max] = [23,34] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Lamelles caudales	[min,max] = [0,11] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Pointue	Unknown	• Saillante	Unknown	• Absente
Chalcolestes viridis	[min,max] = [23,28] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Lamelles caudales	[min,max] = [7,9] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Arrondie • Pointue	Unknown	Unknown	Unknown	• Absente
Lestes barbarus	[min,max] = [23,28] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Lamelles caudales	[min,max] = [7,8] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Pointue	• Environ la longueur totale des lamelles	Unknown	Unknown	• Absente
Lestes dryas	[min,max] = [24,35] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Lamelles caudales	[min,max] = [7,10] mean = null sd = null [uMethLower,uMethUpper] = [null,null]	• Pointue	• Environ la longueur totale des lamelles	Unknown	Unknown	• Absente
Lestes sponsa	[min,max] = [24,31] mean = null sd = null	• Lamelles caudales	[min,max] = [8,8] mean = null sd = null	• Arrondie • Pointue	Unknown	Unknown	Unknown	• Absente

Figure 14. Capture d'écran de la "matrice de comparaison" d'une base de connaissances sur Xper³. Les items sont en ordonnées, les descripteurs en abscisses, et les états de descripteurs valides sont au croisement de chaque item et descripteur.

Une fois la base de connaissances satisfaisante il est possible de générer via l'onglet "identification", une clé d'identification à accès simple qui optimise le nombre d'étapes nécessaires à la description de spécimens, ou une clé d'identification à accès libre telle que celles que j'ai construites au cours de cette thèse (**Figure 15**).

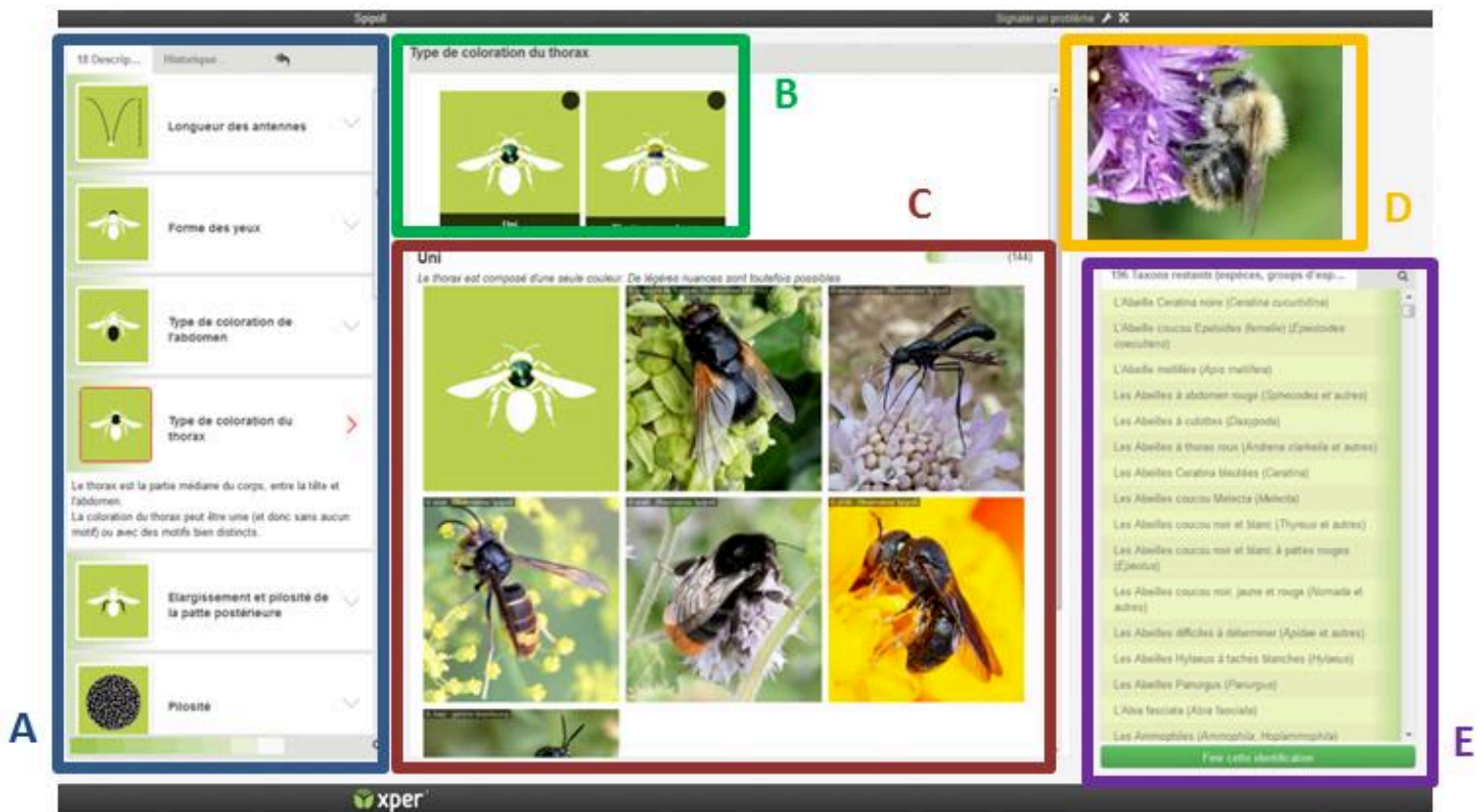


Figure 15. Capture d'écran de la clé d'identification "Insectes" pour le programme Spipoll réalisée avec Xper³.

- A. Liste des descripteurs pouvant être utilisés pour décrire et identifier le spécimen photographié
- B. Propositions d'états de descripteur correspondant au descripteur sélectionné à gauche
- C. Des images pour illustrer l'état de descripteur sélectionné.
- D. Photographie prise par l'utilisateur, représentant l'insecte à identifier
- E. Liste des taxons (items) restant compatibles avec les informations fournies par l'utilisateur.

3.2 Méthodologie pour la construction de clés d'identification numériques sous Xper³

Cette sous-partie traitera en détails des méthodes que j'ai utilisées pour la construction des clés T1 à T7, ainsi que du choix des descripteurs et états de descripteurs pour les bases de données sous Xper³.

3.2.1 Modèle descriptif

Comme présenté dans la partie 2.2, la méthode de construction d'une clé d'identification à accès libre diffère largement de la construction d'une clé à accès simple. Là où pour ces dernières, la principale difficulté réside dans le fait de trouver les cheminements de descripteurs les plus simples, directs et discriminants pour les taxons inclus dans la clé, les clés à accès libre ont à la fois les avantages et les inconvénients du fait de pouvoir inclure autant de descripteurs que souhaité. Il n'y a pas d'ordre imposé pour les descripteurs : l'utilisateur "rentre dans la clé" par le descripteur de son choix, et poursuit l'identification avec le descripteur suivant de son choix jusqu'à la fin de la détermination. La possibilité de créer des descripteurs-fils est également inexistantes dans les clés à accès simple. Pour optimiser l'efficacité de l'outil d'identification et la précision des déterminations, la base de connaissances doit être la plus complète possible, c'est-à-dire qu'il faut limiter les descriptions "inconnues".

Pour la construction des clés d'identification numériques à accès libre sous Xper³, j'ai dû faire des choix, qui me semblaient les plus cohérents par rapport aux outils dont je disposais, aux groupes d'insectes traités, et aux publics ciblées. Ces choix sont laissés à l'appréciation du créateur des bases de données et peuvent grandement varier selon les situations. Dans cette partie, je présente et argumente la méthodologie que j'ai utilisée, qui résulte de mon expérience au cours des années passées sur mon sujet de thèse. Cette expérience a été acquise à partir de ma propre utilisation du logiciel et *via* des échanges avec diverses personnes telles que des étudiants, le personnel de la Direction régionale et

interdépartementale de l'environnement et de l'énergie (DRIEE Ile-de-France), les coordinateurs du programme Spipoll, et des spécialistes des clés d'identification.

Pour rappel, les clés d'identifications T1 à T7 sont destinées à une utilisation par **différents types de publics adultes, sur un large spectre qui inclut des experts des taxons comme des utilisateurs novices en entomologie**. D'autre part, je souhaitais que ces clés incluent **tous les espèces (ou genres)** contenus dans les groupes concernés par chaque clé. Enfin, l'objectif était que les clés d'identification puissent **être utilisées directement sur le terrain**, en tirant avantage du fait que Xper³ est disponible en ligne et accessible depuis tout smartphone doté d'une connexion internet.

Mes critères pour le **modèle descriptif** de toutes les bases de données étaient donc les suivants :

- N'inclure comme descripteurs que des **structures morpho-anatomiques visibles à l'œil nu ou avec une loupe** (grossissement x20 maximum)
- Éviter toute description qui nécessiterait une dissection : j'ai donc par exemple **exclu la description des pièces génitales et de certaines pièces buccales**
- **Simplifier le vocabulaire** utilisé pour les descripteurs et états et le rendre accessible à tous les publics
- Choisir des critères de structures **dépourvus de toute ambiguïté et non variables**, ou des structures qui ne sont pas évidentes à décrire clairement (la plupart des insectes aquatiques)
- **Éviter le chevauchement entre les états de descripteurs**, par exemple "longueur comprise entre 1 et 4 mm" et "longueur supérieure à 2 mm" pour le même descripteur
- Au niveau des **descripteurs numériques**, **prendre le spectre le plus large possible** par rapport aux données extraites des ressources bibliographiques, par exemple inclure une large palette pour les tailles et longueurs possible des individus ou des structures morphologiques. Ainsi, les valeurs extrêmes ne sont pas mises de côté, ce qui réduit le risque de mauvaise détermination du spécimen.

Dans quelques cas, la description au rang le plus précis (espèce, forme, sexe...) n'est pas possible. J'ai fait le choix de laisser éventuellement un doute entre plusieurs taxons proches dans la description finale (par exemple, deux ou trois espèces d'un même genre), plutôt que d'ajouter des descripteurs morpho-anatomiques difficile à observer, aux formes ambiguës ou variables, nécessitant une loupe à fort grossissement ou des dissections.

Pour les bases de données des **larves (T1, T2 et T4)**, mon choix a été de **multiplier les descripteurs** afin d'offrir le plus de choix possible à l'utilisateur, sachant que les larves sont de petits organismes fragiles qui sont souvent endommagés avant, pendant ou après la récolte. Il n'est pas rare qu'il leur manque des branchies, des structures caudales, des pattes, des antennes... Plus la matrice contient de descripteurs, moins l'utilisateur risque d'être limité à des descripteurs correspondant à des structures morpho-anatomiques qui ne sont plus présentes – ou dégradées – sur le spécimen à déterminer. De même j'ai préféré **réduire au maximum le nombre d'états possible pour chaque descripteur**, afin de simplifier le choix à l'utilisateur des clés d'identification et de limiter le risque d'erreur dans la description. C'est pourquoi dans les bases de données T1, T2 et T4, le nombre d'états de descripteur est compris entre 2 et 4 seulement (**Tableau 9**, partie 4.1.1).

Les organismes **imagos** (bases de données T3, T5, T6 et T7) sont moins fragiles que les larves ; il est plus rare que certaines structures soient manquantes et/ou difficiles à observer. Dans ces cas, le choix a été fait de réduire le nombre de descripteurs pour simplifier la matrice – sachant que de manière générale, le nombre de descripteurs nécessaires à la discrimination des items a tendance à augmenter avec le nombre d'items inclus dans la base de connaissances. La réduction du nombre de descripteurs s'est accompagnée d'une augmentation naturelle du nombre d'états pour chaque descripteur (jusqu'à 7 pour la clé T7).

Lorsque les espèces présentaient un dimorphisme sexuel (imagos des éphémères et plécoptères), et que leur description n'était donc pas la même entre le mâle et la femelle, j'ai fait le choix de traiter chaque sexe comme un item. C'est pourquoi la clé T3 des éphémères imagos comporte 78 items pour 39 espèces, et que la clé T5 des plécoptères imagos comporte

20 items pour 10 espèces. Les formes de coloration au sein des espèces de coccinelles ont été traitées de la même manière : chaque forme représente un item (**Tableau 8**, partie 3.3).

En ce qui concerne les **compléments d'information** en dehors du modèle descriptif en lui-même, j'ai fait les choix suivants pour toutes les clés d'identification :

- **Intégrer un maximum de textes explicatifs** pour une compréhension optimale des descripteurs et états et une réduction du risque d'erreur au cours de l'identification
- Inclure des **illustrations claires et informatives**, notamment pour les descripteurs et états les plus difficiles à appréhender, afin de guider l'utilisateur dans sa démarche de détermination
- Proposer pour chaque item une **fiche espèce** contenant des informations supplémentaires : photo du taxon, classification, préférences écologiques, distribution en France métropolitaine, statut de conservation, références bibliographiques utilisées...

Le développement des clés des insectes aquatiques pour les ordres des odonates, éphémères et plécoptères, s'est faite en **collaboration avec des experts de chaque groupe, pour une co-construction des outils d'identification** et la réduction du risque d'erreurs lors du codage de la matrice items-descripteurs-états. Il s'agissait soit d'une co-création directe de la clé (Michel Brulin pour les larves d'éphémères), soit d'une révision précise du travail préliminaire que j'avais fait afin de corriger les erreurs, améliorer l'organisation de la matrice de données, compléter les informations... (Guillaume Doucet, Jacques Le Doaré et Alexandre Ruffoni). La liste complète des experts est disponible dans le **Tableau 7**. La construction des clés des insectes aquatiques imagos (T3, T5 et T6) s'est faite dans le cadre du stage de Mathieu Faure-Brac, en troisième année de Licence de Biologie à l'université de Nantes, qui s'est déroulé de mai à août 2016.

3.2.2 Caractères morpho-anatomiques

Pour chaque base de connaissances, j'ai défini la liste des descripteurs (caractères morpho-anatomiques) et de leurs états principalement d'après des informations issues de guides de terrain, ouvrages naturalistes et clés d'identification existantes. Pour les larves d'éphémères, j'ai pu également m'appuyer sur une collection de référence fournie par l'OPIE, et pour les coccinelles, les collections du MNHN ont représenté une ressource précieuse. Les principales références utilisées sont listées dans le **Tableau 7**.

Tableau 7. Références bibliographiques utilisées pour la construction des clés d'identification, noms de leurs constructeurs, et noms des experts ayant contribué à leur validation.

Code base Xper	Nom de la clé d'identification	Co-constructeurs de la base de données (par ordre alphabétique)	Références principales pour la construction de la base de données	Expert(s) ayant validé la base de données (par ordre alphabétique)
T1	Larves d'odonates d'Ile-de-France	Mathilde Delaunay Aurélia Dubois	Nilsson 1996b Doucet 2010 Tachet <i>et al.</i> 2010 Brochard <i>et al.</i> 2012 Brochard & van der Ploeg 2014 Grand <i>et al.</i> 2014 Base de données Perla	Guillaume Doucet
T2	Larves d'éphémères d'Ile-de-France	Michel Brulin Paola Campos Mathilde Delaunay	Ide 1935 Nilsson 1996a Macadam & Bennett 2010 Tachet <i>et al.</i> 2010 Bauernfeind & Soldán 2012 Base de données Perla	Michel Brulin
T3	Ephémères imagos d'Ile-de-France	Michel Brulin Mathilde Delaunay Mathieu Faure-Brac	Manach 1984 Nilsson 1996a Macadam & Bennett 2010 Tachet <i>et al.</i> 2010 Bauernfeind & Soldán 2012	Michel Brulin
T4	Larves de plécoptères d'Ile-de-France	Mathilde Delaunay	Hynes 1977 Nilsson 1996a Pryce <i>et al.</i> 2007 Tachet <i>et al.</i> 2010 Base de données Perla	Jacques Le Doaré Alexandre Ruffoni
T5	Plécoptères imagos d'Ile-de-France	Mathilde Delaunay Mathieu Faure-Brac	Hynes 1977 Nilsson 1996a Pryce <i>et al.</i> 2007 Tachet <i>et al.</i> 2010	Jacques Le Doaré Alexandre Ruffoni
T6	Coléoptères imagos d'Ile-de-France	Mathilde Delaunay Mathieu Faure-Brac Séverin Jouveau	Bertrand 1928 Nilsson 1996a Hájek 2009 Tachet <i>et al.</i> 2010 Base de données Perla	
T7	Coccinelles imagos de France métropolitaine	Mathilde Delaunay Séverin Jouveau Romain Nattier	Iablokoff-Khnzorian 1982 Baugnée & Branquart 2000 Roy <i>et al.</i> 2013 Derolez <i>et al.</i> 2014	

A partir de ces documents, et d'autres références isolées pour préciser certaines identifications, j'ai combiné tous les descripteurs et états de descripteurs qui correspondaient à mes critères pour la construction des clés d'identification des insectes sous Xper³, selon la méthode présentée au 3.2.1. Il est arrivé que plusieurs états puissent s'appliquer pour un item et un descripteur donné ; j'ai alors codé plusieurs états de descripteur, comme le permet Xper³.

La liste complète des descripteurs et descripteurs-fils inclus dans les bases de données et clés d'identification des insectes aquatiques et des coccinelles peut être consultée dans les **Annexes 2 à 8**.

3.3 Construction de clés Xper³ pour les insectes dans le cadre de la thèse

Cette partie est consacrée à la présentation des 7 bases de données qui ont été construites au cours de ce travail de thèse à l'aide du logiciel Xper³. Ces bases de données ont été publiées en ligne sous la forme de clés d'identification numériques de différents groupes d'insectes de France métropolitaine, accessibles librement en ligne à tous les publics qui le souhaitent.

Le **Tableau 8** fait la synthèse des principales caractéristiques de chaque clé d'identification réalisée. La clé d'identification des coccinelles de France est accessible en anglais car elle a fait l'objet d'une publication dans un journal international à comité de lecture (**Annexe 1**). Les autres clés d'identification ont été traduites en anglais ou sont en cours de traduction, en vue d'une publication du même type (voir **Annexe 9**).

Clés d'identification Xper3 réalisées pendant la thèse

Tableau 8. Zone géographique et rang taxonomique des items, nombre d'items, nombres de genres et d'espèces concernées, et langue de chaque clé d'identification réalisée sous Xper³ pendant la thèse.

Code base Xper ³	Nom de la clé d'identification	Zone géographique	Rang taxonomique	Nombre de genres concernés	Nombre d'espèces concernées	Nombre d'items	Langue
T1	Larves d'odonates	Région Ile-de-France	Espèce	28	59	59	Français
T2	Larves d'éphémères	Région Ile-de-France	Espèce	21	39	39	Français
T3	Ephémères imagos	Région Ile-de-France	Espèce, mâle et femelle	21	39	78 (39 ♂+39 ♀)	Français
T4	Larves de plécoptères	Région Ile-de-France	Espèce	8	10	10	Français
T5	Plécoptères imagos	Région Ile-de-France	Espèce, mâle et femelle	8	10	20 (10 ♂+10 ♀)	Français
T6	Coléoptères imagos	Région Ile-de-France	Genre	68 (dont 30 mono-spécifiques)	245	68	Français
T7	Coccinelles imagos	France métropolitaine	Espèces et formes intraspécifiques	24	47	66 (dont 30 formes de couleurs)	Anglais

Les clés d'identification T1 à T7 sont librement consultables sur une page web dédiée, via le lien web suivant :

infosyslab.fr/?q=fr/node/113

T1 : Larves d'odonates de la région Ile-de-France, à l'espèce



T2 : Larves d'éphémères de la région Ile-de-France, à l'espèce



T3 : Éphémères imagos de la région Ile-de-France, à l'espèce



T4 : Larves de plécoptères de la région Ile-de-France, à l'espèce



T5 : Plécoptères imagos de la région Ile-de-France, à l'espèce

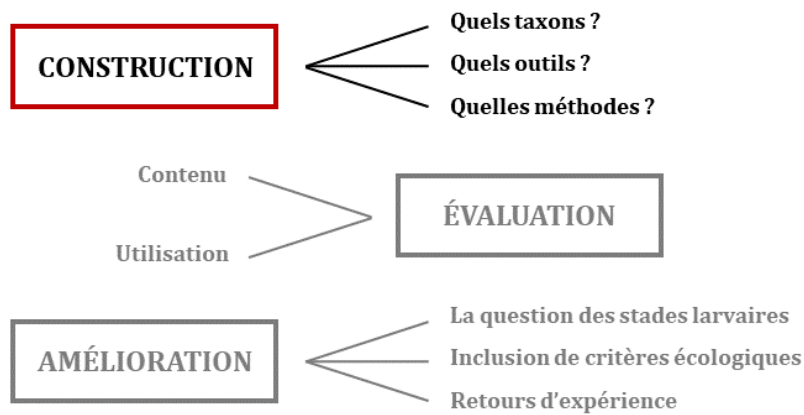


T6 : Coléoptères aquatiques imagos de la région Ile-de-France, au genre



T7 : Coccinelles de France, à l'espèce et formes





Une fois les bases de données pour la description de l'entomofaune et les clés d'identification numériques à accès libre construites, il est intéressant de pouvoir analyser ces dernières. Une analyse en détails permet de mieux comprendre le contenu de ces outils d'identification, comment ils sont utilisés, quels descripteurs ont le meilleur pouvoir discriminant, et comment ces outils peuvent être améliorés et modifiés en fonction des types de publics. Le chapitre 4 portera sur ces questions.

Chapitre 4 – Comment évaluer des clés d'identification numériques ?

Il existe plusieurs critères d'évaluation des clés d'identification numériques à accès libre, parmi lesquels : leur **contenu**, leur **utilisation** (efficacité de l'identification, facilité d'utilisation) et l'analyse de la **perception** que les utilisateurs ont des objets biologiques concernés par les clés d'identification suite à leur utilisation.

Les objectifs sont :

- évaluer l'**efficacité** de la clé en tant qu'outil de détermination des objets biologiques
- tester sa **facilité de prise en main** pour les utilisateurs
- étudier les **possibilités d'amélioration** de la clé d'identification à partir de différentes approches

4.1 Évaluation du contenu des clés d'identification

L'un des avantages des logiciels Xper est l'ensemble des outils qu'ils offrent pour l'analyse du contenu des bases de données de biodiversité. La fonction "checkbase" permet de s'assurer que la matrice de la base de connaissances a été remplie correctement, que toutes les descriptions ont été définies (au moins avec l'option "valeurs inconnues"), que tous les items peuvent être discriminés par la clé d'identification numérique, et qu'aucun item n'a la même description qu'un autre. D'autres fonctions, comme celles qui analysent les descripteurs, permettent d'aller plus loin dans la compréhension de l'outil. Identifier quels descripteurs sont les plus discriminants, par exemple, offre la possibilité de modifier la base de données en conséquence, d'améliorer la clé d'identification, ou encore de changer le "poids" des descripteurs (cet aspect sera discuté dans le chapitre 6).

4.1.1 Contenu général des clés d'identification

La cohérence des bases de données a été vérifiée à l'aide de la fonction "checkbase" de Xper³ : **dans chaque base de connaissances, tous les items sont décrits et aucun item n'a entièrement la même description qu'un autre** (les clés d'identification discriminent chacune tous les items qu'elles contiennent).

Le nombre de descripteurs et d'états de descripteurs contenus dans les clés varie de 12 à 54 descripteurs, et de 24 à 116 états, selon les clés. La liste complète des descripteurs et descripteurs-fils peut être consultée dans les **Annexes 2 à 8**.

Le **Tableau 9** fait la synthèse du contenu des clés d'identifications T1 à T7.

Tableau 9. Tableau récapitulatif du contenu des clés d'identification réalisées ; pour chaque clé : nombre d'items, de descripteurs, de descripteurs-fils, nombre de descripteurs catégoriques (discrets) et numériques (continus), nombre minimal, maximal, moyen et total d'états de descripteurs, nombre d'illustrations pour les items, descripteurs et états de descripteurs, et nombre total d'illustrations.

Code base Xper ³	Nom de la clé d'identification	Nombre d'items	Nombre de descripteurs (dont descripteurs-fils)	Nombre de descripteurs catégoriques/numériques	Nombre d'états de descripteurs (min/max/moyenne)	Nombre d'illustrations pour les items/descripteurs/états (total)
T1	Larves d'odonates	59	54 (12)	52/2	116 (2/4/2,23)	118/54/116 (288)
T2	Larves d'éphémères	39	52 (5)	51/1	115 (2/4/2,25)	77/52/115 (244)
T3	Ephémères imagos	78	25 (9)	24/1	77 (2/7/3,21)	78/25/77 (180)
T4	Larves de plécoptères	10	12 (0)	11/1	24 (2/3/2,18)	10/12/24 (46)
T5	Plécoptères imagos	20	13 (1)	13/0	36 (2/4/2,77)	20/13/36 (69)
T6	Coléoptères imagos	68	22 (4)	21/1	66 (2/7/3,14)	68/22/66 (156)
T7	Coccinelles imagos	66	21 (15)	21/0	98 (2/12/4,67)	66/19/81 (166)

4.1.2 *Pouvoir discriminant des descripteurs*

Il est possible d'obtenir la liste des descripteurs classés par pouvoir discriminant, à l'aide d'une fonction disponible dans le logiciel Xper². Je me suis intéressée aux calculs réalisés avec l'indice Xper, ou "mesure de non compatibilité". Xper² propose deux autres mesures de pouvoir discriminant des descripteurs : l'écart de Jaccard, et l'écart de Sokal et Michener. L'intérêt de l'indice Xper est qu'il trouve plus vite que les deux autres ce qui discrimine complètement les taxons. C'est également la mesure utilisée par défaut par Xper pour calculer le pouvoir discriminant des descripteurs d'une base de connaissance, et classer les descripteurs proposés à l'utilisateur dans les clés (les descripteurs étant listés en ordre

décroissant de leur pouvoir discriminant pour les items restants compatibles avec la description). Pour rappel, avec l'indice Xper le pouvoir discriminant d'un descripteur est compris entre 0 et 1, 1 signifiant que le descripteur permet de discriminer toutes les paires d'items (c'est le rapport du nombre de paires d'items qui se différencient entre eux par ce descripteur sur le nombre total de paires d'items à discriminer). Les **Tableaux 10 à 16** présentent, pour les bases de données T1 à T7, les descripteurs classés par leur pouvoir discriminant (du plus élevé au moins élevé). Par souci de lisibilité, seuls sont montrés les descripteurs ayant un pouvoir discriminant supérieur à 0,33 d'après l'indice Xper (ils discriminent plus du tiers des items de la base).

Tableau 10. Liste des descripteurs de la base de connaissances T1 (larves d'odonates) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Taille générale	1. Longueur du corps (en mm)	0,65
Abdomen	3. Longueur des lamelles caudales	0,53
Abdomen	2. Extrémité de l'abdomen	0,48
Abdomen	14. Forme de l'abdomen	0,48
Abdomen	15. Largeur de l'abdomen	0,48
Masque	42. Type de masque	0,48
Masque	25. Fente médiane à l'extrémité du masque	0,45
Abdomen	22. Épine dorsale sur les tergites abdominaux 4 et/ou 5	0,42
Abdomen	23. Épine dorsale sur le tergite abdominal 7	0,40
Abdomen	4. Extrémité des lamelles caudales	0,34

Tableau 11. Liste des descripteurs de la base de connaissances T2 (larves d'éphémères) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Thorax et abdomen	4. Position des branchies	0,53
Extrémité de l'abdomen	34. Longueur des cerques	0,50
Extrémité de l'abdomen	33. Localisation des soies sur les cerques	0,49
Tête	40. Longueur des antennes	0,44
Thorax et abdomen	8. Nombre de paires de branchies	0,40
Thorax et abdomen	12. Extrémité des branchies	0,40
Thorax et abdomen	15. Angles postérieurs des segments abdominaux	0,34
Extrémité de l'abdomen	35. Taille relative du paracerbe	0,34

Tableau 12. Liste des descripteurs de la base de connaissances T3 (éphémères imagos) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Parties génitales	13. Lobes péniens	0,67
Ailes	1. Ailes postérieures	0,66
Abdomen	25. Coloration (cas des taches)	0,65
Abdomen	23. Coloration de l'abdomen	0,64
Tête	8. Position et taille des yeux composés	0,55
Ailes	6. Nervures intercalaires de l'aile antérieure	0,55
Parties génitales	12. Segmentation des pinces génitales	0,51
Parties génitales	11. Sexe	0,51
Ailes	7. Taille de l'aile antérieure	0,46
Abdomen	21. Paracerque	0,45
Abdomen	24. Type d'ornementation	0,43
Tête	10. Aspect particulier des yeux	0,41
Tête	9. Coloration des yeux	0,38
Ailes	3. Nombre de veines longitudinales (postérieures vestigiales)	0,36
Parties génitales	15. Titillateurs	0,35

Tableau 13. Liste des descripteurs de la base de connaissances T4 (larves de plécoptères) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Abdomen	12. Nombre de segments abdominaux présentant deux parties distinctes (tergite et sternite)	0,71
Thorax	8. Orientation des fourreaux alaires	0,53
Thorax	6. Longueur des articles du tarse	0,47
Thorax	7. Longueur des pattes arrière	0,47
Taille générale	1. Longueur du corps	0,44
Abdomen	10. Protubérances dorsales sur l'abdomen	0,38
Tête	2. Longueur des glosses et paraglosses	0,36
Tête	3. Segment terminal des palpes maxillaires	0,36

Tableau 14. Liste des descripteurs de la base de connaissances T5 (plécoptères imagos) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Abdomen	9. Plaque sous/post génitale	0,72
Thorax	4. Longueur des articles du tarse	0,65
Abdomen	8. Nombre d'articles des cerques	0,64
Tête	3. Forme de la tête	0,44
Thorax	5. Position des ailes au repos	0,44
Abdomen	10. Gonopore	0,35
Thorax	7. Pièce en X	0,35

Tableau 15. Liste des descripteurs de la base de connaissances T6 (coléoptères imagos) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés.

Catégorie	Descripteur	Valeur de l'indice Xper
Aspect général	4. Sillons	0,57
Thorax	12. Motifs du pronotum	0,55
Élytres	19. Motifs des élytres	0,53
Élytres	10. Couleur secondaire des élytres	0,53
Tête	7. Aspect des antennes	0,52
Thorax	14. Couleur principale du pronotum	0,52
Élytres	22. Ponctuation	0,51
Aspect général	3. Ornementation	0,49
Thorax	13. Couleur secondaire du pronotum	0,46
Aspect général	1. Taille de l'organisme	0,46
Élytres	21. Couleur principale des élytres	0,45
Élytres	18. Écusson	0,44
Tête	8. Longueur des antennes	0,44
Tête	9. Rapport des palpes et des antennes	0,41
Aspect général	2. Forme générale du corps	0,35

Tableau 16. Liste des descripteurs de la base de connaissances T7 (coccinelles) les plus discriminants, ainsi que la valeur de leur mesure Xper et la catégorie dans laquelle ils ont été classés (voir **Annexe 1**).

Catégorie	Descripteur	Valeur de l'indice Xper
Pronotum	5. Type of pronotum pattern	0,85
Élytres	8. Number of elytra patterns	0,85
Élytres	10. Number of lateral lines of elytra patterns	0,83
Élytres	11. Number of longitudinal lines of elytra patterns	0,82
Pronotum	4. Number of pronotum patterns	0,71
Élytres	17. Shape of the scutellary spot	0,68
Élytres	2. Elytra main colour (background)	0,53
Élytres	7. Colour of elytra pattern	0,52
Elytra	13. One of the spots reaches the rim of the elytra	0,49
Élytres	18. White marks between the scutellary spot and the elytra basis	0,47
Élytres	16. Scutellary spot	0,46
Élytres	12. A spot in the first third of the elytra	0,40

Les descripteurs les plus discriminants ont tendance à être plus précis et/ou à contenir un nombre plus élevé d'états de caractères : caractères relatifs aux parties génitales ("Lobes pénien", "Plaque sous/post génitale" ...), à des ornements ou des couleurs (coloration des taches ou de l'abdomen, motifs du pronotum ou des élytres, "Sillons" ...), et caractères s'attachant à des longueurs ou des dénombrements (longueurs du corps, des lamelles caudales, des cerques ou des antennes, nombre de segments ou d'articles...).

Malheureusement, les caractères les plus fins (structures morphologiques plus petites) sont également plus difficiles à appréhender par l'utilisateur. Il existe donc un risque accru que les utilisateurs fassent des erreurs dans les identifications.

4.2 Évaluation de l'utilisation des clés d'identification

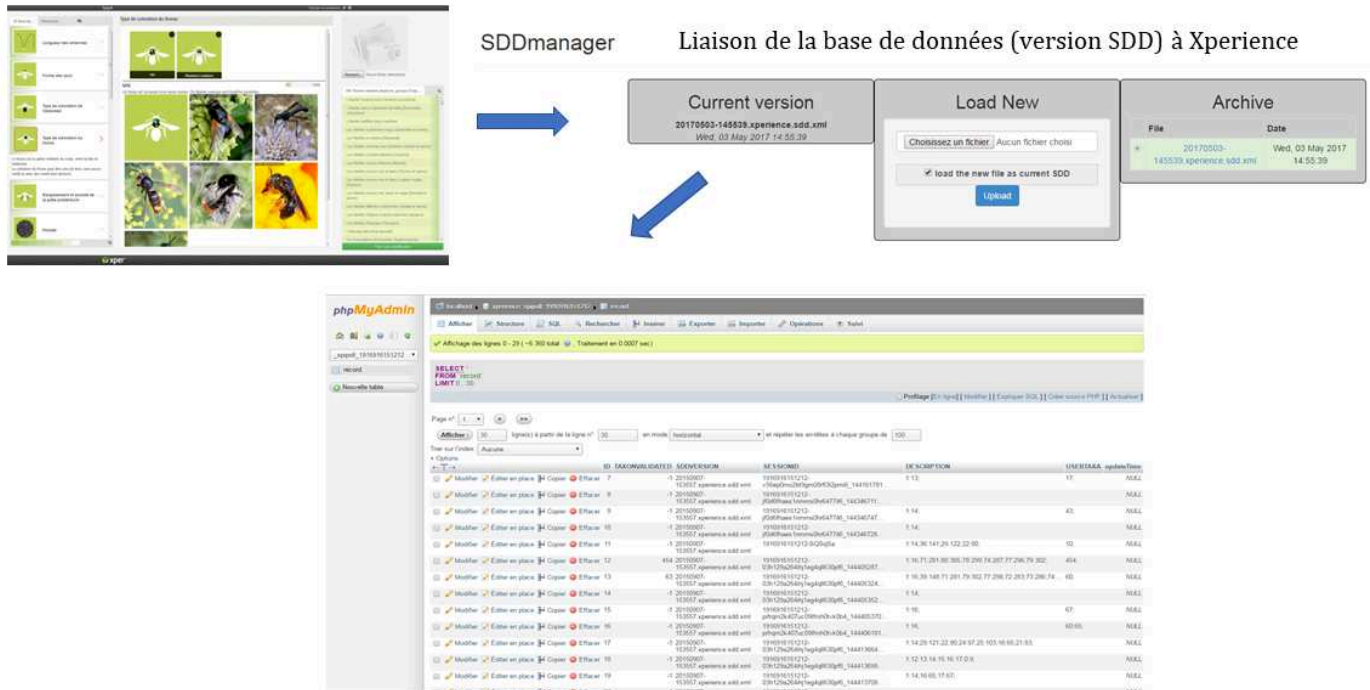
Un autre type d'évaluation des clés qui donne des informations en vue de leur amélioration, est l'analyse de l'utilisation des clés d'identification Xper³ par les utilisateurs. L'analyse de l'utilisation peut se faire par l'étude du retour d'expérience des utilisateurs (qui donnent leur avis sur la qualité de l'outil d'identification et sur les difficultés qu'ils ont éventuellement rencontrées), ou bien par **l'analyse approfondie de toutes les étapes qu'ils ont suivies lors de leurs démarches d'identification**. Un système récemment développé, appelé **Xperience**, permet d'enregistrer toutes les actions effectuées par les utilisateurs des clés Xper³ et d'exporter les données sous forme de tableur pour les exploiter. Au cours de ma thèse j'ai eu l'occasion de tester ce système. J'ai choisi d'exploiter tout d'abord les données extraites à partir d'une **clé d'identification déjà construite, de grande taille**, et développée dans le cadre d'un programme de sciences participatives : le Spipoll. Puis j'ai effectué, avec un panel de participants, un test de l'utilisation d'une des clés construites dans le cadre de mon doctorat : la **clé des larves d'odonates** de la région Ile-de-France (T1).

4.2.1 Le système Xperience

Le système Xperience a été développé par les membres de l'équipe de recherche LIS (Laboratoire informatique et systématique) au sein de l'UMR ISYEB. Lorsqu'on lie une base de connaissances Xper³ à Xperience, ce dernier enregistre toutes les étapes de détermination validées par l'utilisateur au cours de sa démarche d'identification d'un spécimen donné. On obtient alors une nouvelle base de données (base de données Xperience) exploitable sur un tableur de type Excel. Chaque choix de descripteurs et d'états de descripteurs est enregistré, dans l'ordre exact de validation par l'utilisateur. Il en est de même pour "l'identification finale", c'est-à-dire le(s) taxon(s) choisi(s) comme étant correct(s) par l'utilisateur (**Figure**

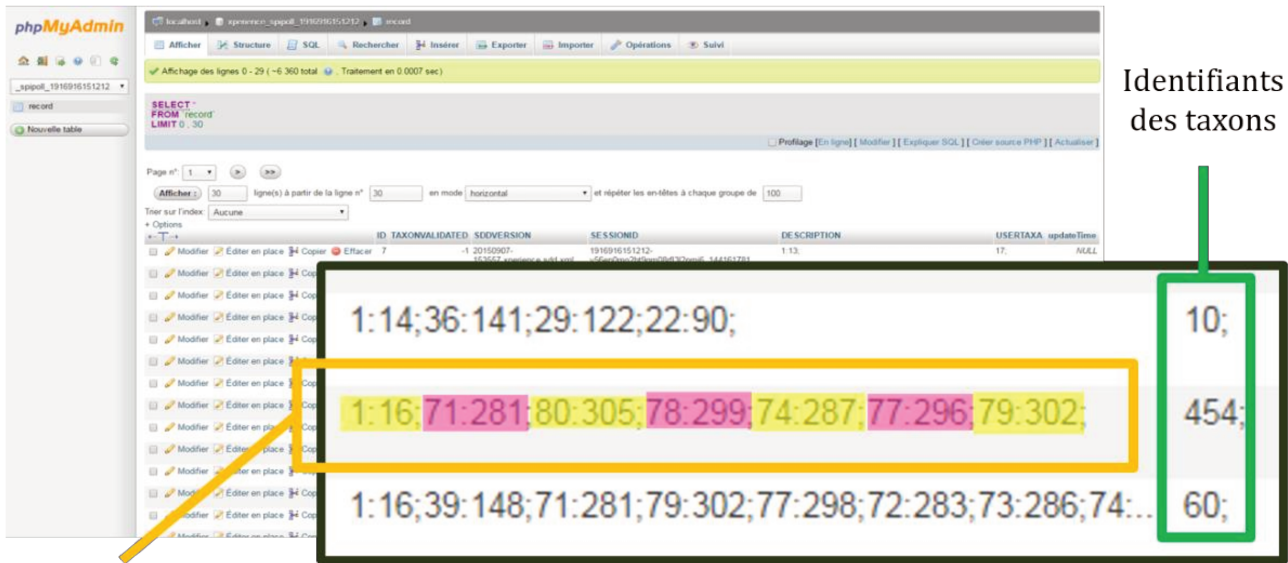
16). Chaque item, descripteur et état de descripteur est associé à un numéro spécifique (code) pour Xperience, permettant l'analyse des chemins d'identification (**Figure 17**).

Base de données / clé d'identification Xper³



Base de données Xperience

Figure 16. Fonctionnement du système Xperience. La base de connaissances est exportée au format SDD (“Structured Descriptive Data” – Hagedorn *et al.* 2005), le format de fichier pour les bases Xper³. Elle est ensuite liée au gestionnaire SDD de Xperience. Après cette étape, toutes les actions effectuées par les utilisateurs de la clé d'identification correspondant à la base de connaissances Xper³ sont enregistrées dans Xperience, et présentées sous forme de tableau pour être analysées.



1 chemin, 7 étapes

Base de données Xperience

Figure 17. Base de données Xperience pour la clé d'identification du Spipoll. Cet exemple présente un chemin d'identification à 7 étapes. La première action de l'utilisateur a été de sélectionner, pour le descripteur #1, l'état de descripteur #16. Son action suivante a été de sélectionner le descripteur #71 et l'état #281. Et ainsi de suite jusqu'à son identification finale, où l'utilisateur a choisi l'item #454 comme étant le taxon qui d'après lui, correspond au spécimen présent sur la photo et à qu'il souhaite déterminer.

4.2.2 Analyse des démarches d'identification pour la clé Spipoll

4.2.2a La clé d'identification "Insectes" du Spipoll

J'ai d'abord utilisé le système Xperience à partir d'une clé d'identification à laquelle il était déjà connecté depuis plusieurs mois : **la clé d'identification des insectes, construite pour le programme Spipoll** (présenté en détails au 2.2.3). Cette clé numérique interactive à accès libre a été créée à l'aide de Xper³. Elle est accessible en ligne à l'adresse suivante : **spipoll.snv.jussieu.fr/mkey/mkey-spipoll.html**. L'un des avantages de cette clé est qu'elle est abondamment utilisée, car les volontaires du Spipoll (les "Spipolliens") sont nombreux et qu'une partie d'entre eux est très fidèle au programme. De nombreuses données sont donc disponibles.

La clé d'identification "Insectes" du Spipoll repose sur une large base de connaissances comportant **630 items, 85 descripteurs et 318 états de descripteurs**. Dans cette base de données, un item peut représenter une espèce – par exemple "L'Abeille mellifère (*Apis mellifera*)" ou un groupe d'espèces, jusqu'à plusieurs familles regroupées dans un même item – par exemple "Les Apions (Curculionidae, Apionidae)" – selon la précision taxonomique qui peut être atteinte grâce à l'identification par la clé d'après des photographies. L'un des objectifs du programme Spipoll est d'enrichir les connaissances sur l'écologie et la distribution des insectes pollinisateurs et sur le fonctionnement des réseaux de pollinisation. De ce fait, les items représentent des groupes taxonomiques et/ou écologiques pertinents pour résoudre ce type de questions.

Les suites d'étapes d'identification (choix des descripteurs et validation d'un ou plusieurs(s) état(s) par descripteur) des utilisateurs de la clé d'identification sont appelées "chemins d'identification". Les chemins d'identification de la clé "Insectes" du Spipoll sont enregistrés depuis septembre 2015 dans le cadre du projet TRASSE (TRANsmission Sémantique des connaissances de la Systématique vers l'Écologie). Ce projet, porté par le MNHN et Sorbonne Université (SU), s'intéresse à l'analyse de la restitution de la connaissance scientifique et de sa perception par les utilisateurs, au travers de l'exemple de la clé d'identification des insectes du Spipoll, afin d'améliorer de façon dynamique les outils d'aide

à l'identification. Le processus implique les données fournies par les Spipolliens lors de leur utilisation de la clé "Insectes", le système Xperience, ainsi que les gestionnaires du Spipoll et chercheurs les chercheurs qui travaillent sur les données de ce programme de sciences participatives (**Figure 18**).

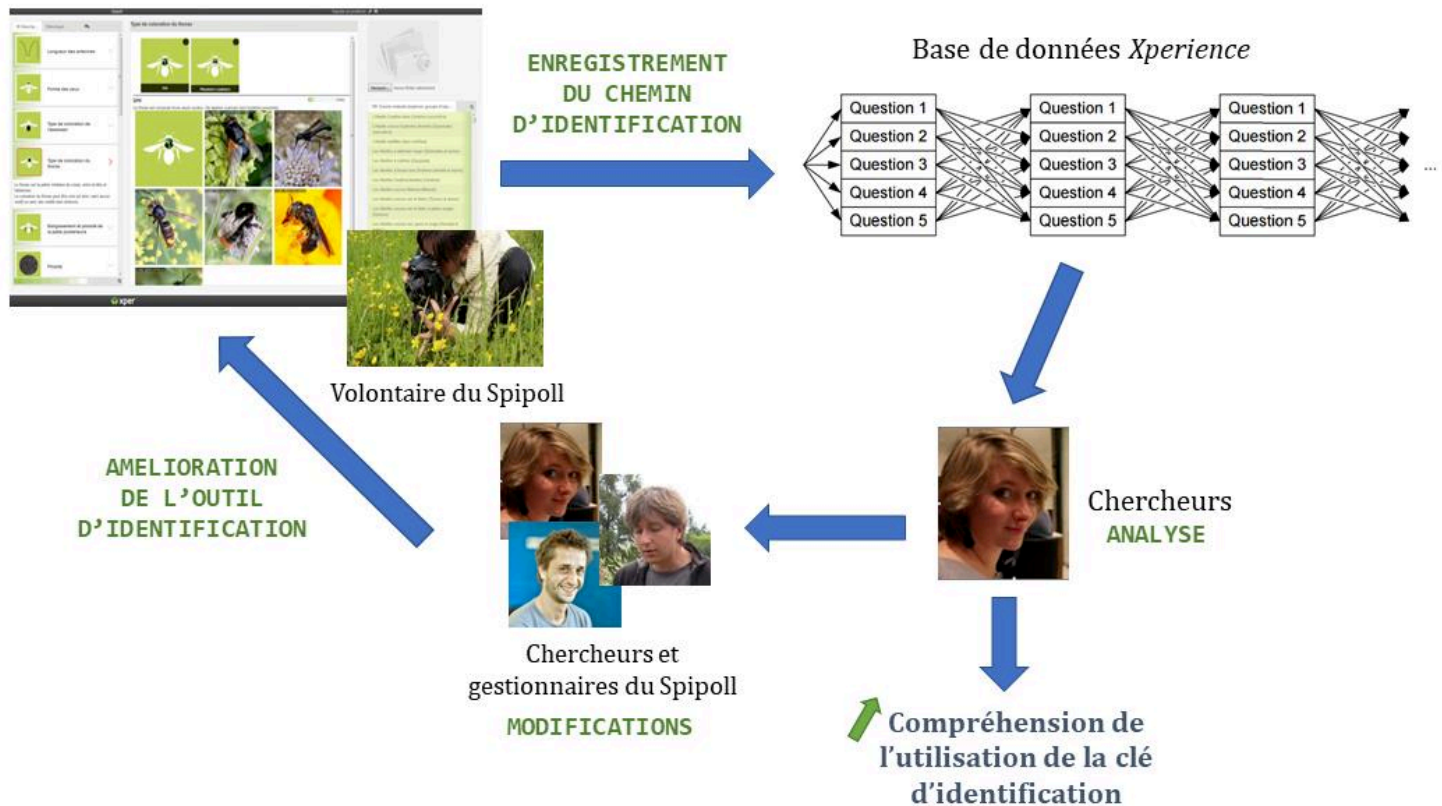


Figure 18. Processus d'extraction et d'analyse des données de la clé d'identification Spipoll, dans le but d'améliorer cet outil d'identification dans le contexte d'un programme de sciences participatives. Les professionnels représentés sont Colin Fontaine, chercheur (MNHN, UMR 7204 CESCO) et Mathieu de Flores (OPIE), gestionnaire, animateur et responsable principal du programme Spipoll.

Pour cette étude, j'ai travaillé à partir de l'enregistrement des chemins d'identification de toutes les utilisations de la clé, sur la **période du 1^{er} septembre 2015 au 31 mai 2016 (8 mois)** ; soit **2114 actions "lancer la clé"** (c'est-à-dire, ouvrir la page web de l'outil d'identification). Cependant, notons que l'action de lancer la clé n'est pas toujours suivie d'une démarche d'identification (l'utilisateur peut refermer la page immédiatement).

Le tableau de données extraites comportait, pour chaque ligne de données dans un tableur Excel, les informations suivantes (dans cet ordre) :

- Code de l'identification (unique)
- Nom du(des) taxon(s) choisi(s) par l'utilisateur à la fin de la démarche d'identification
- Nom du taxon "correct" validé après vérification d'un(e) expert(e) de l'OPIE
- Identifiant (pseudonyme) de l'expert(e)
- Descripteur choisi à l'étape n°1 du chemin d'identification, associé à l'état (ou aux états) de descripteur choisis par l'utilisateur
- Descripteur choisi à l'étape n°2 du chemin d'identification, associé à l'état (ou aux états) de descripteur choisis par l'utilisateur
- Et cætera jusqu'au dernier descripteur choisi dans le chemin d'identification.

Il est important de préciser que comme le système Xperience était nouvellement mis en service, **tous les enregistrements n'ont pu se faire comme voulu**. Notamment, le système n'a enregistré que quelques "identifications finales" associées à leur validation par un expert entomologiste du Spipoll. Je n'ai donc pas pu savoir où les utilisateurs avaient fait des erreurs – que ce soit dans les étapes de détermination ou dans l'identification finale du taxon. C'est pourquoi dans les résultats, je ne présenterai qu'une **analyse des doutes émis par les utilisateurs et pas de leurs erreurs**.

4.2.2b Méthodologie suivie pour l'analyse des démarches d'identification

L'ordre dans lequel l'utilisateur observe et décrit, ses hésitations, ses erreurs éventuelles, sont des informations précieuses permettant d'en savoir davantage sur la compréhension de l'utilisateur sur ce qu'il observe.

Pour traiter ce très grand jeu de données et pouvoir exploiter des résultats cohérents, j'ai tout d'abord divisé les 85 descripteurs de la base de données **en catégories de descripteurs**, selon le type d'élément(s) que l'utilisateur doit observer pour ce descripteur.

A chaque descripteur est assigné un "type de descripteur" et la "partie du corps" de l'insecte concernée par ce descripteur (**Annexe 10**).

9 catégories "type de descripteur" ont été formées :

- "Couleurs" : le descripteur est relatif à la coloration d'une partie ou de la totalité de l'insecte, par exemple "Couleur de l'extrémité de l'abdomen"
- "Détail" : concerne les éléments à observer qui sont de petite taille et peu visibles, en particulier sur des photos ; par exemple "Forme de la cellule marginale" (sur l'aile des hyménoptères). Ces descripteurs peuvent être délicats à appréhender pour l'utilisateur, d'autant plus s'il est novice en entomologie.
- "Forme générale" : on s'intéresse ici à la forme d'une grande partie de l'insecte (ou à la silhouette complète de l'insecte), par exemple "Forme du corps"
- "Forme précise" : le descripteur est relatif à la forme ou la silhouette d'une petite partie ou d'un organe de l'insecte, par exemple "Forme des yeux"
- "Motifs" : on cherche à décrire l'aspect de motifs (colorés ou non) présents sur une partie ou sur la totalité de l'insecte, par exemple "Motifs du dessus des ailes antérieures". Les descripteurs relatifs à la fois à des couleurs et à des motifs (par exemple "Couleurs et motifs des yeux") ont été inclus dans cette catégorie.
- "Présence/Absence" : le descripteur est binaire (il comporte deux états de descripteur) et s'attache à déterminer si une structure est absente ou présente sur l'insecte, par exemple "Présence ou absence d'ailes". Les descripteurs concernant des couleurs ou motifs mais pas des structures ou organes, ont été placés dans les catégories "Couleurs" ou "Motifs" selon le cas.
- "Spatial" : cette catégorie regroupe les trois descripteurs en lien avec des zones géographiques - "Positionnement géographique", "Massif montagneux" et "Localisation Méditerranée"
- "Taille" : le descripteur est relatif à une taille (longueur, épaisseur, envergure...) et peut concerner tout ou partie de l'insecte
- "Temporel" : cette catégorie ne contient que le descripteur "Période d'observation".

De même, 8 **catégories “partie du corps”** ont été formées selon la partie du corps de l’insecte concernée par le descripteur : “Abdomen”, “Ailes”, “Antennes”, “Corps entier” (pour les descripteurs qui s’attachent à plus d’une partie de l’insecte), “Pattes”, “Tête”, “Thorax”, “Yeux”.

J’ai analysé le nombre et le pourcentage de choix des descripteurs, états de descripteurs, taxons finaux choisis, ainsi que les doutes éventuellement exprimés et les erreurs éventuellement faites pour certains taxons, descripteurs et états de descripteurs.

Pour chaque cas, les résultats chiffrés ont été pondérés pour prendre en compte l’importance relative des descripteurs étudiés. Nous avons donc intégré le fait que certains types de descripteurs étaient plus représentés que d’autres dans la clé. Nous avons pondéré les résultats obtenus en tenant compte du nombre de taxons concernés, dans la base de données, par les descripteurs présents dans chaque catégorie de descripteurs. Le calcul a été fait par un simple produit en croix en considérant que chaque catégorie de descripteurs concernait le même nombre de taxons.

4.2.2c Principaux résultats et discussion

❖ Initiation de la démarche d’identification

On s’intéresse ici aux toutes premières étapes des démarches d’identification, et à la façon dont les utilisateurs abordent “l’entrée dans la clé”.

Validation d’une étape au moins

Parmi les 2114 ouvertures de la clé, 1979 ont mené à l’enregistrement d’au moins une réponse et donc, d’un chemin d’identification (93,61%) (**Figure 19**).

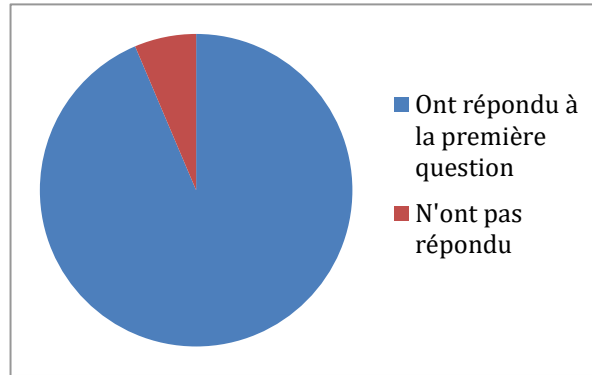


Figure 19. Taux de réponse à au moins une question de la clé d'identification.

L'immense majorité des volontaires du Spipoll qui ouvrent la clé d'identification dans l'objectif de donner un nom au spécimen présent sur la photo répondent au moins à la première question posée par l'outil, à savoir "Quelle est l'allure générale de votre spécimen à identifier ?", présentée comme sur la **Figure 20**. Environ 6% des utilisateurs ferment l'outil avant d'avoir donné une proposition (état de caractère), peut-être parce qu'ils sont rebutés par le design ou la présentation de la clé d'identification, ou alors parce qu'ils ont cliqué par simple curiosité sur le lien pour ouvrir l'outil, ou encore peut-être parce qu'ils ne savent pas répondre à cette première question (bien que cela soit peu probable au vu de sa simplicité).

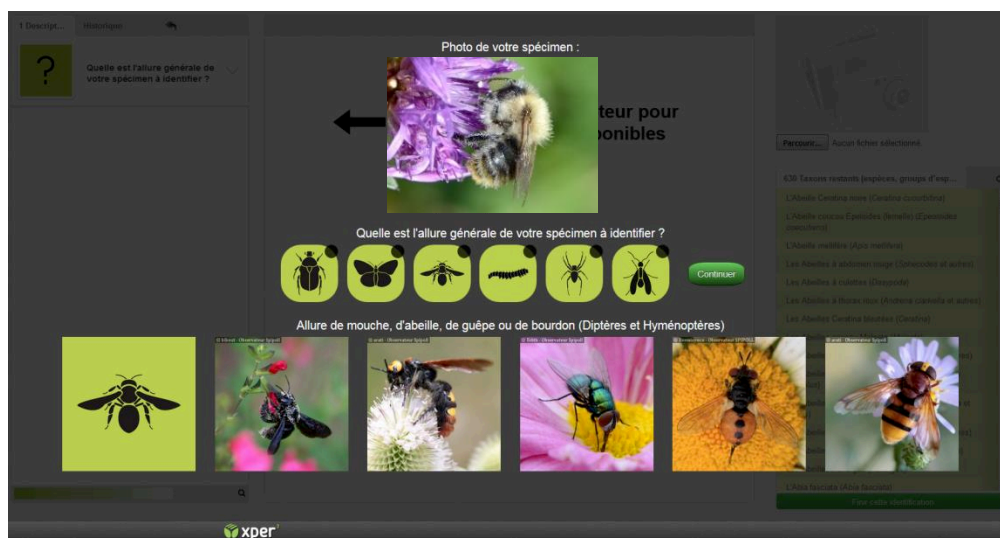


Figure 20. Capture d'écran présentant la première étape d'identification proposée par la clé "Insectes" du Spipoll. En haut se situe la photographie d'un insecte posé sur une fleur, prise sur le terrain par le participant et que ce dernier souhaite identifier grâce à la clé.

Il est intéressant ensuite d'analyser les réponses à cette première question, qui diffère des suivantes par son design particulier et sert de premier "tri" dans l'ensemble des taxons possibles pour l'identification de l'insecte photographié.

Pour quels ordres d'insectes les participants utilisent le plus la clé ?

Les réponses à la première question de la clé ("Quelle est l'allure générale de votre spécimen à identifier ?") se répartissent comme suit (**Tableau 17**) :

Tableau 17. Nombre de choix pour les grands taxons de la clé d'identification

État de caractère	Nombre de choix de l'état
Allure de scarabée ou de punaise (Coléoptères, Hémiptères)	416
Allure de papillon (Lépidoptères)	112
Allure de mouche, d'abeille, de guêpe ou de bourdon (Diptères et Hyménoptères)	1265
Allure de chenille ou d'autre larve	30
Allure d'araignée (Arachnides)	152
Autres aspects	105

Parmi les réponses à la première question (allure générale de l'insecte, pouvant être assimilée à l'ordre), les diptères et hyménoptères représentent presque deux tiers des choix (63,92%). Ils sont suivis par le groupe "coléoptères + hémiptères" (21,02%). Les réponses correspondant à des organismes qui s'apparentent à des chenilles ou à des larves sont minoritaires (1,52%) (**Figure 21**).

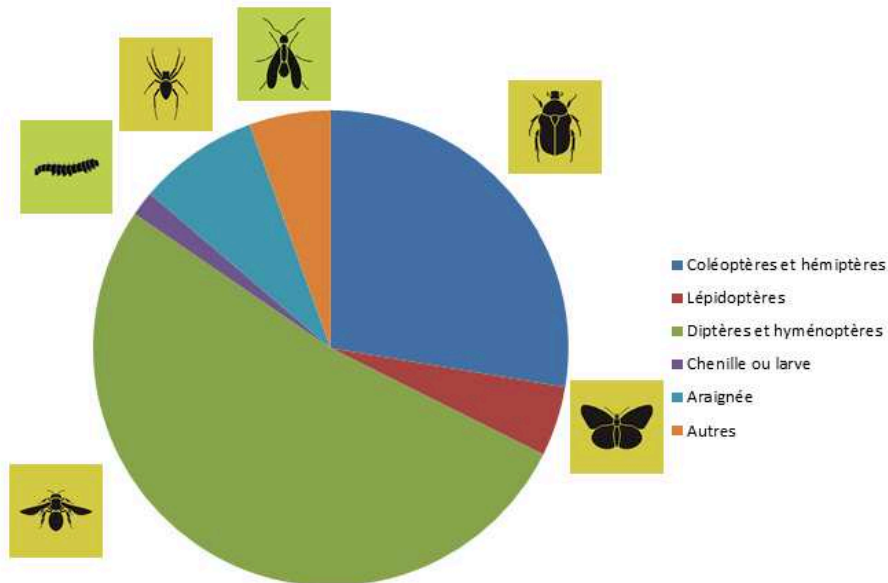


Figure 21. Pourcentages de choix pour les grands taxons de la clé d’identification (première question de la clé).
Silhouettes : spipoll.org.

Les insectes les plus représentés parmi les utilisations de la clé d’identification sont des diptères ou des hyménoptères (“Allure de mouche, d’abeille, de guêpe ou de bourdon”). Ce résultat s’explique notamment par l’importance taxonomique de ces deux ordres – 15 000 espèces en France métropolitaine, pour environ 35 000 espèces d’insectes tous ordres confondus (Berland & Perrier 1981), et par le fait que ce sont des pollinisateurs particulièrement efficaces (Schemske & Horvitz 1984) donc souvent visibles sur les fleurs. Ce sont de ce fait, les ordres les plus représentés dans les photographies de la base de données Spipoll.

Bien que le protocole Spipoll requière que tous les insectes se posant sur les fleurs soient photographiés, dans la pratique, s’il y a beaucoup d’insectes présents simultanément sur l’espèce de fleur choisie, il est difficile voire impossible de prendre un cliché de tous les insectes si l’observateur est seul pendant sa session. Dans ce cas, l’observateur prend en photo prioritairement les insectes les plus visibles, c’est-à-dire souvent les plus grands et/ou les plus colorés. Cette observation est régulièrement faite par les animateurs du Spipoll lors d’ateliers pour présenter le programme au grand public, ou pendant les “Rencontres nationales du Spipoll” (Mathieu de Flores, comm. pers.). Étant volontaire du Spipoll depuis

plusieurs années, c'est un biais que j'ai pu observer lors de mes propres sessions ou pendant des sessions faites en groupe, par exemple au cours de sorties sur le terrain avec une association naturaliste. Les taxons de petite taille sont plus négligés voire manqués (si leur taille est particulièrement réduite, de l'ordre de quelques millimètres).

Les individus à "allure de chenille ou de larve" sont minoritaires dans les utilisations de la clé. En effet, ce type de morphologie est moins présent sur les fleurs que les autres (individus volants notamment) et de taille plus réduite dans l'ensemble ; ses représentants sont donc moins présents sur les photos des Spipolliens. Il est à noter également que la qualité du matériel photo des Spipolliens peut fortement varier, certains ne permettant pas de prendre des clichés nets d'individus très petits (par exemple, la qualité des clichés pris en "mode macro" des appareils compacts est très éloignée de celle obtenue par l'utilisation d'un "objectif macro" sur un appareil réflex).

Il résulte de ce phénomène que les rôles écologiques des "petits" insectes peuvent être largement sous-évalués, entraînant un biais potentiellement important dans les études concernant l'écologie des insectes pollinisateurs et leur place dans les réseaux d'interaction plantes/pollinisateurs. Si possible, l'idéal serait de tenir compte de ce biais lors de la conception de futurs protocoles de terrain tels que celui utilisé pour le programme Spipoll.

Le paragraphe suivant s'attache à **l'analyse approfondie des chemins d'identification**, étape par étape.

❖ *Choix des descripteurs selon leur catégorie : cas de l'ensemble des identifications*

Ici, nous nous intéressons en premier lieu aux choix des descripteurs effectués au cours des identifications de spécimens, **toutes démarches d'identification confondues** (c'est-à-dire pour l'ensemble des taxons). Un descripteur est considéré "choisi" si l'action de l'utilisateur a mené à sa sélection et à la validation d'au moins un état correspondant à ce descripteur.

Choix des descripteurs selon leur "type"

Le **Tableau 18** représente les choix de descripteurs par les utilisateurs (toutes démarches confondues) selon leur catégorie de "type".

Tableau 18. Nombre de choix des descripteurs selon leur type, pour l'ensemble des démarches d'identification.

"Nombre de choix" : somme de l'ensemble des choix de descripteurs effectués pour chaque catégorie

"Nombre de taxons concernés (somme)" : somme des taxons de la base de données qui peuvent être décrits à l'aide de chaque catégorie de descripteurs

"% de choix après pondération" : proportion des choix de descripteurs effectués pour chaque catégorie, en tenant compte de la variation dans les taxons concernés pour chaque catégorie de descripteurs.

Type de descripteur	Nombre de choix	Nombre de taxons concernés (somme)	% de choix après pondération
"Couleurs"	3011	3400	16,92
"Détail"	105	324	6,19
"Forme générale"	1069	1567	13,03
"Forme précise"	1176	1220	18,42
"Motifs"	1219	1650	14,11
"Présence/absence"	679	912	14,22
"Taille"	1109	1239	17,10

La **Figure 22** représente les pourcentages de choix de descripteurs (toutes démarches confondues) en fonction de leur catégorie de "type".

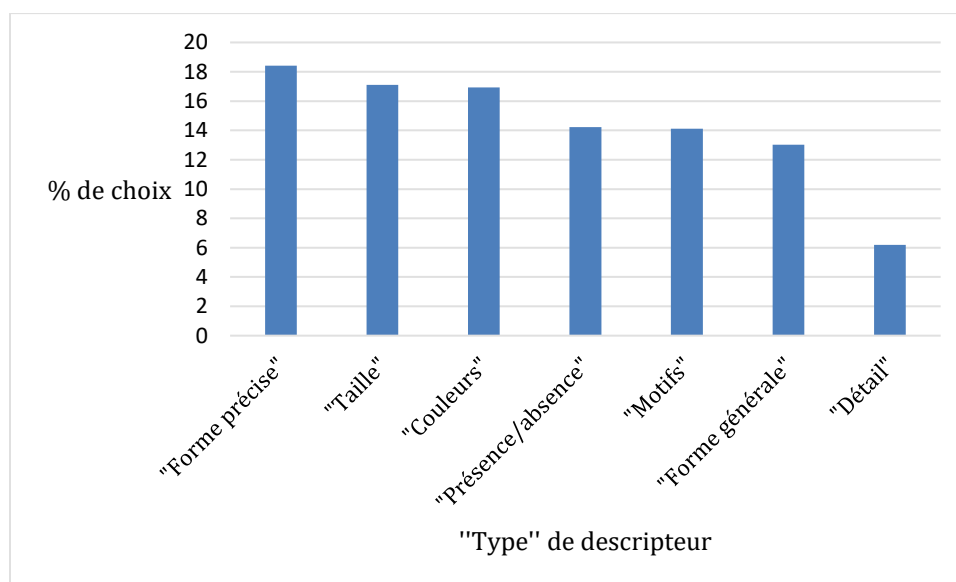


Figure 22. Proportion de choix des descripteurs selon leur type après "pondération".

Les utilisateurs choisissent préférentiellement des descripteurs relatifs à la forme, la taille ou les couleurs. Les descripteurs relatifs aux détails morphologiques sont à l'inverse peu choisis. Ce choix peut s'expliquer par le fait que les notions de forme, de taille et de couleur sont intuitives pour les utilisateurs (par exemple, sélectionner l'état "rouge" quand on doit définir une couleur d'ailes). Leur préférence irait vers des descripteurs plus faciles à appréhender – notamment sur des photographies – et plus faciles à illustrer dans la clé d'identification. Une autre hypothèse (qui renforce à la première) est que les descripteurs de ces trois catégories sont particulièrement discriminants (ils sont efficaces pour distinguer des taxons entre eux), donc ils sont souvent proposés au début des listes de descripteurs et les utilisateurs ont tendance à choisir les descripteurs les plus visibles.

Choix des descripteurs selon la "partie du corps" de l'insecte

La même analyse est faite selon les catégories de descripteurs "partie du corps". Le **Tableau 19** représente les choix de descripteurs par les utilisateurs (toutes démarches confondues) selon leur correspondance aux grandes parties du corps de l'insecte.

Tableau 19 : Nombre de choix des descripteurs selon leur correspondance aux grandes parties du corps de l'insecte, pour l'ensemble des démarches d'identification.

"Nombre de choix" : somme de l'ensemble des choix de descripteurs effectués pour chaque catégorie

"Nombre de taxons concernés (somme)" : somme des taxons de la base de données qui peuvent être décrits à l'aide de chaque catégorie de descripteurs

"% de choix après pondération" : proportion des choix de descripteurs effectués pour chaque catégorie, en tenant compte de la variation dans les taxons concernés pour chaque catégorie de descripteurs.

Partie du corps	Nombre de choix	Nombre de taxons concernés (somme)	% de choix après pondération
Abdomen	1690	906	17,20
Ailes	1499	2856	5,87
Antennes	1156	1337	9,70
Corps entier	1416	2919	6,94
Pattes	426	253	18,89
Tête	194	393	4,98
Thorax	1379	1166	13,68
Yeux	795	392	22,74

La **Figure 23** représente les pourcentages de choix de descripteurs (toutes démarches confondues) en fonction de leur catégorie de “partie du corps”.

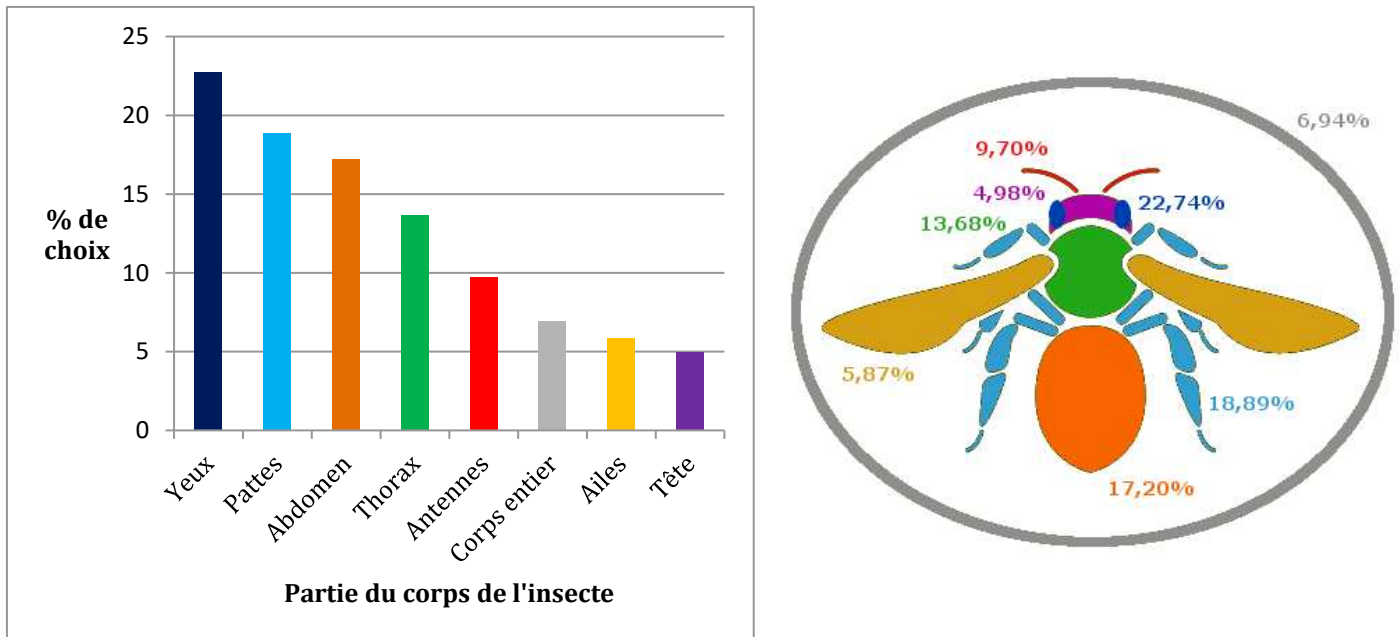


Figure 23. Taux de choix des descripteurs selon la partie du corps après “pondération”.

Les utilisateurs choisissent préférentiellement les descripteurs concernant les yeux. Cela peut s’expliquer par le fait que les seuls 2 descripteurs “yeux” dans la base de données (sur les 81 descripteurs “morpho-anatomiques” considérés pour cette partie) sont assez bien construits et décrits pour être intuitifs pour les utilisateurs. En effet, ils comportent un nombre relativement faible d’états de descripteurs (2 et 4), ce qui rend la description plus facile. D’autre part, le descripteur “Forme des yeux” fait partie des plus aisés à observer et décrire sur des photos d’insectes pollinisateurs, avec ses deux états “Yeux de mouche” ou “Yeux d’abeille, de guêpe ou de bourdon”. Il fait également partie des descripteurs les plus discriminants dans la base de données (rang 14/85, voir **Annexe 11**).

Les insectes sont un groupe extrêmement diversifié, avec des morphologies très variables selon les ordres et même selon les espèces. Naturellement, les descripteurs pertinents pour décrire les spécimens varient grandement selon le grand groupe d’insectes considéré. Le paragraphe suivant s’attache donc à l’analyse des résultats ordre par ordre.

❖ *Choix des descripteurs et doutes*

Certains résultats seront plus spécifiquement analysés pour les deux ordres principaux d'insectes pollinisateurs : hyménoptères et diptères. J'ai également étudié les fréquences de "doute" des utilisateurs au cours de leur démarche d'identification.

Par "doute", on entend que l'utilisateur a validé deux états ou plus pour un même descripteur ; il y a donc hésitation (doute) entre ces différents états de descripteurs. En l'absence de données disponibles assez conséquentes sur les proportions d'erreurs faites au niveau du choix des états de descripteurs, l'analyse des doutes permet cependant d'étudier les facilités et les difficultés que les utilisateurs rencontrent dans leur utilisation de la clé.

Doutes : cas de l'ensemble des identifications

Dans l'ensemble, les utilisateurs doutent peu, malgré la possibilité de sélectionner plus d'un état pour chaque caractère – ce qui est une fonction majeure de Xper³. Sur les 10494 étapes d'identification prises ensemble, seules 522 comportent des doutes (soit 4,97%). Cela peut s'expliquer par le fait que l'interface de la clé d'identification n'indique pas clairement à l'utilisateur qu'il a la possibilité de douter.

Le **Tableau 20** liste les descripteurs entraînant le plus de doute (15% des réponses et plus) auprès des utilisateurs (toutes démarches d'identification confondues).

Tableau 20. Descripteurs qui entraînent le plus de doutes (15% des réponses totales ou plus).

N° du descripteur	Descripteur	Partie du corps	Type	% de doutes
49	Motifs sur le dessus des ailes orange	Ailes	Motifs	57,14
51	Motifs sur le dessus des ailes marron à fauve	Ailes	Motifs	42,86
61	Silhouette de votre papillon "de nuit"	Corps entier	Forme générale	25,00
76	Type de motifs sur l'abdomen	Abdomen	Motifs	21,43
26	Taches claires sur le corps, les pattes, la face	Corps entier	Motifs	18,75
47	Couleur principale du dessus des ailes	Ailes	Couleur	17,86
43	Couleur principale des élytres ou demi-élytres	Ailes	Couleur	16,67

21	Couleurs principales de l'abdomen	Abdomen	Couleur	16,38
52	Couleur principale du dessous des ailes	Ailes	Couleur	15,63
14	Couleur du thorax (Coléoptère, Hémiptère)	Thorax	Couleur	15,38
68	Positionnement géographique	N/A	Spatial	15,00

Dans la majorité des cas (9 sur 11), il s'agit de **descripteurs relatifs à des motifs ou à des patrons de couleurs** ; c'est une première indication sur le type de descripteurs qui pose problème aux utilisateurs pour interpréter les états. Cependant, la plupart des descripteurs du **Tableau 20** ne concernent que certains taxons – par exemple, exclusivement les lépidoptères pour les trois premiers descripteurs du tableau. Je me suis donc penchée sur les **résultats pour deux ordres d'insectes : les hyménoptères et les diptères**. Ces deux ordres, qui incluent la majorité des insectes pollinisateurs, sont également les plus représentés dans la base de données des enregistrements pour la clé Spipoll.

Choix des descripteurs et doutes : cas des hyménoptères

186 chemins d'identification concernent les hyménoptères, pour 1371 étapes d'identification. Les **Figures 24 et 25** sont une représentation schématique des résultats (nombre de choix et pourcentage de doutes) selon chaque catégorie de descripteurs.

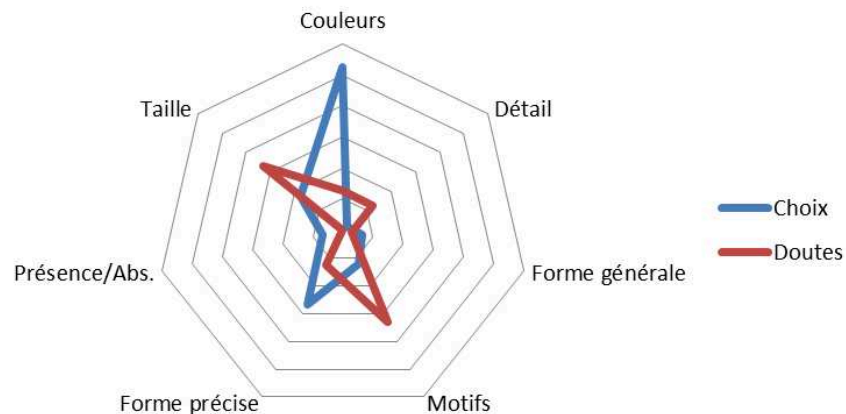


Figure 24. Nombre de choix des descripteurs et pourcentage de doute pour ces mêmes descripteurs selon leur catégorie de "type", pour l'ordre des hyménoptères. En bleu : nombre de choix, avec une échelle comprise entre 0 et 600. En rouge : pourcentage de doutes, avec une échelle comprise entre 0% et 30%.

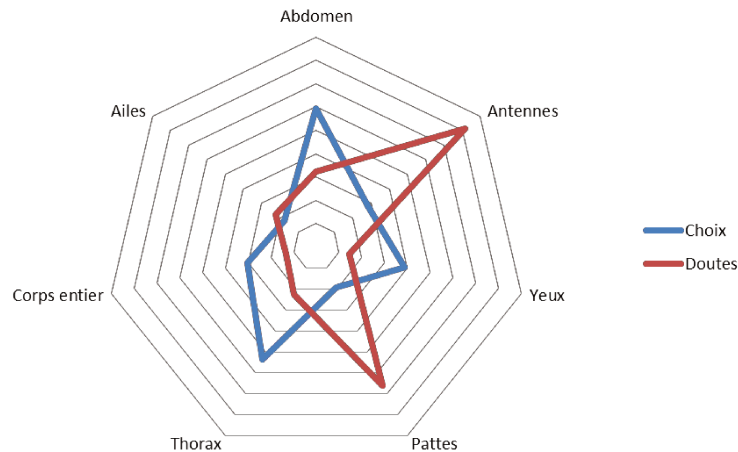


Figure 25. Nombre de choix des descripteurs et pourcentage de doute pour ces mêmes descripteurs selon leur correspondance aux grandes parties du corps de l'insecte, pour l'ordre des hyménoptères. En bleu : nombre de choix, avec une échelle comprise entre 0 et 500. En rouge : pourcentage de doutes, avec une échelle comprise entre 0% et 21%.

Sur la **Figure 24** on voit notamment que **les descripteurs relatifs aux couleurs sont massivement choisis pour les identifications, et qu'ils entraînent peu de doutes** (environ 6%). Cette catégorie représente donc pour les hyménoptères des descripteurs "de qualité", bien construits et expliqués dans la base de connaissances, que les utilisateurs apprécient – probablement parce qu'il est aisé de répondre aux questions qui les concernent et que les illustrations qui les accompagnent sont claires et informatives.

La **Figure 25** montre à l'inverse, deux catégories de **descripteurs qui à la fois sont relativement peu choisies par les utilisateurs, et qui entraînent également beaucoup de doutes** (respectivement 20,42% et 16,49%) : **les descripteurs relatifs aux antennes et aux pattes**. La qualité de ces descripteurs n'est donc pas optimisée dans la clé d'identification. Il serait intéressant de comprendre pourquoi elles entraînent autant de doutes et pourquoi elles semblent rebuter les utilisateurs au cours de l'identification.

Il est important de noter que ces pourcentages de doutes sont dus en grande partie à deux descripteurs particulièrement délicats. Tout d'abord le descripteur "Longueur des antennes", qui contient 3 états de caractères : "Antennes courtes ou à peine visibles", "Antennes de taille moyenne", "Antennes longues à très longues" (**Figure 26**). Si pour l'état

“Antennes longues à très longues” l’outil d’identification précise bien que les antennes doivent “dépasser nettement la longueur tête + thorax”, et que cette grande longueur porte difficilement à confusion, pour les deux autres états la nuance est plus délicate à percevoir, surtout lorsqu’on ne possède qu’une photo sur laquelle l’angle de l’insecte ou d’autres paramètres peuvent faire que les antennes sont en partie cachées ou difficiles à percevoir. Ce descripteur – qui apparaît en premier dans la liste juste après la sélection de l’état de caractère “Allure de mouche ou d’abeille” – est donc logiquement associé à une forte proportion de doutes: 20,57%, quand la moyenne de doutes dans l’ensemble des descripteurs concernant les hyménoptères est de 8,16%.

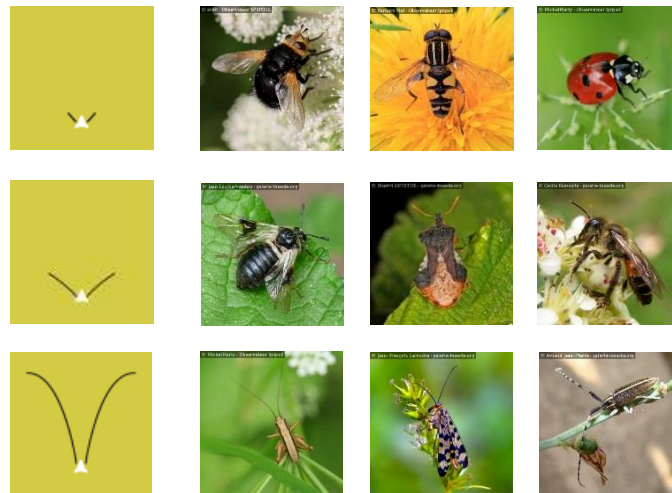


Figure 26. Illustrations du descripteur “Longueur des antennes” dans la clé d’identification © Spipoll. Etats de descripteur, de haut en bas : “Antennes courtes ou à peine visibles” ; “Antennes de taille moyenne” ; “Antennes longues à très longues”.

Le second descripteur posant particulièrement problème est “Elargissement et pilosité de la patte postérieure”, seul descripteur de la catégorie “Pattes” pour les hyménoptères. Ce descripteur, très discriminant au sein des taxons d’hyménoptères, est également difficile à appréhender. Il s’attache à des organes de petite taille (les pattes) et requiert l’observation de nuances précises. Les quatre états de descripteur sont “Elargie, aplatie sur toute la longueur et peu velue”, “Non élargie et velue”, “Non élargie et peu ou pas velue”, “Fémur élargi peu ou pas velu” (**Figure 27**). Sachant que les utilisateurs de la clé n’ont à leur disposition qu’une photographie d’insecte, parfois floue, il n’est pas étonnant que ce

descripteur soit particulièrement difficile à appréhender. Ainsi, il entraîne 16,49% de doutes – pour une moyenne de doutes dans l’ensemble des descripteurs concernant les hyménoptères égale à 8,16%.

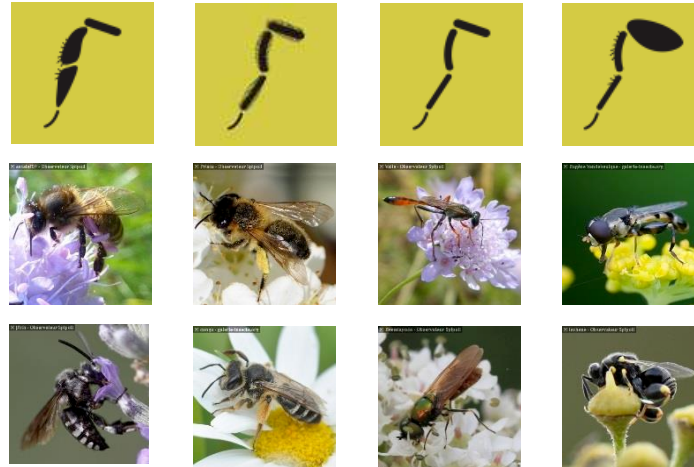


Figure 27. Illustrations du descripteur “Elargissement et pilosité de la patte postérieure” dans la clé d’identification © Spipoll. Etats de descripteur, de gauche à droite : “Elargie, aplatie sur toute la longueur et peu velue” ; “Non élargie et velue” ; “Non élargie et peu ou pas velue” ; “Fémur élargi peu ou pas velu”.

Choix des descripteurs et doutes : cas des diptères

136 chemins d’identification concernent les diptères, pour 870 étapes d’identification. Les **Figures 28 et 29** sont une représentation schématique des résultats (nombre de choix et pourcentage de doutes) selon chaque catégorie de descripteurs.

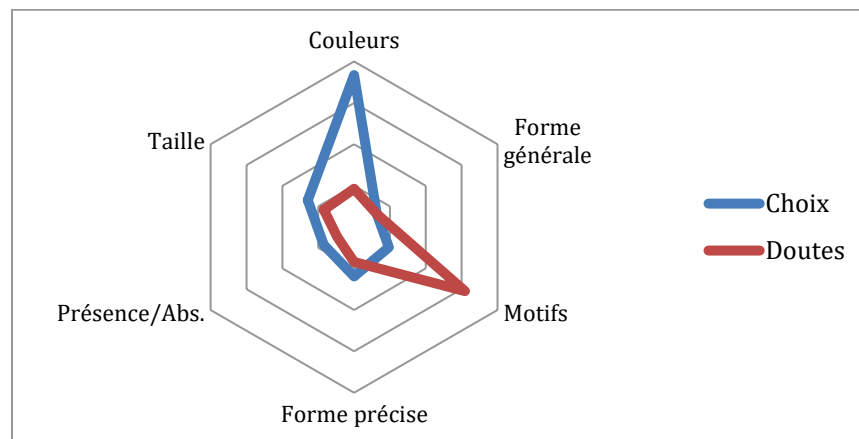


Figure 28. Nombre de choix des descripteurs et pourcentage de doute pour ces mêmes descripteurs selon leur catégorie de “type”, pour l’ordre des diptères. En bleu : nombre de choix, avec une échelle comprise entre 0 et 400. En rouge : pourcentage de doutes, avec une échelle comprise entre 0% et 20%.

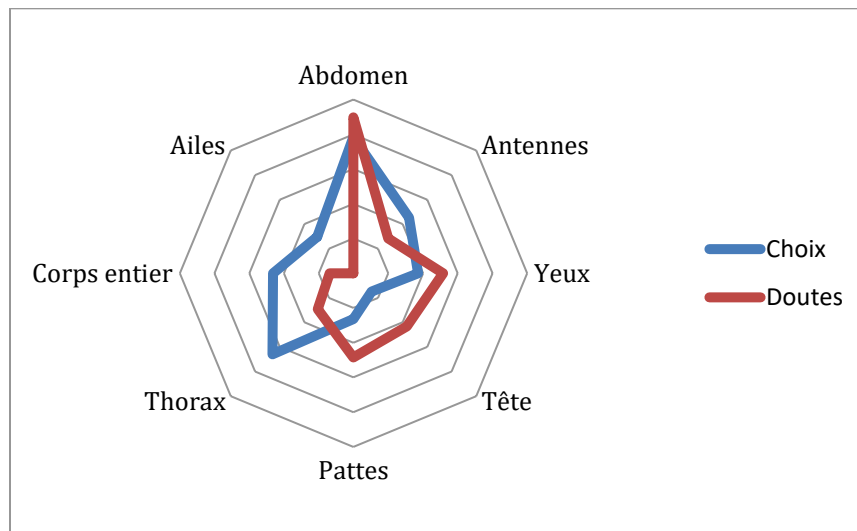


Figure 29. Nombre de choix des descripteurs et pourcentage de doute pour ces mêmes descripteurs selon leur correspondance aux grandes parties du corps de l'insecte, pour l'ordre des diptères. En bleu : nombre de choix, avec une échelle comprise entre 0 et 500. En rouge : pourcentage de doutes, avec une échelle comprise entre 0% et 21%.

Comme chez les hyménoptères, on observe ici de “bonnes” catégories de descripteurs – par exemple ceux qui concernent **les couleurs et le thorax** – qui sont beaucoup choisies par les utilisateurs de la clé tout en entraînant peu de doutes.

Une catégorie de descripteurs est néanmoins plus problématique : la catégorie “Abdomen”. On observe que **ces descripteurs sont les plus fréquemment choisis**, ce qui signifie que les utilisateurs se sentent confiants quant à leur capacité à sélectionner les bons états pour ces descripteurs. **Cependant, on observe également que ces descripteurs relatifs à l'abdomen sont aussi la catégorie “partie du corps” qui entraîne le plus de doutes.** Le fait de douter réduit l'efficacité d'une étape d'identification pour discriminer les espèces restantes, car l'algorithme doit prendre en compte l'assignation d'au moins deux états de caractère. Il est donc important de réduire les pourcentages de doutes – en particulier pour les descripteurs qui sont beaucoup choisis par les utilisateurs – par l'amélioration de leur description et de leur illustration dans la clé.

❖ Doutes et erreurs dans les identifications finales

L'analyse de cette problématique consiste à s'intéresser aux doutes et erreurs observés lors de l'assignation finale, par l'utilisateur, d'une identification (nom de taxon) pour l'insecte situé sur la photographie. Cependant, cette question ne peut être abordée qu'avec prudence pour le moment, en raison du faible nombre de chemins enregistrés dans le jeu de données. En effet, certaines erreurs techniques dues au système d'enregistrement ont fait que beaucoup de taxons finaux, assignés par les volontaires suite à l'utilisation de la clé d'identification, n'ont pu être enregistrés et liés aux chemins d'identification correspondants. D'autre part, seule une petite partie des assignations de taxons avaient pu être corrigées et/ou validées par un expert entomologiste au moment de l'étude. Cependant, nous pouvons analyser les données disponibles pour dégager des tendances.

La **Figure 30** indique la proportion de chemins d'identification qui ont effectivement mené à l'assignation d'un nom de taxon et parmi ces assignations, la proportion qui a été validée ou corrigée par un expert entomologiste du Spipoll.

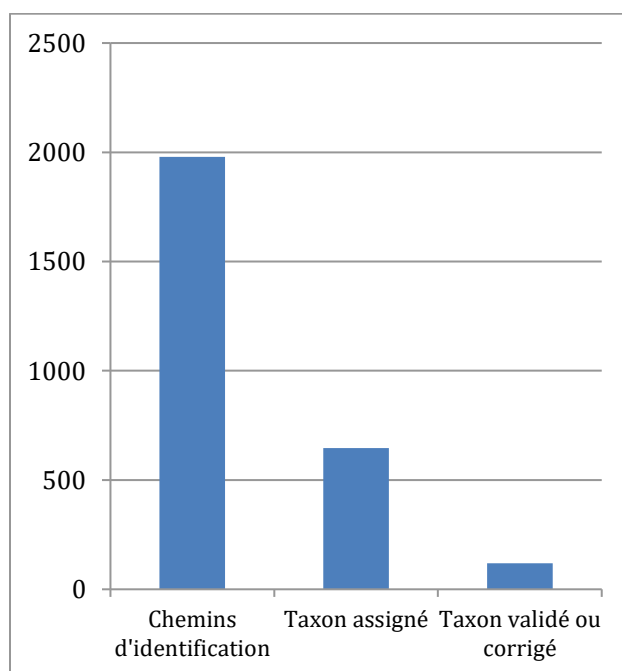


Figure 30. Nombre total de chemins d'identification enregistrés ; nombre de chemins qui ont mené à une assignation de taxon ; nombre d'assignations qui ont été validées ou corrigées par un expert.

Sur les 1979 chemins d'identification, 646 assignations de taxons ont été enregistrées (32,64%). 119 identifications ont été validées ou corrigées par un expert (18,42% des assignations ; 6,01% de l'ensemble des chemins d'identification).

Seul un tiers des chemins d'identification (donc des photos) mène à l'enregistrement d'au moins un taxon associé. Il peut s'agir du dysfonctionnement du système évoqué précédemment, qui empêche parfois l'enregistrement d'un couple "nom de taxon / chemin d'identification correspondant" dans la base de données du site. Il existe également des cas de démarches d'identification avortées : l'utilisateur ne va pas jusqu'à la dernière étape de sélection d'un taxon, il ferme la clé dès qu'il obtient un résultat qui lui paraît adéquat, sans l'envoyer au site. Ou bien encore l'utilisateur n'est pas satisfait de la liste de propositions de taxons car il pense qu'elles ne correspondent pas à son spécimen photographié, donc il n'enregistre pas d'identification finale.

Un peu moins de 20% des assignations de taxons ont été vérifiées par un expert entomologiste. Cela est dû au fort écart entre la quantité de données récoltées pour le Spipoll et les moyens humains disponibles pour la vérification de ces données. Pour l'heure, il est impossible de valider l'ensemble des identifications par vérification de chaque photo.

Doutes dans les assignations de taxons

On peut également regarder la proportion d'assignations finales qui ont mené à des doutes (proposition de deux noms de taxons ou plus) (**Figure 31**).

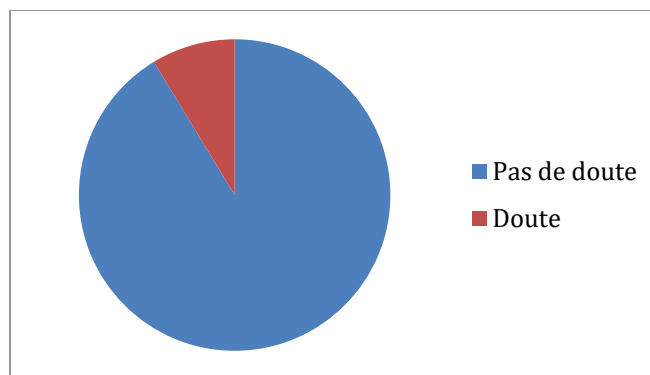


Figure 31. Proportion de doutes dans les assignations de taxons.

Il y a eu 56 doutes dans les assignations de taxons sur un total de 646, soit 8,67%. Lorsque l'utilisateur a proposé deux taxons ou plus et que l'expert a validé un de ces taxons, on ne considère pas que l'utilisateur a fait une erreur – le doute étant permis et prévu dans l'outil.

Le pourcentage de doutes est faible, surtout lorsqu'on prend en compte le fait qu'on observe environ 20% d'erreurs dans la désignation finale des taxons correspondant aux photographies (voir **Figure 32**), un taux non négligeable. Les utilisateurs ont tendance à préférer "se tromper" plutôt qu'à proposer plusieurs taxons qu'ils estiment compatibles. Cela rejoint la "réticence" des utilisateurs à "douter" au cours de la démarche d'identification, observée précédemment.

Erreurs dans les assignations de taxons

Grâce à la validation et à la correction (si nécessaire) des noms de taxons par les experts entomologistes de l'OPIE, nous pouvons étudier les erreurs faites par les utilisateurs lors de l'assignation des identifications finales (**Figure 32**).

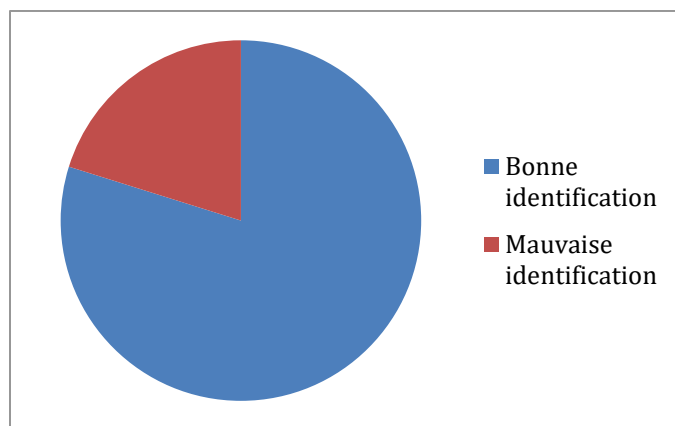


Figure 32. Proportion d'erreurs dans les assignations de taxons.

Parmi les 119 identifications vérifiées par un expert, 24 étaient incorrectes (20,17%).

❖ *Les erreurs des utilisateurs*

Le pourcentage d'erreur dans les assignations de taxons pour la clé Spipoll ($\approx 20\%$) se situe à la limite haute des taux d'erreurs observés en moyenne dans les programmes de sciences participatives : en effet, ces taux d'erreurs sont habituellement compris entre 10% et 20% (Cohn 2008, Bonter & Cooper 2012, Gardiner *et al.* 2012). Il est toutefois important de noter que 20% d'erreurs est considéré comme un taux restant acceptable pour des études en écologie (Cohn 2008).

Pour mieux comprendre où se situent les erreurs – dans les états de descripteurs et dans les taxons finaux assignés – un travail supplémentaire de validation et de correction sera inévitable, en se concentrant sur les identifications déjà effectuées (taxons assignés à un chemin d'identification).

La validité scientifique et l'exactitude des données du programme Spipoll gagneraient à voir ce taux d'erreurs de 20% baisser dans les mois et années à venir. Cela réduirait également la charge de travail des experts validateurs du Spipoll en permettant un maximum d'assignations correctes de taxons en amont.

4.2.3 Test de l'utilisation de la clé des larves d'odonates d'Ile-de-France

J'ai souhaité utiliser le système Xperience avec l'une des clés d'identification réalisées pour le doctorat (larves d'odonates d'Ile-de-France), en utilisant la même méthodologie que pour la clé du Spipoll mais avec un **panel de testeurs**, afin de mieux comprendre comment est utilisé cet outil d'identification et quels problèmes sont éventuellement rencontrés par les utilisateurs.

4.2.3a Échantillonnage taxonomique et méthodologie

Pour tester l'efficacité du système Xperience et analyser l'utilisation d'une des clés réalisées, mon choix s'est porté sur la clé d'identification des larves d'odonates pour deux raisons. La première est que leur identification est la plus aisée comparée aux autres ordres d'insectes aquatiques, notamment à cause de leur relative grande taille. La seconde est qu'il est relativement facile de trouver des populations de larves d'odonates de plusieurs espèces, sur le terrain en Ile-de-France, tout au long de l'année.

Nous avons **collecté des spécimens de larves de trois espèces courantes en Ile-de-France**, dans une mare localisée sur le campus de l'université Paris Saclay, à Orsay (coordonnées GPS : 48,702480 ; 2,171445). Ce nombre de trois espèces a été choisi afin d'avoir plusieurs exemples proposés à chaque testeur de la clé, tout en limitant la durée de la session à des temps raisonnables (entre 15 et 45 minutes selon la rapidité des testeurs). J'ai sélectionné des espèces représentant une **variété taxonomique et morphologique**, et pour chaque espèce le spécimen en meilleur état parmi l'échantillonnage de terrain. Les deux premières espèces étaient *Libellula quadrimaculata* (Anisoptera, Libellulidae) et *Pyrrhosoma nymphula* (Zygoptera, Coenagrionidae) (**Figure 33**). Les deux spécimens étaient des larves matures proches de l'émergence. Le troisième taxon ne pouvait pas être déterminé au niveau spécifique car le spécimen (femelle), pas assez mature (plusieurs stades larvaires avant l'émergence), ne présentait pas toutes les structures morphologiques nécessaires à une identification à l'espèce avec la clé des larves d'odonates d'Ile-de-France. Le spécimen de ce troisième taxon pouvait être attribué avec la clé à trois espèces : *Anax imperator*, *Anax*

parthenope ou *Hemianax ephippiger* (Anisoptera, Aeshnidae) (**Figure 33**). Une identification complète et correcte par l'utilisateur devait donc mener à la proposition de ces trois espèces restantes dans la liste des items compatibles dans Xper³. Cela s'accorde avec la réalité du terrain : il arrive souvent que les larves ne soient pas assez âgées pour pouvoir être identifiées au niveau de l'espèce avec des caractères morpho-anatomiques uniquement.

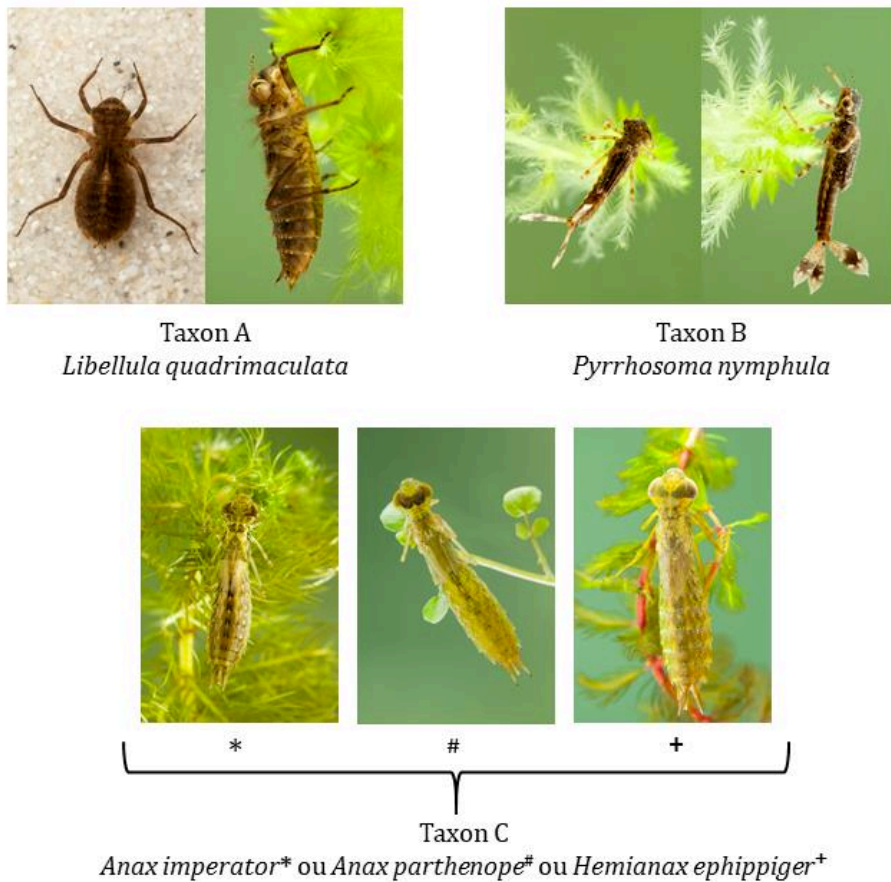


Figure 33. Nymphes des espèces correspondant aux trois spécimens présentés aux testeurs de la clé d'identification des larves d'odonates d'Ile-de-France. Même au stade nymphal, ces trois espèces sont similaires morphologiquement. Dans notre clé d'identification, *H. ephippiger* se distingue des *Anax* par une légère différence de longueur du corps, et les deux espèces d'*Anax* peuvent être discriminées en mesurant la longueur de l'extension de leur épiprocte par rapport à celle de leurs cerques (au niveau de la "pyramide anale"). Les larves plus jeunes de ces trois espèces ne peuvent être distinguées morphologiquement à l'aide de la clé. Photographies : Christophe Brochard.

Pour tester le bon fonctionnement du système Xperience et l'utilisation de la clé d'identification T1, j'ai fait appel à un panel de 24 "testeurs" appartenant à cinq catégories de personnes susceptibles d'utiliser ce type d'outil d'identification pour leurs activités personnelles ou professionnelles. Les cinq catégories étaient ainsi divisées :

- 7 chercheurs entomologistes
- 7 naturalistes amateurs (membres de l'association naturaliste Timarcha⁵⁶)
- 4 professeurs de collège et lycée
- 4 étudiants du Master "Systématique, Évolution, Paléontologie" (MNHN-SU)
- 2 professionnels de terrain en hydrobiologie et suivis de biodiversité aquatique.

L'expérience s'est faite avec une version traduite en anglais de la clé T1 des larves d'odonates (publication en préparation, voir **Annexe 9**) (**Figure 34**). La même loupe binoculaire a été utilisée pour toutes les identifications, de la marque Olympus, avec une capacité d'agrandissement allant de x7,5 à x64. Les spécimens étaient éclairés à l'aide d'une fibre optique (**Figure 35**). Une solution d'éthanol à 70% et deux paires de pinces (une rigide et une souple) ont été fournis à chaque testeur. Les spécimens étaient présentés aux testeurs sur une coupelle d'observation, et conservés dans une solution d'éthanol à 90% lorsqu'ils n'étaient pas utilisés pour un test.

⁵⁶ timarcha.org

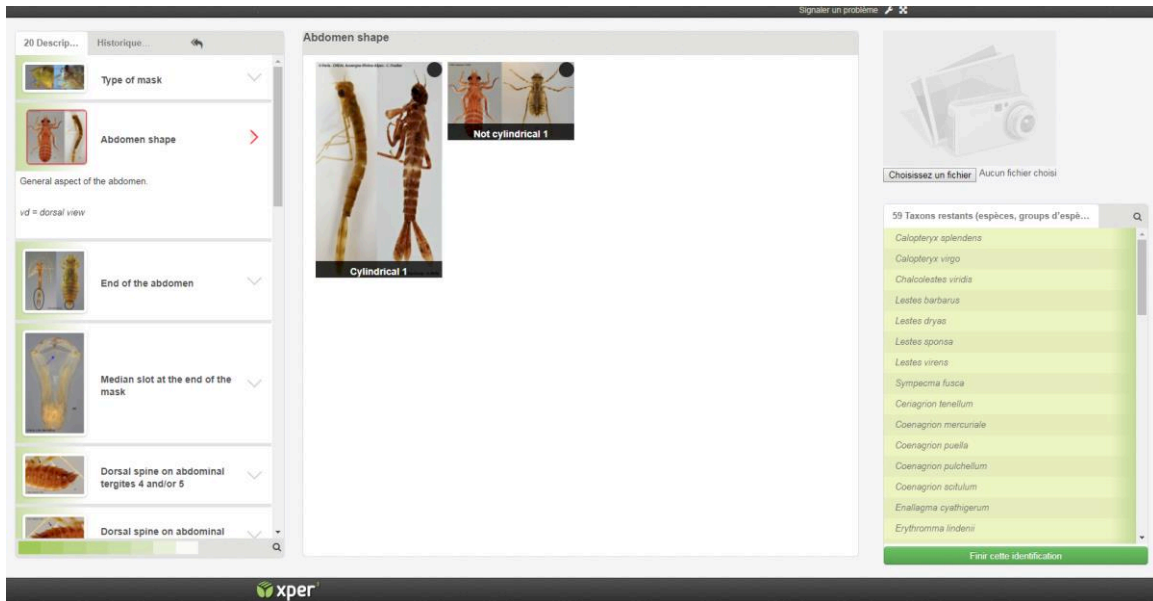


Figure 34. Interface de la clé d'identification en anglais des larves d'odonates l'Ile-de-France, grâce à laquelle les testeurs ont effectué les identifications des trois spécimens récoltés sur le terrain.



Figure 35. Une participante au test de la clé des larves d'odonates.

Avant de débiter les identifications, chaque testeur a visionné une vidéo de deux minutes présentant le fonctionnement des clés Xper³. Ce tutoriel précisait qu'il était possible de sélectionner plusieurs états pour un même descripteur. **Les trois spécimens ont été présentés successivement aux testeurs. L'ordre de présentation s'est fait aléatoirement** : le premier, second et troisième spécimen à identifier n'étaient pas les mêmes pour chaque testeur, afin de limiter le "biais d'apprentissage" qui peut faire que l'identification du premier spécimen est plus souvent erronée que les suivantes. Aucune limite de temps n'a été imposée pour l'expérience. Aucune aide n'a été donnée aux testeurs pendant leur démarche d'identification, sauf s'ils rencontraient des difficultés techniques avec le logiciel ou avec l'anglais (tous les testeurs étaient francophones).

Tous les étapes d'identification et tous les choix finaux de taxons ont été enregistrés dans une base de données Xperience comme pour la clé "Insectes" du Spipoll. J'ai analysé les données exportées sous forme de tableur Excel.

L'échantillonnage limité de participants, ainsi que le faible nombre de testeurs dans chaque catégorie, ne permettent pas une analyse statistique des résultats. Cependant nous pouvons observer des tendances : le paragraphe suivant présente les résultats obtenus.

4.2.3b Résultats

Ici, contrairement à la base de données issue de l'utilisation de la clé "Insectes" du Spipoll, le jeu de données est relativement petit, toutes les démarches ont pu être intégralement enregistrées par Xperience jusqu'à la validation de l'identification finale, et nous savons à quel taxon appartient chaque spécimen identifié par les testeurs. Il nous est donc possible d'effectuer une **analyse plus fine** que dans le 4.2.2, c'est-à-dire **analyser les descripteurs individuellement (sans les regrouper par catégories), ainsi que le pourcentage de doutes et d'erreurs pour chaque descripteur et identification finale.**

❖ *Identifications finales et doutes*

Dans l'ensemble, on observe une nette **augmentation du pourcentage d'identifications correctes entre le premier spécimen observé et les suivants.**

Le nombre moyen de pas (validation d'au moins un état pour un descripteur) nécessaires pour obtenir l'identification finale est homogène entre les différentes catégories (6,46 pour l'ensemble des testeurs) (**Tableau 21**).

Tableau 21. Résultats totaux et par catégorie de testeurs pour le nombre moyen d'étapes nécessaires à l'obtention d'une identification finale, et le pourcentage d'identification correcte du "spécimen 1" (le premier spécimen présenté aux testeurs au cours de l'expérience), du "spécimen 2" (second spécimen présenté) et du "spécimen 3" (troisième et dernier spécimen présenté).

Catégorie de testeurs	Nombre de testeurs	Nombre moyen de pas	% d'identification correcte du spécimen 1	% d'identification correcte du spécimen 2	% d'identification correcte du spécimen 3
Entomologistes	7	6,33	57,14	71,43	71,43
Naturalistes	7	6,95	28,57	57,14	57,14
Enseignants	4	6,33	0	25	25
Étudiants	4	6,08	25	25	50
Professionnels terrain	2	6,17	50	100	50
<i>Total</i>	<i>24</i>	<i>6,46</i>	<i>33,33</i>	<i>54,17</i>	<i>54,17</i>

Les professionnels de terrain spécialisés dans les suivis de populations de “macro-invertébrés aquatiques” **sont la catégorie de testeurs ayant réalisé le plus haut taux d’identifications finales correctes** (tous spécimens et chemins d’identifications compris), avec 66,67% d’identifications finales correctes. Ils sont suivis par les chercheurs entomologistes (61,90% d’identifications finales correctes), les naturalistes amateurs (42,86%), les étudiants en master de sciences naturelles (33,33%) et les enseignants du second degré (16,67%) (**Figure 36**).

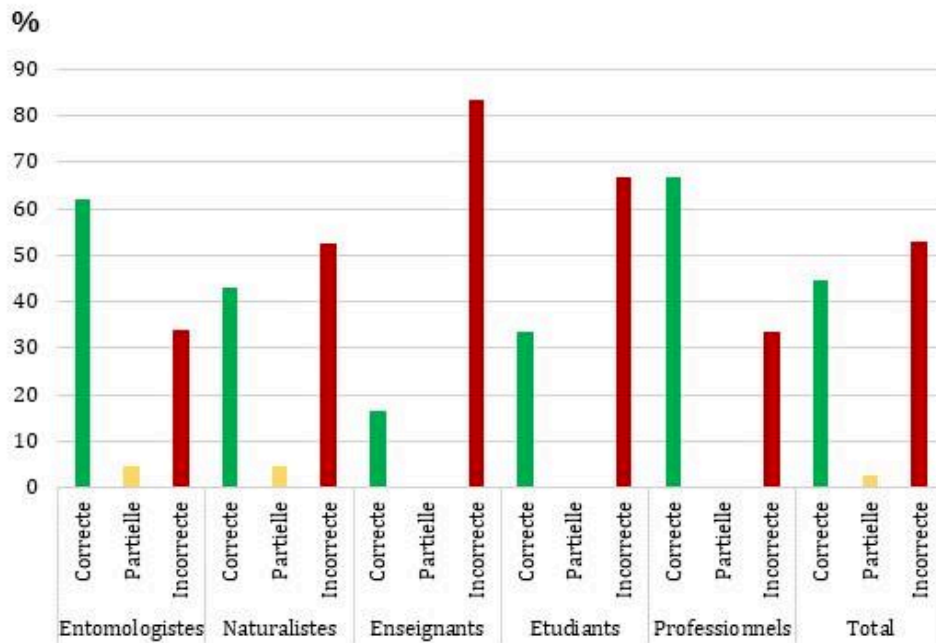


Figure 36. Représentation, au total et par catégorie de testeurs, du pourcentage d’identifications correctes, partiellement correctes et incorrectes, tous spécimens et chemins d’identification compris. Une identification finale est considérée “partiellement correcte” si l’item correspondant au spécimen à déterminer fait partie des propositions finales d’items, mais que d’autres items sont également présents dans la liste finale.

Au total, **le spécimen a été correctement identifié près de la moitié du temps** (44,44%). **Le taux d’erreurs est d’un peu plus de 50%** (52,78%). **Les testeurs doutent très peu** lors du choix des états de descripteurs (3,01% au total) malgré le fait qu’ils ont été invités à le faire (**Tableau 22**).

Tableau 22. Nombre et pourcentage totaux d'identifications correctes, partiellement correctes et incorrectes ; nombre et pourcentage de doutes exprimés sur l'ensemble des descripteurs, tous spécimens et chemins d'identification compris.

	Nombre	%
Identification finale (correcte/partiellement correcte/incorrecte)	32/2/38	44,44/2,78/52,78
Doutes (choix de plusieurs états de descripteur)	14	3,01

❖ Proximité taxonomique des identifications

Lorsque l'identification était "partiellement correcte" ou "incorrecte" (40 démarches d'identification), j'ai regardé à quel point elle était proche du taxon correct (**Tableau 23**). Par exemple, si le spécimen à identifier était *Libellula quadrimaculata* et que le testeur l'a déterminé comme étant *Cordulia aenea*, alors la proximité taxonomique est au niveau de la super-famille (ces deux taxons appartiennent à la super-famille Libelluloidea).

Tableau 23. Proximité taxonomique des identifications partiellement correctes et incorrectes, selon qu'elles appartiennent à la même famille, la même super-famille, le même sous-ordre ou le même ordre que le taxon correct correspondant au spécimen présenté au testeur.

Proximité taxonomique	Nombre	%
Même famille	21	52,5
Même super-famille	8	20
Même sous-ordre	8	20
Même ordre	3	7,5

Plus de la moitié des identifications partiellement correctes et incorrectes sont relativement peu éloignées taxonomiquement du taxon correct (52,5% appartiennent à la même famille que ce dernier).

❖ Choix des descripteurs, doutes et erreurs

Je me suis également penchée sur les erreurs faites au niveau des descripteurs (mauvais choix d'état(s) de descripteurs). Le **Tableau 24** présente tous les descripteurs ayant été choisis au moins une fois au cours de la démarche d'identification d'un testeur (toutes démarches et tous spécimens confondus) ainsi que le taux d'erreur dans le choix des états pour chacun de ces descripteurs.

Tableau 24. Liste de tous les descripteurs ayant été choisis au moins une fois au cours de la démarche d'identification d'un testeur (toutes démarches et tous spécimens confondus), le taux d'erreur dans le choix des états pour chacun de ces descripteurs, et leur "niveau d'erreur". Niveau d'erreur vert : le descripteur a entraîné moins de 5% d'erreurs ; orange : entre 5% et 15% d'erreurs ; rouge : plus de 15% d'erreurs.

Descripteur	Nombre de fois qu'il a été choisi	% d'erreurs	Niveau d'erreur
End of the abdomen	55	0	Green
Diameter of the eye	40	7,50	Orange
Type of mask	30	13,33	Orange
Median slot at the end of the mask	29	3,45	Green
Abdomen shape	25	4,00	Green
Dorsal spine on abdominal tergite 7	22	31,82	Red
End of the caudal gills	21	19,05	Red
Dorsal spine on abdominal tergites 4 and/or 5	20	15,00	Orange
Furrow at the base of the mask	19	26,32	Red
Lateral expansions at the back of the head	18	11,11	Orange
Lateral spines on abdominal tergite 5	17	17,65	Red
Dorsal spine on abdominal tergite 9	16	0	Green
Number of long setae inside of the mask	15	13,33	Orange
The ovipositor reaches or exceeds the posterior margin of sternite 9	11	0	Green
Lateral spines on abdominal tergite 7	11	0	Green
Relative length of the cerci and paraprocts	10	20,00	Red
Antenna shape	10	0	Green
Pairs of legs 1 and 2	10	20,00	Red
Row of spines on abdominal stergite 2	8	12,50	Orange
Lateral spines on abdominal tergite 6	8	37,50	Red
Lateral sides of the head	6	33,33	Red
Horizontal node in the middle of the caudal gills	4	0	Green
Shape of the prementum	4	0	Green
On top of the anal pyramid, an expansion with a pointed end	3	0	Green
Teeth on the labial palps	3	0	Green
Protuberance(s) between the eyes	3	0	Green
Setae on the posterior margin of abdominal tergites 8 and 9	3	0	Green
Central caudal gill	2	0	Green
Relative length of the epiproct expansion and of the cerci	1	0	Green
Number of setae in the movable hook of the labial palps	1	0	Green

La **Figure 37** permet de visualiser à la fois le nombre de fois que chaque descripteur du **Tableau 24** a été choisi, et le pourcentage d’erreurs associé aux descripteurs.

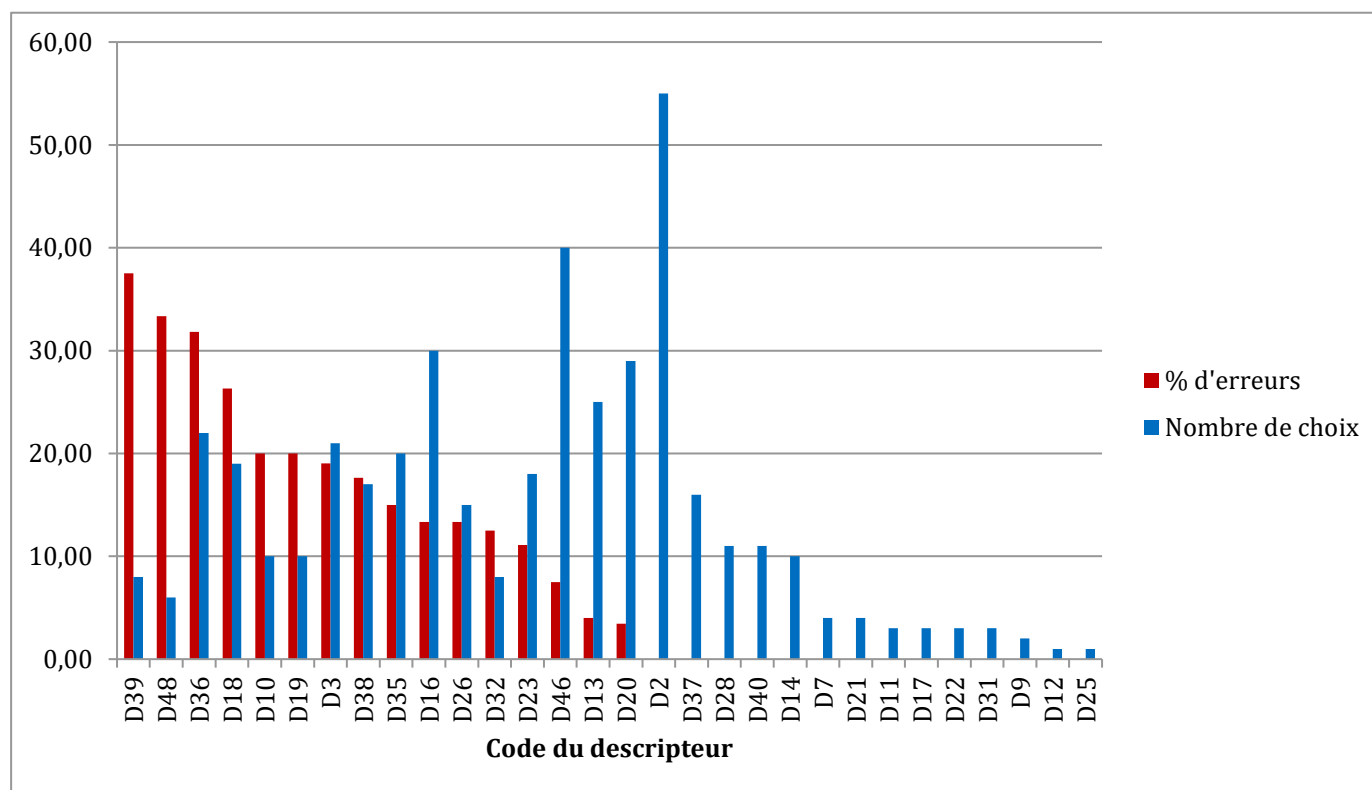


Figure 37. Nombre de choix et pourcentage d’erreurs pour les 30 descripteurs ayant été choisis au moins une fois au cours des démarches d’identification des testeurs. En bleu : nombre de fois que chaque descripteur a été choisi. En rouge : pourcentage d’erreurs associé à chaque descripteur.

Cette analyse permet de mettre en évidence quels sont les **descripteurs “efficaces”** – **beaucoup choisis et entraînant peu d’erreurs** – et lesquels sont **“inefficaces”** – **entraînant beaucoup d’erreurs**. Elle montre notamment que :

- Le descripteur D13 et tous les autres descripteurs à sa droite sur la **Figure 37** peuvent être considérés comme des descripteurs efficaces : ils n’entraînent pas ou peu (moins de 5%) d’erreurs dans les choix d’états de descripteur.
- Le descripteur D2 “End of the abdomen” est de loin le plus choisi et il n’entraîne pas d’erreur. C’est une bonne chose car c’est aussi l’un des descripteurs les plus discriminants (indice Xper égal à 0,48 – voir **Tableau 10**) : il permet de faire la distinction entre les zygotères et les anisoptères.

- Les descripteurs D39 "Lateral spines on abdominal tergite 6", D48 " Lateral sides of the head", D36 "Dorsal spine on abdominal tergite 7" et D18 "Furrow at the base of the mask" posent problème car ils entraînent le plus grand pourcentage d'erreurs (plus de 25%).

- Les descripteurs les plus problématiques sont D36, D18, D3 "End of the caudal gills" et D38 "Lateral spines on abdominal tergite 5". En effet, non seulement ils entraînent plus de 15% d'erreurs, mais ils font également partie des 10 descripteurs les plus choisis, toutes démarches d'identification confondues. Ce sont donc les "points faibles" de cette clé de détermination pour les taxons A, B et C qui ont fait l'objet de cette expérience de test.

Il est important de noter un biais dans les choix des descripteurs. Pour rappel, l'interface d'identification de Xper³ présente la liste des descripteurs du plus discriminant au moins discriminant, et cet ordre est mis à jour à chaque étape de l'identification pour correspondre aux items restant compatibles avec la description faite par l'utilisateur.

J'ai observé que pour les trois taxons de ce test, **les descripteurs les plus choisis à la première étape d'identification sont systématiquement ceux qui apparaissent en premier dans la liste** des descripteurs proposés à l'utilisateur (donc les descripteurs les plus discriminants). Le **Tableau 25** montre que les trois descripteurs les plus choisis à la première étape d'identification sont aussi les trois descripteurs listés en premier dans l'interface de la clé testée.

Tableau 25. Liste des descripteurs choisis au moins une fois à la première étape de l'identification, du plus choisi au moins choisi, selon le spécimen présenté au testeur.

Descripteurs choisis à la première étape	Nombre de fois qu'ils ont été choisis	Position dans la liste des descripteurs proposés
Taxon A : <i>Libellula quadrimaculata</i>		
Type of mask	10	1
Abdomen shape	9	2
End of the abdomen	3	3
Dorsal spine on abdominal tergites 4 and/or 5	1	5
Diameter of the eye	1	12
Taxon B : <i>Pyrrhosoma nymphula</i>		
Abdomen shape	9	2
End of the abdomen	7	3
Type of mask	7	1
Median slot at the end of the mask	1	4
Taxon C : "complexe <i>Anax/Hemianax</i>"		
Type of mask	8	1
End of the abdomen	7	3
Abdomen shape	4	2
Median slot at the end of the mask	2	4
Number of long setae inside of the mask	1	8
Dorsal spine on abdominal tergites 4 and/or 5	1	5
Diameter of the eye	1	12

4.2.3c Discussion et conclusion

L'analyse de l'utilisation de la clé des odonates par les testeurs a permis de mettre en évidence des défauts représentant des obstacles à une identification juste des spécimens.

Sans surprise, les catégories de testeurs ayant fait le moins d'erreurs dans les identifications (plus de 60% d'identifications correctes) regroupent des personnes qui sont souvent amenées à effectuer des déterminations de spécimens d'insectes : les professionnels effectuant des suivis de biodiversité sur le terrain et les chercheurs en entomologie.

Le taux de "réussite" pour les identifications avec cette clé des larves d'odonates est d'environ 50%. D'autre part, plus de la moitié des identifications incorrectes de spécimens ont mené à l'attribution d'un item relativement proche du taxon correcte (même famille).

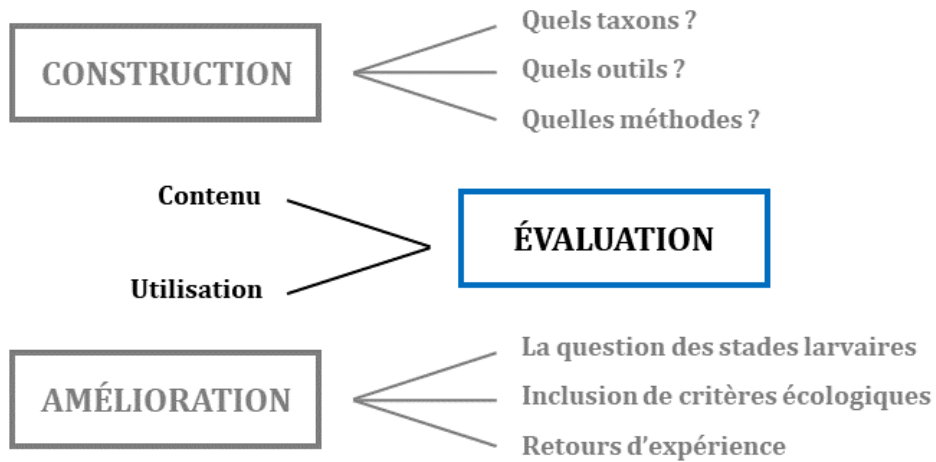
Comme observé pour la clé d'identification du Spipoll, les utilisateurs doutent très peu, que ce soit au niveau du choix des états de descripteurs ou pour les identifications

finales. On peut se demander si cette “réticence” par rapport au doute ne viendrait pas d’un conditionnement social acquis dès le jeune âge, du fait qu’au cours de la scolarité et de l’apprentissage en général, il est le plus souvent demandé de donner une seule réponse qui serait l’unique réponse “correcte” attendue. Il serait intéressant de demander aux utilisateurs pourquoi ils n’ont pas exprimé de doutes.

De même, les utilisateurs ont tendance à choisir les descripteurs “les plus visibles” dans la clé de détermination, soit les premiers dans la liste de descripteurs proposés.

Ces tendances doivent être considérées au moment de la démarche de construction de telles clés d’identification.

Les descripteurs entraînant le plus haut taux d’erreurs concernent des structures de petite taille ou des détails, par exemple des épines sur l’abdomen, des caractères présents sur les pattes ou les antennes, ou encore des détails au niveau du masque. Ces types de descripteurs doivent faire l’objet d’une attention particulière lors de la construction de clés d’identification numériques à accès libre. Il est crucial de soigner le codage des états de ces descripteurs, ainsi que la présentation de ces descripteurs et de leurs états (illustrations et textes explicatifs) pour réduire les risques d’erreurs.



Les résultats obtenus après analyse de l'utilisation de ces deux clés d'identification numériques sous Xper³ serviront de base aux réflexions sur les modifications à apporter afin d'améliorer ces clés – du point de vue de la prise en main de la part des utilisateurs mais aussi de la fiabilité des identifications – afin d'augmenter les taux de déterminations correctes des spécimens. L'amélioration des clés d'identification sera l'objet du chapitre 5.

Chapitre 5 – Comment améliorer des clés d'identification numériques ?

L'évaluation des clés d'identification numériques a permis de mettre plus concrètement en évidence les points forts de ce type d'outils, ainsi que de soulever certains problèmes courants qui peuvent s'avérer des obstacles à une détermination fiable des spécimens d'entomofaune. Dans ce chapitre, nous allons explorer **plusieurs approches pour améliorer les outils d'identification numériques** (certaines des clés réalisées au cours de la thèse et la clé d'identification "Insectes" du Spipoll) :

- L'ajout de critères non morpho-anatomiques dans les clés (et l'apport de données moléculaires pour préciser l'écologie fine des larves)
- Les modifications faites suite à analyse fine de l'utilisation des clés
- La prise en compte des retours d'expérience des utilisateurs des clés pour l'amélioration de ces outils (meilleure fiabilité des déterminations et utilisation facilitée pour différents types de publics).

5.1 Intégration de critères écologiques dans les clés d'identification

La majorité des clés d'identification existantes se basent sur des caractères morpho-anatomiques uniquement. J'ai souhaité tester l'intérêt d'intégrer dans les clés d'identification des larves insectes aquatiques (odonates et éphémères d'Ile-de-France) des critères qui ne sont pas des caractères morpho-anatomiques.

5.1.1 Le cas des larves d'insectes aquatiques : odonates et éphémères

Les larves d'insectes aquatiques sont particulièrement intéressantes pour les études d'écologie et les suivis de populations. D'une part, ne possédant pas d'ailes comme les adultes, les larves **ne peuvent pas disperser loin de leur lieu de vie** principal. D'autre part, les individus **peuvent rester longtemps sous forme de larve** – jusqu'à deux ans pour les éphémères et cinq ans pour les odonates, ce qui représente la majeure partie de leur cycle de vie (Snodgrass 1954, Tachet *et al.* 2010). L'étude des larves représente donc une source d'information fiable en ce qui concerne la localisation précise des populations. Ce sont

d'ailleurs les individus vivant dans les cours d'eau (des larves pour la plupart des ordres d'insectes aquatiques) qui sont à la base des prélèvements de terrain pour les suivis de biodiversité en milieu aquatique et le calcul d'indices tels que l'IBG-DCE (cf. 1.3.1).

Les clés d'identification T1 à T7 présentées aux chapitres 3 et 4 concernent les **nymphes (derniers stades larvaires)** ou les imagos. Cependant la "phase larvaire" n'est pas homogène ; elle se décompose en plusieurs **mues larvaires** qui peuvent avoir lieu sur un temps long. Chez les odonates par exemple, on observe – avec des variations selon les espèces – de 8 à 18 mues larvaires, mais la majorité des espèces en effectue entre 11 et 13 (Daniel Grand *et al.* 2014) ; chez les éphémères, ce nombre peut monter jusqu'à une cinquantaine selon l'espèce (Ide 1935, Bauernfeind & Soldán 2012).

À chaque nouvelle mue l'aspect de l'individu évolue, et de nouvelles structures morpho-anatomiques se modifient ou apparaissent. Les stades larvaires les plus jeunes (peu de temps après l'éclosion) ne présentent donc qu'une petite partie de toutes les structures qui sont observables chez les nymphes et qui servent comme critères pour l'identification des espèces (**Figures 38 et 39**).

Les clés T1 et T2, pour les larves d'odonates et d'éphémères d'Ile-de-France, ne permettent donc de distinguer les espèces entre elles que si l'individu a effectué toutes ses mues larvaires ou la quasi-totalité. Cela restreint les possibilités d'identification, sachant que la plupart des spécimens prélevés sur le terrain pour les suivis de populations n'ont pas effectué la plupart de leurs mues larvaires. **Si l'on considère uniquement les structures morpho-anatomiques présentes et visibles dès les premiers stades larvaires, la détermination à l'espèce est impossible pour la plupart des taxons d'odonates et d'éphémères.**

Notre objectif était donc d'**intégrer des critères supplémentaires** dans les bases de connaissances Xper³ des larves d'odonates et d'éphémères, **afin de tester si leur capacité d'identification peut être affinée** par rapport à l'utilisation de caractères morpho-anatomiques seuls.

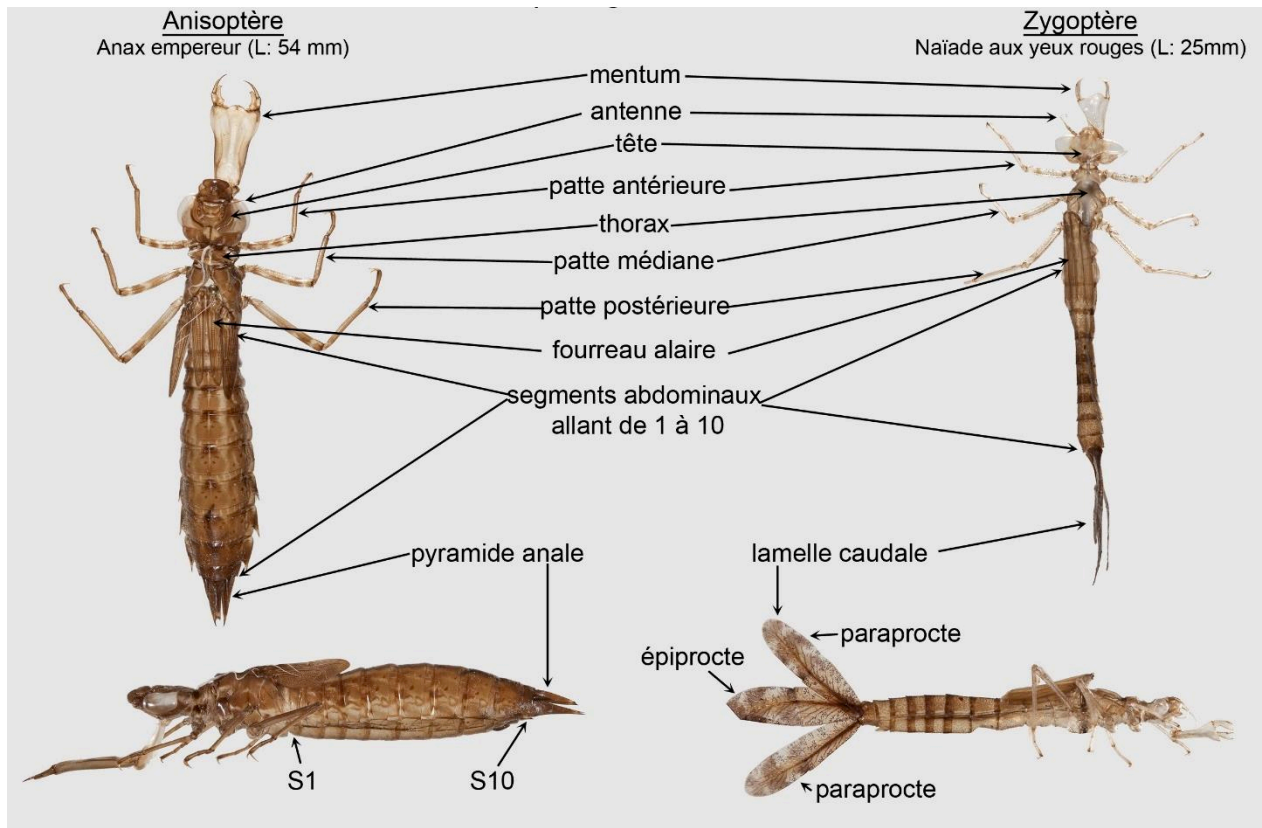


Figure 38. Représentation des principales structures morphologiques reconnaissables sur une exuvie d'odonate. D'après des photographies de Guillaume Doucet.

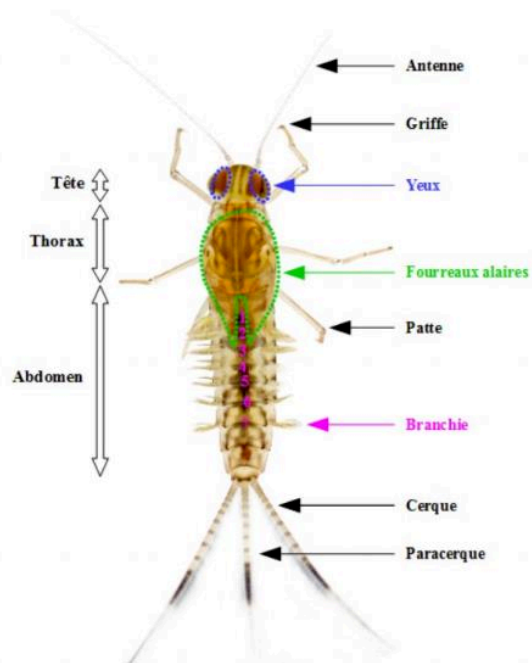


Figure 39. Représentation des principales structures morphologiques reconnaissables sur une nymphie d'éphémère. Figure réalisée par Paola Campos.

Cette étude a été réalisée dans le cadre de deux stages de niveau Master 1 : le stage d'**Aurélia Dubois**, en M1 "Écologie, Biodiversité, Évolution" qui s'est déroulé d'avril à juin 2015, et celui de **Paola Campos**, en M1 "Systématique, Évolution, Paléontologie" qui s'est déroulé d'avril à août 2017. Le stage d'Aurélia s'intéressait aux larves d'odonates, tandis que celui de Paola a porté sur les larves d'éphémères.

La première étape a été de **déterminer quelles structures morpho-anatomiques étaient présentes et visibles chez tous les stades larvaires** pour ces deux ordres, y compris chez les tous premiers stades après éclosion. Pour cela, nous nous sommes appuyés sur des documents de la littérature (Ide 1935, Degrange 1960), sur nos observations personnelles à partir de larves prélevées par la DRIEE ou par nous-mêmes sur le terrain, et sur les conseils d'experts de ces groupes (Guillaume Doucet et Michel Brulin).

Chez toutes les larves d'odonates et d'éphémères :

- Dès la première mue larvaire, les structures suivantes sont présentes et visibles à l'œil nu ou grâce à une loupe binoculaire : tête, yeux, antennes, pattes, thorax, abdomen, lamelles caudales ou pyramide anale (odonates), cerques et au moins une partie des branchies (éphémères)
- Chaque mue s'accompagne d'une augmentation de la taille et du volume des individus, et de la taille des lamelles caudales et branchies
- Les pièces buccales des larves les plus jeunes sont visuellement différentes de celles des larves les plus âgées
- Les fourreaux alaires n'apparaissent pas avant les dernières mues.

D'autre part, chez les larves d'éphémères :

- Chaque mue s'accompagne de l'ajout d'un à quatre segment(s) sur les cerques et paracerque
- Chez les espèces aux branchies ramifiées, cette ramification n'est pas observée avant les dernières mues larvaires.

A partir de ces éléments, nous avons construit de **nouvelles versions des clés d'identification T1 et T2**, que nous avons appelées **T1bis et T2bis**. Nous y avons intégré les **caractères morpho-anatomiques observables à partir de la seconde mue (troisième stade larvaire)**. En effet avant cela, les branchies des éphémères ne sont pas visibles et les individus mesurent (hors antennes et cerques) moins d'un millimètre, ce qui rend l'observation trop compliquée pour des non-spécialistes des groupes et si on n'a pas un matériel d'observation assez puissant.

Les descripteurs présents dans les bases de connaissances T1bis et T2bis sont les suivants.

T1bis : Odonates d'Ile-de-France, tous stades larvaires, morphologie uniquement

1. Extrémité de l'abdomen
2. Taille des antennes
3. Type de masque
4. Trou dans le prémentum
5. Taille du premier article antennaire par rapport aux autres articles

T2bis : Éphémères d'Ile-de-France, tous stades larvaires, morphologie uniquement

1. Longueur du corps
2. Carapace dorsale
3. Branchies visibles dorsalement
4. Aplatissement dorso-ventral
5. Position des branchies
6. Seconde paire de branchies
7. Taille relative du paracerque
8. Aspect des branchies

Nous nous sommes posé les questions suivantes :

- Quels types de critères non morpho-anatomiques pouvons-nous intégrer dans nos clés d'identification ?
- Quels sont les degrés de discrimination des clés d'identification avec et sans l'ajout des critères écologiques en plus des caractères morphologiques ?

- Comment obtenir et utiliser des données moléculaires pour affiner les bases de données en ce qui concerne l'écologie fine des larves présentes en Ile-de-France ?

5.1.2 L'apport des critères écologiques

5.1.2a Problématiques

Comme présenté au 2.1.3, les clés d'identification peuvent se baser sur plusieurs types de caractères pour la description des spécimens : morpho-anatomie, écologie et traits d'histoire de vie, chants et sons, séquences d'ADN. Dans le cas des larves d'odonates et des éphémères d'Ile-de-France, seules leurs **préférences écologiques** sont assez bien connues pour pouvoir être exploitées dans ce travail.

Nous avons vu dans la partie précédente que chez ces ordres, les structures morphologiques présentes et assez développées dès les premiers stades larvaires sont peu nombreuses. Elles ne permettent pas à elles seules de discriminer toutes les espèces de notre échantillon.

Notre objectif était de tester si l'ajout, dans les clés d'identification, de critères relatifs à l'écologie et au mode de vie des différents taxons, permettait d'affiner la puissance de discrimination de nos clés d'identification des larves d'odonates et d'éphémères.

5.1.2b L'écologie des larves d'odonates et d'éphémères

Les odonates sont l'un des ordres d'insectes les mieux connus en termes de taxonomie et d'écologie. Toutes les espèces sont prédatrices et dépendantes de la présence d'eau (Dupont 2010). Certains odonates sont présents uniquement dans des eaux courantes (ruisseaux, canaux, rivières, fleuves...); la plupart sont inféodés à des habitats aux eaux stagnantes (étangs, mares, lacs, tourbières...), relativement présents en Ile-de-France,

notamment en forêt (**Figure 40**). Quelques espèces peuvent se trouver dans les deux types de milieux (Grand & Boudot 2006, Houard & Merlet 2014).

Divers critères écologiques caractérisent les milieux dans lesquels se trouvent les différentes espèces : température, turbidité et pH de l'eau, vitesse du courant, abondance de la végétation, quantité de proies...



Figure 40. Mare naturelle abritant des larves d'odonates, à Boussy-Saint-Antoine (91). Photographie : M. Delaunay

Les larves d'éphémères présentent des habitudes de vie sédentaire et leurs préférences écologiques sont variées. Toutes les espèces sont dépendantes de la présence d'eau (**Figure 41**). Il s'agit d'un groupe diversifié au niveau de leurs habitats (eaux stagnantes à courantes), de leur mode de locomotion (larves fouisseuses, rampantes ou nageuses), et de leur type d'alimentation (broyeurs détritivores, racleurs de substrats, filtreurs, prédateurs...).



Figure 41. Berge de ruisseau à faible courant, un des milieux de vie des Éphémères. © Natureparif

Les bases de données sur l'écologie des larves d'insectes aquatiques sont incomplètes et fractionnées. Cependant, il existe une grande base de données en ligne sur l'écologie, la biologie et la distribution des organismes d'eau douce (téléostéens, insectes, autres "macro-invertébrés", macrophytes, diatomées et phytoplancton) au niveau européen, nommée "Fresh Water Ecology"⁵⁷ (Schmidt-Kloiber & Hering 2015). Cette base de données a été réalisée dans le cadre d'un projet financé par l'Union Européenne et implique des experts de la taxonomie et de l'écologie des groupes, appartenant à huit pays européens. L'objectif est de regrouper à terme toutes les connaissances scientifiques existantes sur l'écologie des organismes d'eau douce. La base de données sur le site web est régulièrement mise à jour. Trois ouvrages papier ont également été publiés pour les données issues de cette base, relatives aux ordres d'insectes aquatiques Tricoptera, Plecoptera et Ephemeroptera (Graf *et al.* 2008, Graf *et al.* 2009, Buffagni *et al.* 2009).

5.1.2c Sélection des critères écologiques et seconde version des clés

Pour sélectionner les critères écologiques pertinents dans le cadre de notre travail, nous nous sommes principalement référés au site web Fresh Water Ecology et au livre correspondant pour les éphémères (Buffagni *et al.* 2009). Nous avons également utilisé des

⁵⁷ www.freshwaterecology.info

ouvrages naturalistes (pour les odonates : Dommanget 1998, Dijkstra & Lewington 2006, Grand & Boudot 2006, Doucet 2010, Grand *et al.* 2014 ; pour les éphémères : Bauernfeind & Soldán 2012).

Nous avons sélectionné des critères écologiques pertinents pour la région Ile-de-France (par exemple, l'altitude et la salinité de l'eau ont été écartées). Ces critères devaient également être assez précis pour pouvoir faire la distinction entre les différentes espèces, mais nous avons aussi fait attention à la possibilité de les coder correctement et sans ambiguïté dans le système Xper³.

Nous avons ensuite construit de **nouvelles versions plus complètes des clés d'identification T1bis et T2bis**, que nous avons appelées **T1ter et T2ter**. Ces clés comprennent donc à la fois les caractères morphologiques présents dès la seconde mue chez les larves d'odonates et d'éphémères, et des critères écologiques propres à chaque espèce (au nombre de 17 pour les odonates et 8 pour les éphémères).

Les descripteurs présents dans les bases de connaissances T1ter et T2ter sont les suivants.

T1ter : Odonates d'Ile-de-France, tous stades larvaires, morphologie et écologie

Morphologie

1. Extrémité de l'abdomen
2. Taille des antennes
3. Type de masque
4. Trou dans le prémentum
5. Taille du premier article antennaire par rapport aux autres articles

Écologie

6. Intensité du courant
7. Microhabitat
8. pH minimum
9. Qualité de l'eau
10. Détectabilité
11. Autochtonie
12. Cortège d'espèces
13. Milieu artificiel

14. Milieu temporaire
15. Dynamique du milieu
16. Ensoleillement du milieu
17. Pollution organique
18. Pollution inorganique
19. Présence de poissons
20. Profondeur de l'eau
21. Pente de la berge
22. Végétation aquatique

T2ter : Éphémères d'Ile-de-France, tous stades larvaires, morphologie et écologie

Morphologie

1. Longueur du corps
2. Carapace dorsale
3. Branchies visibles dorsalement
4. Aplatissement dorso-ventral
5. Position des branchies
6. Seconde paire de branchies
7. Taille relative du paracercue
8. Aspect des branchies

Écologie

9. Zonation amont/aval
10. Niveaux d'eau et connectivité
11. Intensité du courant
12. Température de l'eau
13. Microhabitat
14. Mode d'alimentation
15. Mode de locomotion
16. Période de développement larvaire

Le **Tableau 26** fait la synthèse du contenu des différentes versions des clés d'identification T1 et T2. Les clés T1bis, T1ter, T2bis et T2ter ne comportent pas de descripteur-fils.

Tableau 26. Tableau récapitulatif du contenu des clés d'identification T1, T1bis, T1ter, T2, T2bis et T2ter. Pour chaque clé : nombre d'items, de descripteurs morpho-anatomiques et écologiques, nombre de descripteurs catégoriques (discrets) et numériques (continus), nombre minimal, maximal, moyen et total d'états de descripteurs.

Code base Xper ³	Nom de la clé d'identification	Nombre d'items	Nombre de descripteurs morpho-anatomiques	Nombre de descripteurs écologiques	Nombre de descripteurs catégoriques/numériques	Nombre d'états de descripteurs (min/max/moy)
T1	Larves d'odonates	59	54	0	52/2	116 (2/4/2,23)
T1bis	Odonates tous stades larvaires, morphologie seule	59	5	0	5/0	10 (2/2/2)
T1ter	Odonates tous stades larvaires, morpho et écologie	59	5	17	22/0	74 (2/9/3,36)
T2	Larves d'éphémères	39	52	0	51/1	115 (2/4/2,25)
T2bis	Éphémères tous stades larvaires, morphologie seule	39	8	0	7/1	16 (2/4/2,29)
T2ter	Éphémères tous stades larvaires, morpho et écologie	39	8	8	15/1	64 (2/11/4,27)

5.1.2d Amélioration de la détermination grâce à une approche multi-critères dans les clés d'identification

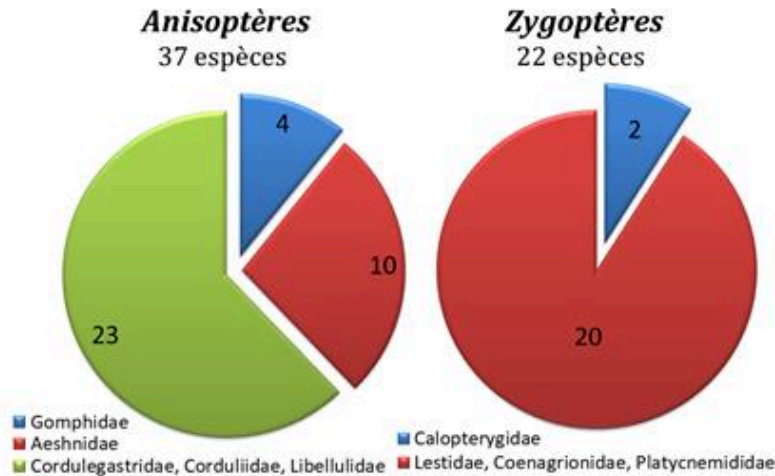
Pour les larves d'odonates et d'éphémères, **l'intégration de critères écologiques affine la capacité de discrimination des clés d'identification** : on observe une augmentation du nombre de groupes discriminés (**Figures 42 et 43**).

❖ *Odonates*

Amélioration de l'identification grâce aux données écologiques

Les résultats de cette étude ont fait l'objet d'un poster scientifique présenté lors du colloque "Les invertébrés dans la conservation et la gestion des espaces naturels" qui a lieu du 13 au 16 mai 2015 à Toulouse (**Annexe 12**).

Morphologie seule
Distinction de certaines familles



Morphologie + écologie
Distinction plus fine de certains genres et espèces

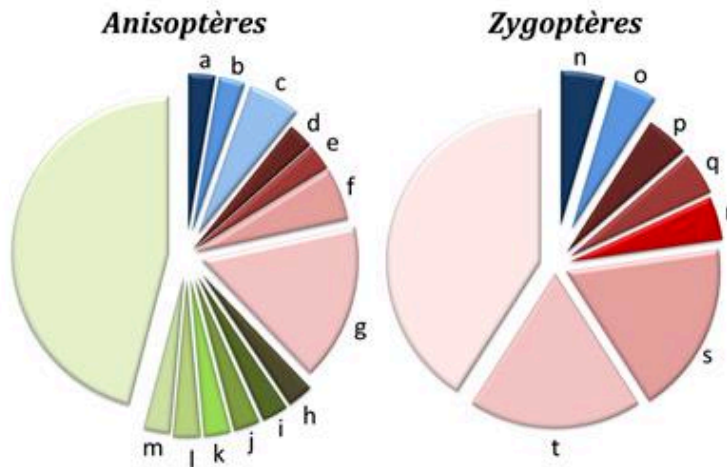


Figure 42. Comparaison de la puissance de discrimination des clés T1bis (morphologie seule, tous stades larvaires), en haut, et T1ter (morphologie et écologie, tous stades larvaires), en bas.

a. *Onychogomphus forcipatus* ; b. *Gomphus vulgatissimus* ; c. *Gomphus pulchellus* + *Gomphus simillimus* ; d. *Boyeria irene* ; e. *Hemianax ephippiger* ; f. *Anax imperator* + *Anax parthenope* ; g. toutes les espèces du genre *Aeshna* + *Brachytron pratense* ; h. *Epitheca bimaculata* ; i. *Leucorrhinia caudalis* ; j. *Leucorrhinia pectoralis* ; k. *Libellula quadrimaculata* ; l. *Orthetrum coerulescens* ; m. *Oxygastra curtisii*

n. *Calopteryx splendens* ; o. *Calopteryx virgo* ; p. *Platycnemis acutipennis* ; q. *Platycnemis pennipes* ; r. *Sympetma fusca* ; s. toutes les espèces du genre *Coenagrion* ; t. toutes les espèces du genre *Lestes*.

Chez les odonates, l'utilisation seule de caractères morphologiques présents chez tous les stades larvaires (clé T1bis) permet de distinguer des groupes de familles :

- Anisoptères : trois groupes formés, aucune espèce discriminée
- Zygoptères : deux groupes formés, aucune espèce discriminée.

Lorsque des variables écologiques sont également incluses (clé T1ter), on obtient une discrimination plus fine de certains genres et espèces :

- Anisoptères : 14 groupes formés, 10 espèces discriminées (sur 37)
- Zygoptères : 8 groupes formés, 5 espèces discriminées (sur 22).

L'utilisation de critères écologiques est efficace pour améliorer la détermination des jeunes larves d'odonates. Toutefois, les données écologiques disponibles pour notre étude ne sont pas suffisantes pour qu'une clé d'identification discrimine toutes les espèces.

Précisions sur les variables écologiques

Les zygoptères sont plus sensibles que les anisoptères à des conditions environnementales défavorables. Le descripteur écologique le plus discriminant de la clé T1ter est la présence/absence de pollution organique et inorganique, une problématique fondamentale en Ile-de-France. Un descripteur peut discriminer peu d'espèces mais se révéler important au sein d'un genre donné. Par exemple, les deux espèces du genre *Leucorrhinia* se distinguent uniquement par leur possible présence ou non dans une pièce d'eau abritant des poissons (Grand & Boudot 2006). De même, *Gomphus vulgatissimus* peut être discriminé des autres espèces de *Gomphus* par la nécessité stricte d'une végétation abondante dans son habitat (Dommanget 1998).

Certains critères écologiques peuvent être difficiles à appréhender à cause d'éventuelles variations au cours de l'année, ou la difficulté de les mesurer précisément. Par exemple, la vitesse de courant peut varier en cas de crue ou de sécheresse. La quantité d'ensoleillement d'une mare est difficile à mesurer et peut varier fortement selon les saisons.

De plus, certaines larves d'anisoptères peuvent se déplacer sur une courte distance, par exemple entre deux mares proches (Grand et Boudot 2006).

❖ *Éphémères*

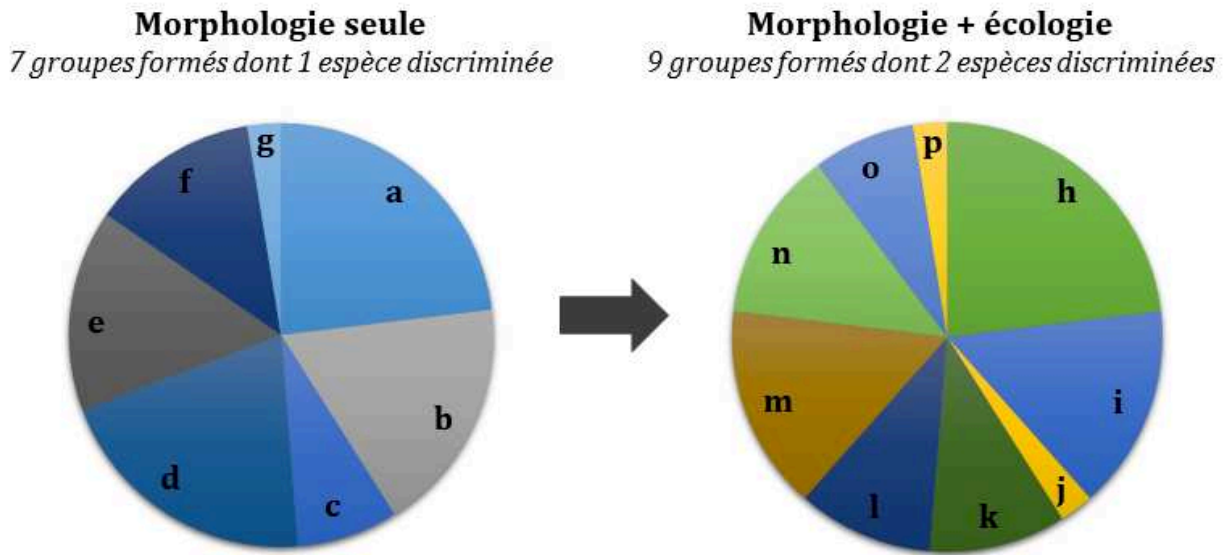


Figure 43. Comparaison de la puissance de discrimination des clés T2bis (morphologie seule, tous stades larvaires), et T2ter (morphologie et écologie, tous stades larvaires).

a. toutes les espèces du genre *Baetis* + *Labiobaetis atrebatinus* + *Metreletus balcanicus* ; b. toutes les espèces des genres *Cloeon* et *Procloeon* + *Baetopus tenellus* + *Centroptilum luteolum* + *Leptophlebia marginata* ; c. *Habrophlebia lauta* + *Paraleptophlebia submarginata* + *Potamanthus luteus* ; d. toutes les espèces du genre *Caenis* + *Brachycercus harrisellus* + *Cercobrachys minutus* ; e. toutes les espèces du genre *Ephemera* + *Serratella ignita* + *Ephoron virgo* ; f. toutes les espèces des genres *Ecdyonurus* et *Heptagenia* + *Electrogena ujhelyii* ; g. *Prosopistoma pennigerum*

h. toutes les espèces du genre *Baetis* + *Labiobaetis atrebatinus* + *Metreletus balcanicus* ; i. toutes les espèces des genres *Cloeon* et *Procloeon* + *Centroptilum luteolum* + *Leptophlebia marginata* ; j. *Baetopus tenellus* ; k. *Caenis macrura* + *C. pusilla* + *Brachycercus harrisellus* + *Cercobrachys minutus* ; l. *Caenis horaria* + *C. lactea* + *C. luctuosa* + *C. robusta* ; m. toutes les espèces du genre *Ephemera* + *Serratella ignita* + *Ephoron virgo* ; n. toutes les espèces des genres *Ecdyonurus* et *Heptagenia* + *Electrogena ujhelyii* ; o. *Habrophlebia lauta* + *Paraleptophlebia submarginata* + *Potamanthus luteus* ; p. *Prosopistoma pennigerum*.

Les caractères morphologiques seuls permettent de distinguer 7 groupes d'espèces pour les larves d'éphémères quel que soit leur stade larvaire. Une espèce peut être totalement discriminée grâce à la clé T2bis.

L'ajout de critères écologiques permet de distinguer 9 groupes d'espèces pour les larves d'éphémères quel que soit leur stade. Deux espèces (*Baetopus tenellus* et *Prosopistoma pennigerum*) peuvent être totalement discriminées grâce à la clé T2ter.

Comme chez les odonates, **l'utilisation de critères écologiques est efficace pour améliorer la détermination des jeunes larves d'éphémères**. Les nouveaux groupes d'espèces permettent de réduire le champ des possibles lors des identifications (des Caenidae notamment, divisés en deux groupes de quatre espèces). La clé T2ter est ainsi plus efficace et plus discriminante que la clé T2bis.

5.1.2e Limites de l'utilisation des données écologiques issues de la littérature

Les données écologiques issues de la littérature que nous avons utilisées pour notre étude et la réalisation de nouvelles clés d'identification ne s'avèrent pas suffisantes pour compenser le faible nombre de caractères morphologiques et permettre une détermination à l'espèce des jeunes larves d'odonates et d'éphémères. Nous faisons face à un **manque de connaissances sur l'écologie fine des larves**, notamment sur les échelles plus réduites qui nous intéressent dans le cas des larves d'Ile-de-France. En effet, la base de données Fresh Water Ecology et les ouvrages naturalistes sont très inclusifs : ils intègrent tous les états possibles pour une espèce au niveau d'un pays entier voire de plusieurs pays. D'autre part, de nombreuses informations sont indisponibles pour certaines espèces, et le nombre de critères écologiques renseignés est limité. Enfin, la description de certains états est trop vague, par exemple "lent", "moyennement lent" ou "rapide" pour la vitesse de l'eau (un descripteur numérique ou des catégories chiffrées de vitesse seraient plus pertinents). Pour certaines spécimens indistinguables par la morphologie et des données écologiques générales, la détermination à l'espèce peut être impossible, ou nécessiter un séquençage d'ADN pour comparer avec les bases de données moléculaires correspondantes (barcoding).

On peut s'attendre à ce l'intégration de critères écologiques plus nombreux, plus détaillés et spécifiques à la région Ile-de-France améliorent encore la capacité de

détermination des clés d'identification. C'est ce que nous avons souhaité faire pour les larves d'éphémères, en partenariat avec la Direction régionale et interdépartementale de l'environnement et de l'énergie (DRIEE) Ile-de-France. La prochaine partie de ce document présente cette étude.

5.1.3 Une étude moléculaire pour affiner les critères écologiques

5.1.3a Principes de l'étude

La DRIEE Ile-de-France est un service du Ministère de l'Environnement qui applique les priorités d'actions en matière d'énergie et d'environnement issues du Grenelle de l'Environnement. Le recensement des populations d'animaux et de végétaux représente ainsi une large partie des travaux réalisés par les techniciens de la DRIEE chaque année.

En ce qui concerne les "macro-invertébrés d'eau douce", les déterminations des spécimens relevés sur le terrain se limitent au genre, tel que recommandé pour les protocoles IBGN et IBG-DCE. La DRIEE compte 113 stations d'études de cours d'eau en Ile-de-France. 58 d'entre elles sont associées à des relevés hydrobiologiques et 54 à des relevés physico-chimiques.

La DRIEE sauvegarde toutes les données de terrain d'une année sur l'autre. Elle possède une base de données numérique pour les occurrences d'espèces et le nombre d'individus récoltés à chaque station, mais également des collections physiques de spécimens récoltés sur le terrain. En ce qui concerne les insectes aquatiques, 10 spécimens de chaque genre sont conservés par station et par année, dans une solution d'alcool à 70°.

Nous nous sommes focalisés sur les larves d'éphémères, car la DRIEE ne fait des relevés que dans les cours d'eau ; or la plupart des espèces d'odonates sont inféodées aux eaux stagnantes. Nous avons donc accès à plus de données pour les éphémères.

Comme dit précédemment, les connaissances sont incomplètes en ce qui concerne les préférences écologiques des larves d'éphémères (micro-habitats, température et qualité de l'eau, vitesse du courant...). D'autre part, sur GenBank⁵⁸ seules quelques espèces présentes en Ile-de-France sont représentées.

Les objectifs de cette étude étaient les suivants (**Figure 44**) :

- 1. Obtenir une collection moléculaire de référence pour les éphémères d'Ile-de-France, à partir d'une collection de nymphes d'éphémères fournie par le spécialiste du groupe Michel Brulin (OPIE), comportant 35 espèces clairement identifiées par ce dernier**
- 2. Séquencer des spécimens de larves d'éphémères récoltés par la DRIEE et les comparer (barcoding) à la collection moléculaire de référence, afin de les identifier à l'espèce sans ambiguïté**
- 3. Faire correspondre les noms d'espèces ainsi obtenus avec les données écologiques fines des stations où les spécimens ont été récoltés**
- 4. Intégrer ces nouvelles données écologiques dans la clé d'identification "morphologie + écologie" dans le but d'affiner la capacité de discrimination de la clé.**

⁵⁸ www.ncbi.nlm.nih.gov/genbank

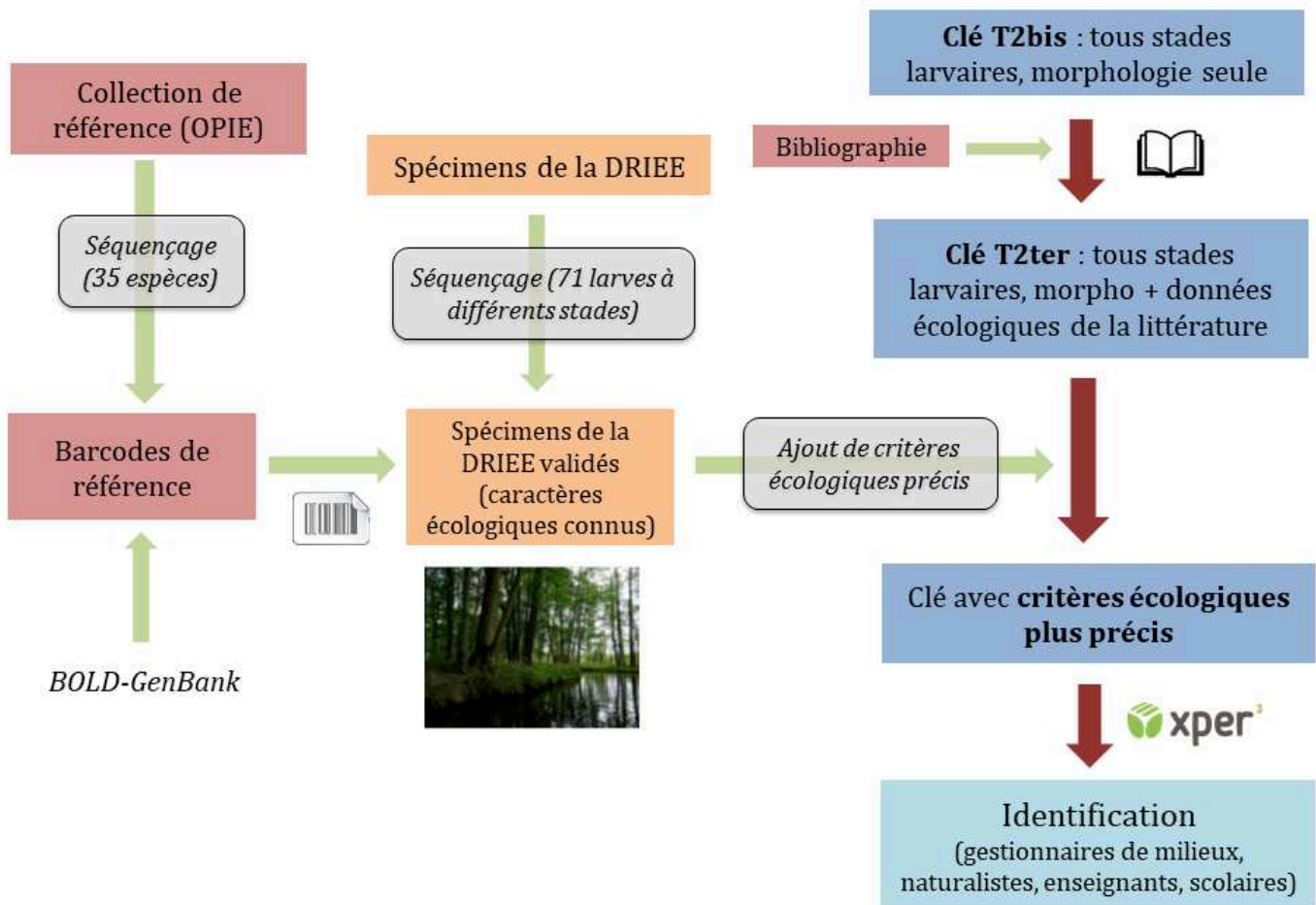


Figure 44. Schéma récapitulatif de l'ensemble des étapes prévues dans l'étude sur les larves d'éphémères.

5.1.3b Matériel et méthodes

❖ Échantillonnage taxonomique

Collection de référence de l'OPIE

Michel Brulin (OPIE) nous a fourni une collection de référence (constituée entre 2014 et 2016), en alcool, des larves d'éphémères d'Ile-de-France au stade nymphal. Cet expert des

éphémères a pu confirmer de façon certaine l'identification, à l'espèce, de chaque individu de la collection (un à trois par espèce).

Toutes les espèces d'Ile-de-France sont représentées dans la collection (voir le **Tableau 3**, 3.1.1a) sauf quatre : *Caenis lactea*, *Ephemera lineata*, *Electrogena ujhelyii* et *Prosopistoma pennigerum*. Les trois premières sont des espèces assez rares en Ile-de-France, et l'OPIE n'avait pas assez de spécimens de ces espèces en bon état à sa disposition pour nous en fournir. Quant à *Prosopistoma pennigerum*, cette espèce est probablement éteinte régionalement voire à l'échelle nationale : les dernières captures d'individus remontent à 1958 dans le Rhône près de Lyon⁵⁹. La collection de référence dont nous disposons comprend donc 35 espèces.

Spécimens récoltés par la DRIEE

Le personnel de la DRIEE a mis à notre disposition leurs collections des années 2014 et 2015. Face à la très grande quantité de spécimens disponibles et au fait que notre temps et les fonds pour effectuer des analyses moléculaires étaient limités, nous avons choisi de nous pencher sur le genre *Baetis* seulement. En effet les espèces de ce genre varient peu entre elles morphologiquement, y compris au stade nymphal. Lorsque les larves sont plus jeunes, il est souvent impossible de les distinguer grâce à des caractères morfo-anatomiques uniquement. Leur identification grâce au barcoding est donc particulièrement intéressante.

Pour rappel, seuls 10 spécimens maximum de chaque genre sont conservés par station et par année de prélèvements. Ces spécimens sont sélectionnés par les techniciens de terrain de la DRIEE ; il s'agit la plupart du temps de spécimens bien préservés (peu dégradés par les manipulations lors des relevés de terrain). Il est important de noter que ce choix de réduire la collection à 10 spécimens maximum par genre peut entraîner un biais en ce qui concerne la diversité des espèces de *Baetis* réellement en présence dans un milieu donné.

⁵⁹ Données OPIE

Nous nous sommes rendus au laboratoire d'hydrobiologie de la DRIEE en mai 2017 afin d'effectuer des prélèvements sur les spécimens d'intérêt. Les spécimens doivent rester accessibles dans les collections de la DRIEE pour attester de la bonne réalisation des relevés de terrain chaque année. Nous avons néanmoins été autorisés à prélever une patte de chaque individu. Au total, nous avons prélevé une patte pour 71 spécimens.

❖ *Choix des stations et caractères écologiques DRIEE*

Il nous était nécessaire d'effectuer un choix réfléchi des stations de prélèvements d'où venaient les spécimens que nous avons étudiés. En effet, pour obtenir des informations sur des préférences écologiques plus spécifiques des éphémères, nous devions sélectionner des stations qui avaient chacune des paramètres physico-chimiques et hydrobiologiques différents (dans le cas contraire, il n'est pas possible de discriminer des taxons sur la base de leurs préférences écologiques).

Les données relatives à l'hydrobiologie pour les années 2014 et 2015 ont été fournies par Claire David, personnel DRIEE. Elles suivent le protocole de l'IBG-DCE : 12 points de relevés pour lesquels sont réalisées plusieurs mesures : nature de la berge (naturelle ou artificielle), recouvrement du substrat (superficie relative en %), dominance/marginalité de 12 différents substrats, vitesse du courant (scindée en 4 classes de vitesse en fonction de la mesure en cm/s) et la profondeur de l'eau. D'après les données issues du site web de la DRIEE⁶⁰, 22 stations présentent les deux types de relevés (hydrobiologiques et physico-chimiques).

Parmi ces 22 stations, cinq ont été sélectionnées car elles présentaient les plus fortes amplitudes de différences écologiques : "le Ru du Dragon à Longueville", "le Ru d'Ancoeul à Moisenay", "l'Orge à Sermaise", "le Grand Morin à Villeneuve-la-Lionne" et "la Vancouleurs à Villette". Ces stations appartiennent toutes à des unités hydrographiques (périmètres définis dans un SDAGE – Schéma directeur d'aménagement et de gestion des eaux) différentes (**Figure 45**).

⁶⁰ www.driee.ile-de-france.developpement-durable.gouv.fr/donnees-qualite-des-eaux-r1063.html

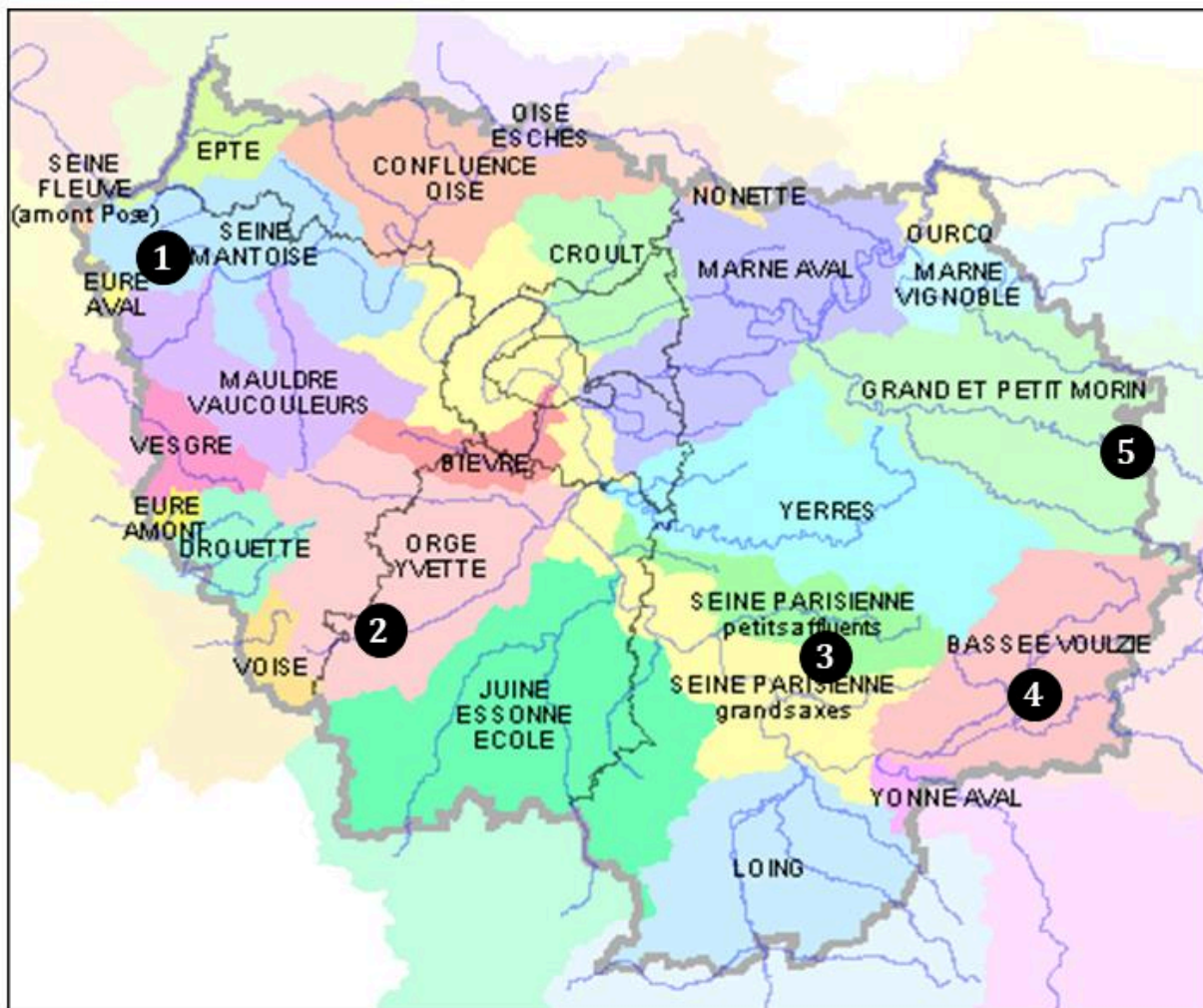


Figure 45. Carte des unités hydrographiques de l’Ile-de-France et localisation des cinq stations sélectionnées pour notre étude. (1) “la Vancoeurs à Villette”, (2) “l’Orge à Sermaise”, (3) “le Ru d’Ancoeuil à Moisenay”, (4) “le Ru du Dragon à Longueville”, (5) “le Grand Morin à Villeneuve-la-Lionne”.

De manière générale les données écologiques associées aux zones de relevés de la DRIEE présentent une définition plus fine que celles de la littérature. Certaines variables permettent également de valider les informations issues de la littérature (par exemple, “vitesse aux différentes profondeurs” correspondant au critère “Intensité du courant” de la base de données Fresh Water Ecology).

En ce qui concerne les critères physico-chimiques, nous en avons retenus cinq qui étaient pertinents pour notre étude : le pH moyen, les concentrations d’oxygène dissous, les

concentrations de nitrites, nitrates et phosphore. Ils définissent la capacité du milieu à accueillir des formes de vie grâce à une concentration d'oxygène nécessaire à la respiration branchiale, la présence de nitrites et nitrates inclus dans le cycle de l'azote, et la mesure du phosphore témoignant du lessivage des sols et de la pollution (eutrophisation notamment).

Le lien entre espèce de *Baetis* et préférences écologiques peut être fait après caractérisation moléculaire des spécimens prélevés depuis les collections de la DRIEE.

❖ *Obtention des séquences d'ADN*

Nous avons choisi de séquencer le gène mitochondrial universel du cytochrome c oxydase, sous-unité I (COI). La méthodologie a été la même pour les spécimens de la collection de référence et pour les spécimens issus des relevés de la DRIEE.

L'extraction d'ADN s'est faite à partir des spécimens entiers pour la collection de référence, et à partir des pattes pour les spécimens de la DRIEE. Nous avons suivi le protocole DNeasy® Blood & Tissue Kit (Qiagen).

L'amplification par PCR a suivi plusieurs protocoles utilisant les amorces universelles CO1-L1490 et CO1-H2198 (Folmer *et al.* 1994), et des températures d'hybridation et nombre de cycle variables (48°C et 50°C, et 35 à 40 cycles). L'efficacité des différents protocoles a été testée par des électrophorèses et des mesures de quantité et qualité d'ADN. Ces tests nous ont amenés à réévaluer la pertinence d'utiliser des amorces universelles. Pour les spécimens répondant moins bien au protocole, les amorces CO1-H2198 et CO1-L1490 ont été remplacées par les amorces spécifiques LepF1 et LepR1 (Hebert *et al.* 2004) ou les amorces Uni-Minibar F1 et R1 (Meusnier *et al.* 2008).

Le séquençage a été externalisé au laboratoire Eurofins. Le nettoyage des séquences (*forward* et *reverse*) a été réalisé à l'aide du logiciel Sequencher™ (v. 5.1) afin d'obtenir des séquences consensus.

5.1.3c Résultats et discussion

Sur les 35 espèces de la collection de référence, 14 espèces (21 échantillons) ont abouti à une amplification d'ADN suffisante pour effectuer un séquençage efficace.

Les séquences finalement obtenues étaient de qualité variable et seules 10 étaient exploitables. Elles correspondent à 8 espèces (sur les 14 envoyées au séquençage) : *Baetis rhodani*, *Centroptilum luteolum*, *Cloeon simile*, *Procloeon bifidum*, *Caenis horaria*, *Serratella ignita*, *Ecdyonurus venosus* et *Paraleptophlebia submarginata*. Ces séquences pourront enrichir la base de données GenBank, notamment celle d'*Ecdyonurus venosus*, pour qui aucune séquence de référence n'existe à l'heure actuelle.

Notre travail a donc permis de compléter les données moléculaires pour les éphémères d'Ile-de-France, et d'obtenir un ensemble de références pour l'identification des spécimens du genre *Baetis* provenant de la DRIEE.

En ce qui concerne les spécimens de la DRIEE, aucune amplification n'a abouti avec les amorces universelles CO1-L1490 et CO1-H2198. Certaines ont fonctionné avec les couples d'amorces LepF1-LepR1 et Uni-Minibar F1 et R1.

Sur 71 spécimens, 38 (53,52%) ont abouti à une amplification de qualité satisfaisante pour être envoyée au séquençage, avec des variations selon les stations (**Tableau 27**).

Tableau 27. Nombre de spécimens et nombre d'amplifications satisfaisantes pour chacune des cinq stations sélectionnées pour notre étude.

Station	Nombre de spécimens	Nombre d'amplifications
Le Grand Morin à Villeneuve-la-Lionne	15	8
L'Orge à Sermaise	10	4
Le Ru d'Ancoeul à Moisenay	8	7
La Vaucouleurs à Villette	18	6
Le Ru du Dragon à Longueville	20	13
<i>Total</i>	<i>71</i>	<i>38</i>

Cette étude n'a pas encore été finalisée, mais elle **ouvre d'intéressantes possibilités pour l'identification moléculaire des larves d'éphémères quel que soit leur stade, et pour l'amélioration des connaissances sur leurs préférences écologiques fines.**

La nouvelle clé d'identification finalement obtenue inclura les caractères morphologiques présents dès les premiers stades larvaires, les variables écologiques issues de la littérature et les données écologiques fines pour les larves d'éphémères d'Ile-de-France. Elle représentera une source fournie d'informations multi-critères pour les larves d'éphémères de la région.

Ce type de clé est novateur par rapport aux outils d'identification qui se basent sur des caractères morpho-anatomiques présents relativement tardivement dans le cycle de vie des insectes aquatiques (derniers stades larvaires et imago). Les clés d'identification incluant des critères relatifs à l'écologie des espèces permettent une détermination des spécimens dès le début de leur vie post-éclosion et tout au long de l'année.

5.2 Prise en compte de l'expérience des utilisateurs des clés d'identification

5.2.1 Pour la clé "Insectes" du Spipoll

L'analyse fine de l'utilisation par les volontaires du Spipoll (4.2.2) a confirmé que **certaines parties de la clé d'identification "Insectes" du Spipoll sont plus faciles à appréhender** par les utilisateurs. Certains types de caractères tels que **la forme, la taille et la coloration** sont plus intuitifs que d'autres pour les participants. Il s'agit de caractères morphologiques facilement visibles même sur des photographies, qui peuvent être floues ou de mauvaise qualité. On pourrait s'attendre à ce qu'une préférence pour ces types de caractères soit encore plus marquée au sein d'un public plus jeune (les enfants).

5.2.1a Descripteurs méritant une attention particulière

Une attention particulière devrait être portée aux descripteurs relatifs aux couleurs. En effet, ces descripteurs font appel à la vision subjective des volontaires, qui varie beaucoup selon les individus. Dans le meilleur des cas, ils doutent et sélectionnent alors les deux états (ce qui réduit la puissance de discrimination de l'identification), sinon, ils n'en sélectionnent qu'un et prennent le risque de faire une erreur (ce qui peut mener à une assignation finale de taxon incorrecte). D'autre part, les nuances de teintes sont délicates à décrire de façon précise dans un outil tel que la clé Spipoll, et les couleurs sont hautement variables sur les photos en raison des différences d'éclairage (environnement ensoleillé ou non) et des paramètres de l'écran sur lequel sont visionnées les photos.

Les utilisateurs sont sensibles à la qualité et la précision des caractères et états de caractères. Plus un descripteur possède d'états, plus il est difficile à appréhender pour les utilisateurs – surtout s'il s'agit d'une liste d'états peu tranchés, par exemple des variations de couleurs ou des gradients de taille. Les descripteurs binaires (qui ne proposent que deux états) posent moins de problèmes, c'est le cas par exemple des descripteurs de type "Présence/Absence".

Les descripteurs de type “Motifs” sont les plus difficiles à décrire finement, à expliquer et à bien illustrer de sorte à faciliter le choix des utilisateurs lors de l’identification. En effet, les états décrits peuvent paraître très similaires, par exemple pour les lépidoptères les états “Ailes orange avec points blancs” et “Ailes orange avec motif blanc en damier”. Les images qui accompagnent la description de ces états de caractères, qui ont pour but d’aiguiller l’utilisateur et de l’aider dans ses choix, ne remplissent pas toujours leur fonction avec succès. Le risque de confusion entre états est fort, et par là le risque de mauvaise identification également.

5.2.1b L'utilisateur et le doute

Les utilisateurs ont tendance à très peu douter et ce, même si l'on insiste sur la possibilité de le faire. J’ai pu également observer ce comportement lors d’un atelier réalisé dans le cadre de la Fête de la Science, avec des enfants de 5 à 12 ans. Il serait intéressant de mettre en place un protocole d’étude dans lequel on demande directement aux participants de détailler les raisons pour lesquelles ils favorisent un choix unique (quitte à mettre en péril toute la démarche d’identification) plutôt qu’un choix multiple en cas de doute. Avec les paramètres actuels, il suffit d’une erreur pour que l’identification du taxon final soit incorrecte.

Cette tendance qu’ont les participants à peu exprimer de doute au cours de leurs utilisations de la clé **doit être prise en compte pour l’amélioration** de l’outil. L’algorithme actuel autorise une erreur dans le chemin d’identification, c’est-à-dire qu’au besoin il ignore cette erreur et ne restreint pas en conséquence la liste des taxons compatibles restants. Il pourrait être judicieux d’autoriser, dans le futur, au moins deux erreurs dans le chemin d’identification (cela se paramètre simplement sous Xper³). Ainsi, la liste des taxons compatibles proposée à la fin du chemin d’identification serait certes peut-être plus longue, mais le taxon “correct” en ferait partie de façon certaine. À l’utilisateur ensuite de regarder

attentivement les photos associées à chaque taxon suggéré pour affiner son choix et/ou de proposer plusieurs taxons qui lui paraissent compatibles avec le spécimen de sa photo.

5.2.1c Pallier les défauts de la clé d'identification

La première action à entreprendre serait de **pondérer le "poids" des descripteurs**, en fonction du nombre de taxons qu'ils concernent, et du pourcentage de doutes et d'erreurs qui leur sont associés. En effet, par défaut tous les descripteurs ont un poids défini sur "3" (le poids peut être réglé entre 1 et 5). Si l'on abaisse le poids d'un descripteur entraînant beaucoup de doutes et d'erreurs de "3" à "1", Xper³ abaissera artificiellement son pouvoir de discrimination. Ainsi, le descripteur sera situé plus bas dans la liste des descripteurs proposés à l'utilisateur lors de l'identification. De même, il est possible de passer le poids d'un "bon" descripteur (très choisi et/ou entraînant peu d'erreurs) de "3" à "4" ou "5", ce qui le "remontera" dans la liste de descripteurs afin de guider l'utilisateur vers des caractères plus fiables.

Des informations complémentaires devraient également être ajoutées dans les fiches descriptives des taxons afin de mettre en garde les utilisateurs sur les confusions taxonomiques les plus fréquentes, et de pointer les caractéristiques nécessitant une attention particulière. Les messages pourront ressembler aux suivants : "Attention à ne pas confondre le taxon x et le taxon y", ou "Prenez la peine de bien lire la description des arrangements des motifs sur les ailes avant de sélectionner l'état de descripteur correspondant".

L'étude réalisée pour cette thèse (cf. 4.2.2) peut, si besoin, être complétée par des entretiens semi-directifs pour mieux comprendre en direct les actions des participants : pourquoi ils choisissent ce descripteur ou cet état à un instant donné, pourquoi ils expriment (ou non) un doute, quels sont leurs ressentis pendant la démarche d'identification, etc.

5.2.2 Pour la clé des larves d'odonates d'Ile-de-France

5.2.2a Remarques générales

La tendance qu'ont les participants à peu exprimer de doute au cours de leurs utilisations de la clé doit être prise en compte pour l'amélioration de l'outil. Il faut soigner au maximum le codage des descripteurs et états : peu d'états, formalisation claire des états que ce soit dans leur nom, leur précision (pas d'ambiguïté ni de recouvrement avec les autres états), le texte explicatif qui les accompagne ou encore leurs illustrations.

L'algorithme actuel de la clé n'autorise aucune erreur dans le chemin d'identification : dès qu'une erreur est faite dans le choix d'un état de descripteur, l'identification est faussée. Il pourrait être judicieux d'autoriser, dans le futur, une ou deux erreurs dans le chemin d'identification.

Comme dans le cas de la clé du Spipoll, les poids des descripteurs gagneraient à être ajustés en fonction des résultats obtenus par le test de la clé (cf. 4.2.3), ceci afin, encore une fois, d'améliorer la qualité et la fiabilité de cette clé d'identification.

5.2.2b Seconde version de la clé

Je suis en train de construire une seconde version de la clé d'identification des larves d'odonates d'Ile-de-France, à partir des résultats des tests de la clé. L'objectif est de prendre en compte le maximum d'éléments obtenus grâce au système Xperience pour pallier les défauts de la clé actuelle et proposer un nouvel outil qui limiterait les erreurs couramment faites par les utilisateurs, et qui serait ainsi plus fiable pour la détermination des larves d'odonates.

Cette nouvelle version de la clé des larves d'odonates sera publiée dans un avenir proche, ainsi que les clés d'identification des larves d'éphémères et de plécoptères, et la méthodologie et les résultats du test de la clé des larves d'odonates (**Annexe 9**).

Plusieurs types de modifications sont envisagées :

- **Fusion de descripteurs.** Les descripteurs "End of the abdomen" et "Abdomen shape" discriminent tous les deux entre les anisoptères et les zygoptères. La fusion de ces deux descripteurs annulerait les erreurs associées à "Abdomen shape", alors que "End of the abdomen" n'entraîne jamais d'erreurs.
- **Amélioration de l'iconographie** (ajout de schémas et de photos sous différents angles) **et des textes** explicatifs, pour les descripteurs concernant des structures morpho-anatomiques fines et difficiles à appréhender, par exemple "Type of mask" et "Pairs of legs 1 and 2"
- **Augmentation du poids des descripteurs** associés à aucune erreur, tels que "Dorsal spine on abdominal tergite 9" et "Antenna shape"
- **Diminution du poids des descripteurs** qui entraînent beaucoup d'erreurs, par exemple "Dorsal spine on abdominal tergite 7" et "End of the caudal gills".

5.3 Questionnaire et expérience des volontaires au Spipoll

Un autre moyen efficace de cibler les défauts d'un outil d'identification et les difficultés rencontrées par ses utilisateurs, est de se renseigner directement auprès des utilisateurs sur ces questions. C'est pourquoi au cours de ma thèse, et dans le cadre de la validation de mon diplôme de Mastère Spécialisé "Politiques Publiques et Stratégies pour l'Environnement" d'AgroParisTech, j'ai mis en place un **questionnaire en ligne adressé aux volontaires du programme Spipoll** (les "Spipolliens"). Ce sondage incluait plusieurs types de questions portant sur :

- L'utilisation de la clé d'identification "Insectes" du Spipoll et la satisfaction des utilisateurs
- L'expérience en tant que Spipollien
- Le rapport à la nature, l'engagement pour la préservation de l'environnement, et l'influence de la participation à des programmes de sciences participatives sur ces éléments.

5.3.1 Contexte

Environ la moitié des citoyens s'engage dans une quelconque activité de bénévolat (ou volontariat), c'est-à-dire une activité effectuée librement, non rémunérée en direction d'autrui (Clary & Snyder 1999). Cette activité peut être de nature diverse, par exemple : apprentissage aux enfants et adultes illettrés, aide aux activités ménagères, conseils pour les personnes en difficulté, garde malade... Les sciences participatives représentent une catégorie particulière d'activité de volontariat.

Les citoyens qui s'impliquent dans les programmes de sciences participatives n'obtiennent pas de compensation financière pour le temps qu'ils y consacrent. Leur implication dépend donc directement de leurs motivations, des avantages divers qu'ils espèrent obtenir à travers leur participation à ce type de projets. Clary & Snyder (1999) ont proposé une hypothèse appelée "*Matching hypothesis*", qui stipule que les volontaires s'engagent si l'activité de bénévolat répond à des besoins et envies particulièrement

importants pour eux, et qu'ils poursuivent leur activité bénévole seulement si elle s'accorde avec les motivations qui les ont poussés à faire du bénévolat initialement – autrement dit, si la satisfaction et les bénéfices obtenus grâce au bénévolat ont été en accord avec leurs motivations.

Ces mêmes auteurs ont mis en place un système comportant 6 fonctions, dans lesquelles peuvent se classer toutes les motivations des citoyens en faveur d'activités de bénévolat (**Tableau 28**).

Tableau 28. Description des 6 fonctions définies par Clary & Snyder (1999).

Fonction	Définition du concept : ce que recherche le bénévole	Exemple
Valeurs	Agir pour des valeurs importantes telles que l'humanisme	Je pense qu'il est important d'aider les autres
Compréhension	Apprendre des choses et/ou à exercer des compétences	Le bénévolat m'aide à apprendre grâce à des expériences directes et pratiques
Amélioration	Grandir et se développer psychologiquement	Le bénévolat me fait me sentir mieux dans ma vie
Carrière	Gagner de l'expérience utile pour sa carrière	Le bénévolat m'aide à me rapprocher d'une structure dans laquelle j'aimerais travailler
Social	Renforcer ses relations sociales	Des gens que je connais sont intéressés par ces activités bénévoles
Protection	Réduire des sentiments négatifs tels que la culpabilité ou résoudre des problèmes personnels	Le bénévolat est une échappatoire à mes propres problèmes

En particulier, les volontaires déclarent que les fonctions "Valeurs", "Compréhension" et "Amélioration" sont les plus importantes. Il existe logiquement des variations selon les groupes de gens interrogés (par exemple, les jeunes accordent plus d'importance à la fonction "Carrière" que les personnes âgées).

Dans le cas précis de l'écologie et de la conservation, des études sur plusieurs programmes de sciences participatives à travers le monde (par exemple : Gardiner *et al.* 2012, Domroese & Johnson 2016) montrent que les principales motivations des participants sont souvent liées au désir de participer à un projet de conservation de la biodiversité, à l'envie de contribuer à un projet de recherche, et à la curiosité de découvrir ou d'apprendre davantage sur les organismes ciblés par le programme. Cependant, l'attrait pour certains

aspects de l'activité plutôt que d'autres varie selon les programmes de sciences participatives.

Le programme Spipoll fonctionne depuis 2010, ce qui correspondait à sept saisons de pollinisation au moment de mon étude. Il était intéressant et utile de faire le point à ce moment de l'histoire du programme, et notamment d'obtenir des retours détaillés de la part des Spipolliens. L'objectif est d'en savoir davantage sur leurs motivations initiales pour la participation au programme, ce qui leur a plu ou posé des difficultés lors de leur participation, savoir si l'expérience a répondu à leurs attentes...

Je souhaitais répondre à ces trois questions principales :

- 1. Que nous apprennent les retours d'expérience des utilisateurs de la clé d'identification "Insectes" du Spipoll ?**
- 2. Quelles sont les motivations des citoyens pour les sciences participatives en général, et pour le Spipoll en particulier ?**
- 3. Quels ont été les bénéfices de leur participation au programme Spipoll ?**

Pour cela, j'ai mis en place un sondage sous forme de questionnaire écrit et accessible en ligne, **diffusé à toutes les personnes ayant participé au moins une fois au programme Spipoll.**

Cette partie du document présentera l'analyse détaillée de ce questionnaire. Le suivi des motivations, problèmes et satisfactions des participants aux programmes de sciences participatives est crucial pour pouvoir améliorer ces programmes, faire en sorte qu'ils répondent au mieux aux attentes des participants, et parvenir à maintenir intacte leur motivation pour les "fidéliser" dans le temps. On fait également l'hypothèse que si le volontaire a apprécié ses expériences dans les sciences participatives en rapport avec la conservation de la biodiversité, sa connexion avec la nature sera renforcée après sa participation, et que potentiellement le volontaire s'engagera plus fortement dans des actions liées à la protection de la nature.

5.3.2 Matériel et méthodes

Le sondage a été envoyé fin 2016 à **l'ensemble des adresses email associées à un compte Spipoll actif, soit 1283 personnes** ayant posté au moins une collection sur le site web du programme depuis son lancement en 2010.

J'ai réalisé le questionnaire grâce aux outils proposés par la plateforme Eval&Go®⁶¹. Il comprend **23 questions** affichées sur trois pages, ainsi qu'une page d'accueil et une page de fin du questionnaire.

Il est possible d'accéder à l'intégralité du questionnaire *via* le lien suivant (lien "test", les réponses données ne sont pas enregistrées) :

<https://app.evalandgo.com/s/?p=1&id=JTk1bSU5MWoIOUYIQUI=&a=JTk1aiU5NWklOTclQUQ=>

Les 5 pages du sondage sont également consultables en **Annexe 13**.

Le sondage était complètement anonyme ; cependant il a été indiqué aux répondants qu'ils pouvaient, s'ils le souhaitaient, contacter l'auteur à l'adresse mdelaunay@mnhn.fr pour toute question ou remarque.

Pour la conception des questions en rapport avec les motivations des participants, je me suis inspirée de questions incluses dans les sondages de certaines études récentes sur le même sujet (Clary & Snyder 1999, Domroese & Johnson 2016).

L'analyse des résultats au questionnaire a été faite, pour certaines questions, avec l'aide de l'outil "Rapport" intégré à la plateforme Eval&Go® ; pour d'autres questions, j'ai refait des graphiques avec le programme Microsoft Excel® dans l'optique de mettre en évidence certains résultats en particulier.

⁶¹ evalandgo.fr

5.3.3 Résultats et discussion

5.3.3a Taux de réponses

Les réponses de 271 personnes différentes ont été enregistrées entre le 17 novembre et le 1^{er} décembre 2016. **Le taux de réponse est donc de 21,12%**. 50 personnes (soit 18,5% des répondants) n'ont pas répondu à l'ensemble des questions du sondage et l'ont quitté avant la fin (**Figure 46**), réduisant la quantité de données obtenues pour certaines questions.

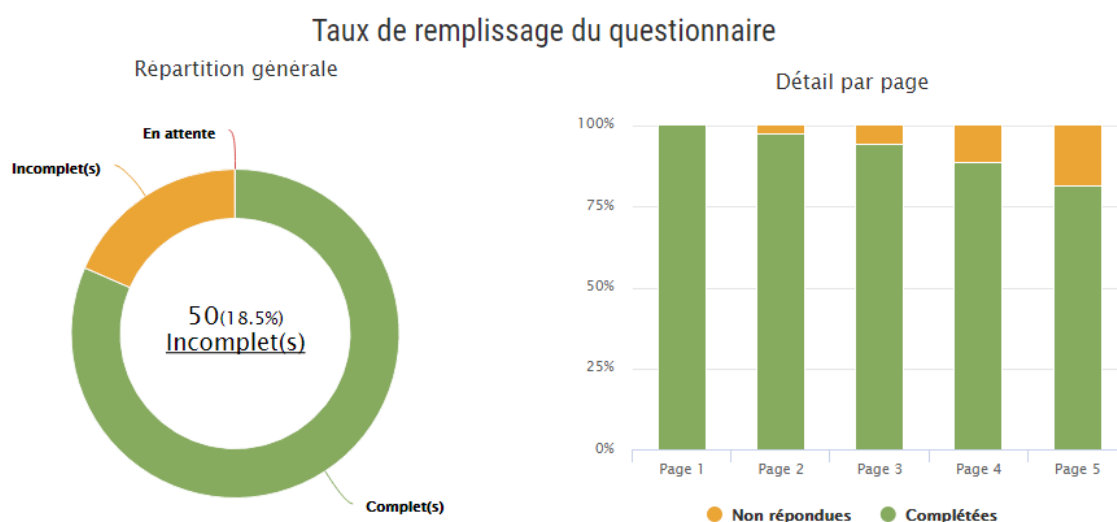


Figure 46. Taux de réponses au questionnaire. A gauche : répartition générale (sur 271 répondants) des questionnaires complétés et incomplets. A droite : répartition page par page.

5.3.3b Profils des Spipolliens

Les réponses s'équilibrent globalement entre les hommes et les femmes (**Figure 47**).

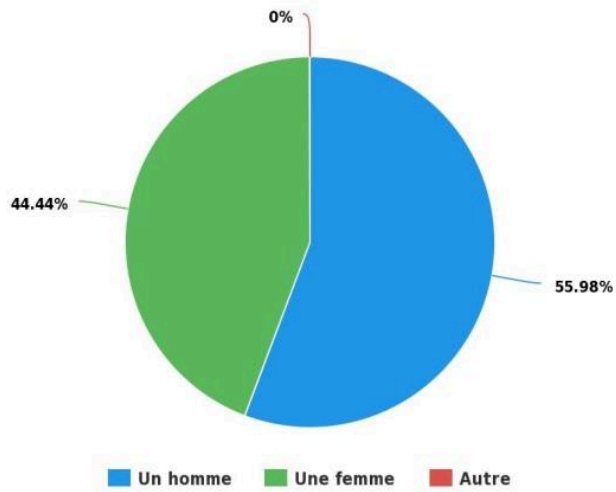


Figure 47. Proportion d'hommes et de femmes ayant répondu au questionnaire (sur 234 personnes).

La moyenne d'âge des répondants est d'environ 46 ans. Les personnes âgées de 40 à 59 ans sont majoritaires : leurs deux classes d'âge représentent à elles seules presque la moitié (46,75%) des répondants (**Figure 48**).

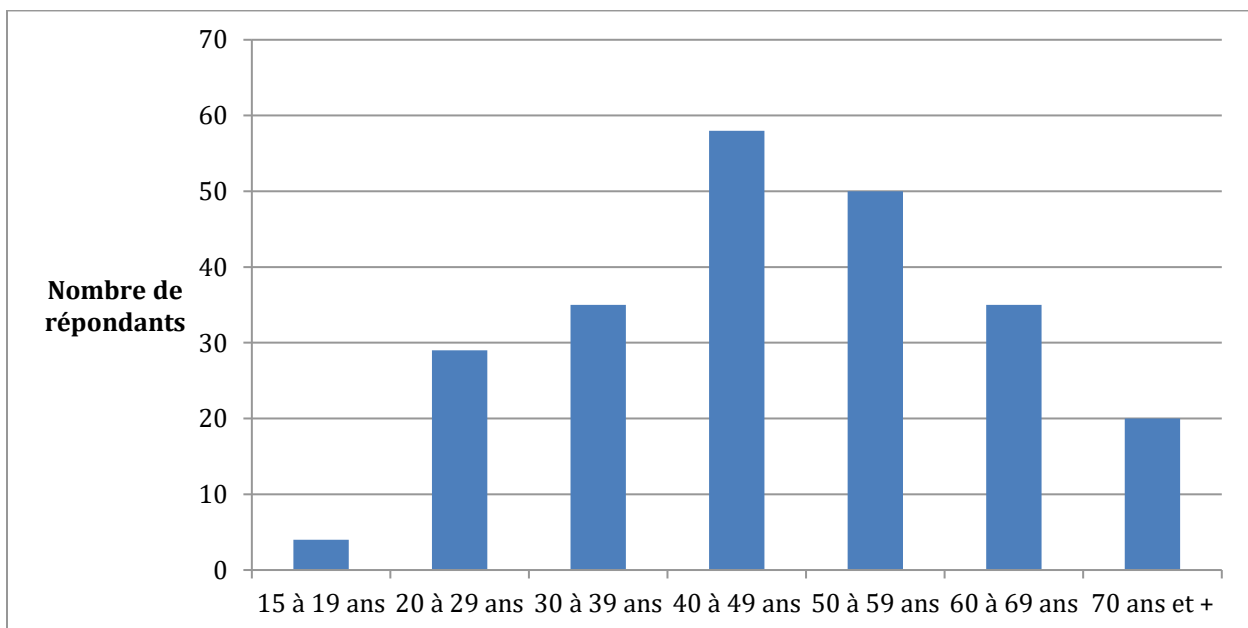
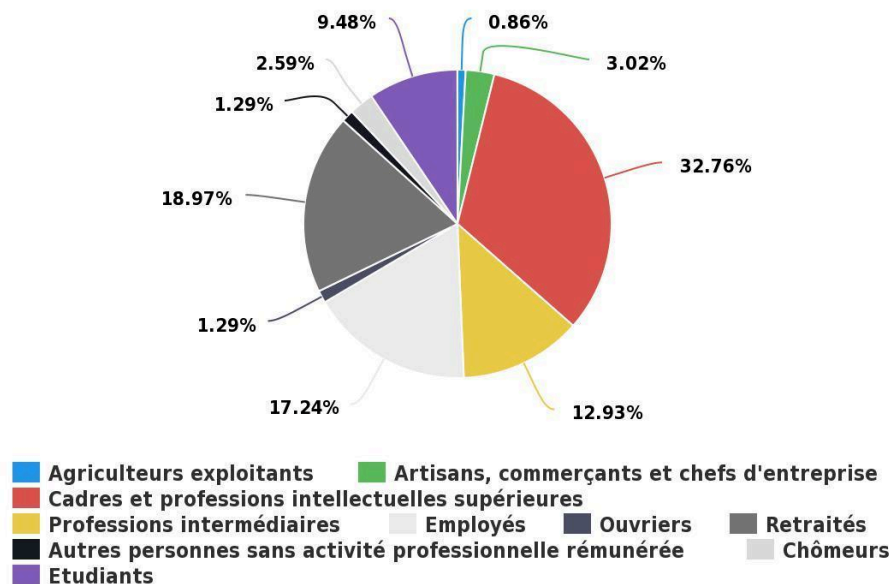


Figure 48. Nombre de répondants par classe d'âge (sur 231 personnes).



Les deux catégories socio-professionnelles dominantes parmi les volontaires sont les “cadres et professions intellectuelles supérieures” et les “retraités” (**Figure 49**).

Figure 49. Catégories socio-professionnelles des répondants (sur 232 personnes).

Les profils des participants au Spipoll ne diffèrent pas significativement de ceux d'autres programmes de sciences participatives comparables (relatifs aux sciences et à l'environnement) et récents (Guiney & Oberhauser 2009, Crall *et al.* 2013, Wright *et al.* 2015). Cependant, l'âge moyen des participants est légèrement moins élevé que dans la moyenne des autres programmes, où il dépasse le plus souvent les 50 ans. De même, les étudiants sont légèrement sous-représentés dans l'échantillon des répondants à notre questionnaire par rapport aux pourcentages habituellement observés (Domroese & Johnson 2016).

Les participants au Spipoll sont 70% à vivre en ville (plus de 2000 habitants), une proportion comparable à celle observée sur l'ensemble du territoire de la France métropolitaine (77,5%⁶²) (**Figure 50**).

⁶² Insee 2007

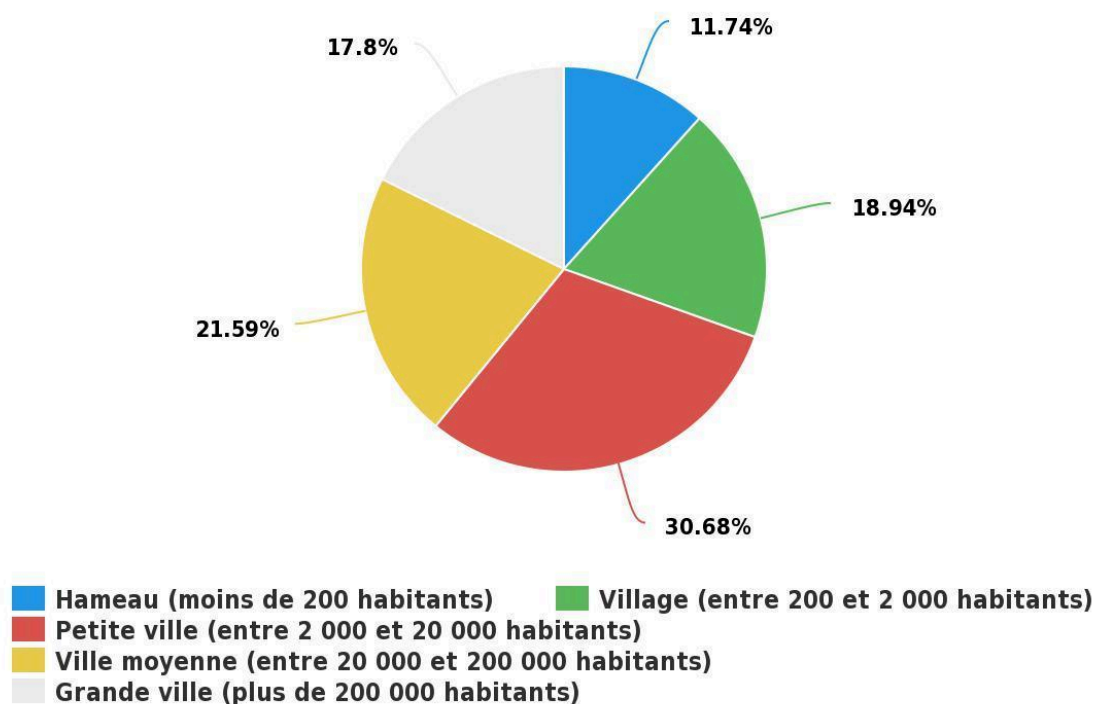


Figure 50. Lieux de vie des répondants (sur 232 personnes).

Comme observé globalement parmi les volontaires des sciences participatives en lien avec la biodiversité, **les Spipolliens se sentent en phase avec la nature**. Ils ont évalué leur lien avec cette dernière à 3,99/5 en moyenne (Figure 51).

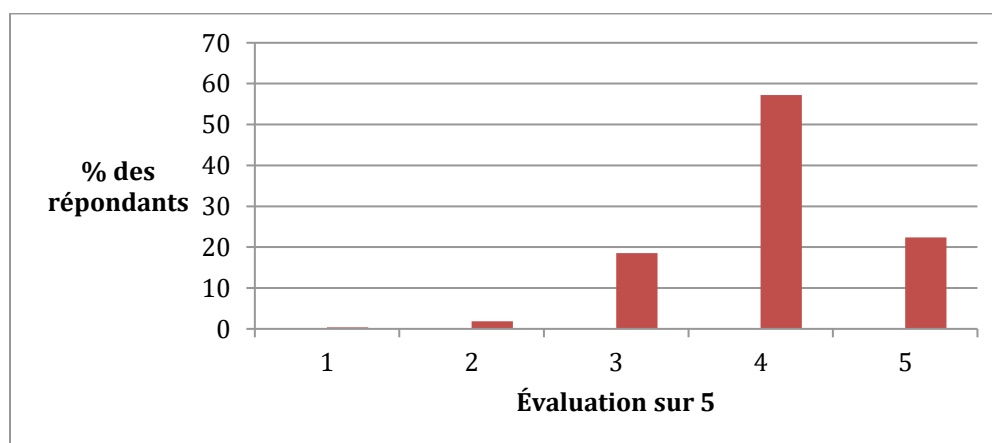


Figure 51. Réponses à la question "Sur une échelle de 1 à 5, à quel point vous sentez-vous en phase avec la nature ?" (sur 264 personnes)

Une part non négligeable des répondants indique **participer à au moins un programme de sciences participatives en dehors du Spipoll**. Pour la majorité, il s'agit d'autres programmes de la plateforme Vigie-Nature (**Figure 52**). On peut remarquer que les Spipolliens sont particulièrement attirés par les programmes relatifs aux insectes et aux oiseaux.

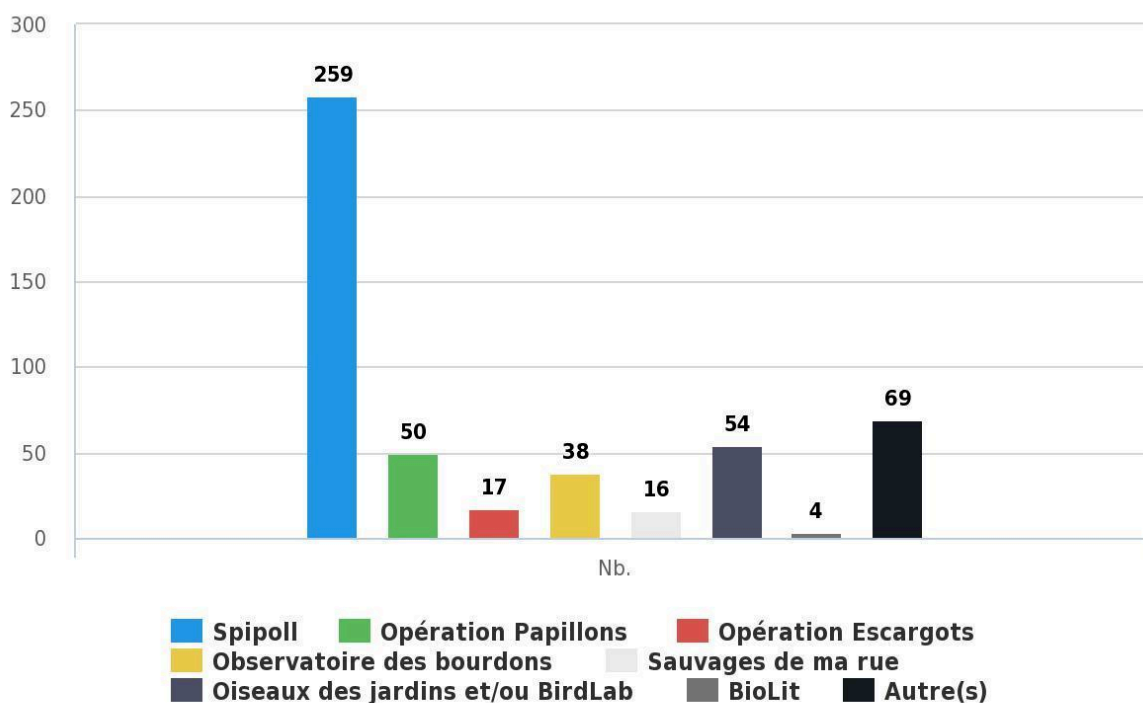


Figure 52. Nombre de réponses à la question à choix multiple “Quels sont les programmes de sciences participatives auxquels vous participez ou avez participé ?” (sur 264 personnes)

5.3.3c. Retours d'expérience du Spipoll

Ce questionnaire a représenté une occasion rare de connaître directement l'avis des volontaires sur leur participation au Spipoll. Les questions posées ont touché à tous les aspects du programme, depuis l'expérience de terrain jusqu'à l'utilisation de la clé d'identification “Insectes”, en passant par la qualité du site web. Certains de ces résultats ont

fait l'objet d'un poster présenté au congrès YNHM (Young Natural History scientists' Meeting) qui a eu lieu du 7 au 11 février 2017 au MNHN (**Annexe 14**).

❖ *Importance et fréquence de la participation au programme*

Un peu plus des deux tiers (68,15%) des répondants ont déposé moins de 10 collections sur le site du Spipoll, toutes années confondues (**Figure 53**). **La majorité des participants au programme n'est donc pas fidélisée. La part des Spipolliens "très actifs" (plus de 100 collections mises en ligne) représente 12,1% des répondants. Le programme se repose sur un petit "noyau" d'une trentaine de "fidèles" pour l'acquisition de la grande majorité de ses données.**

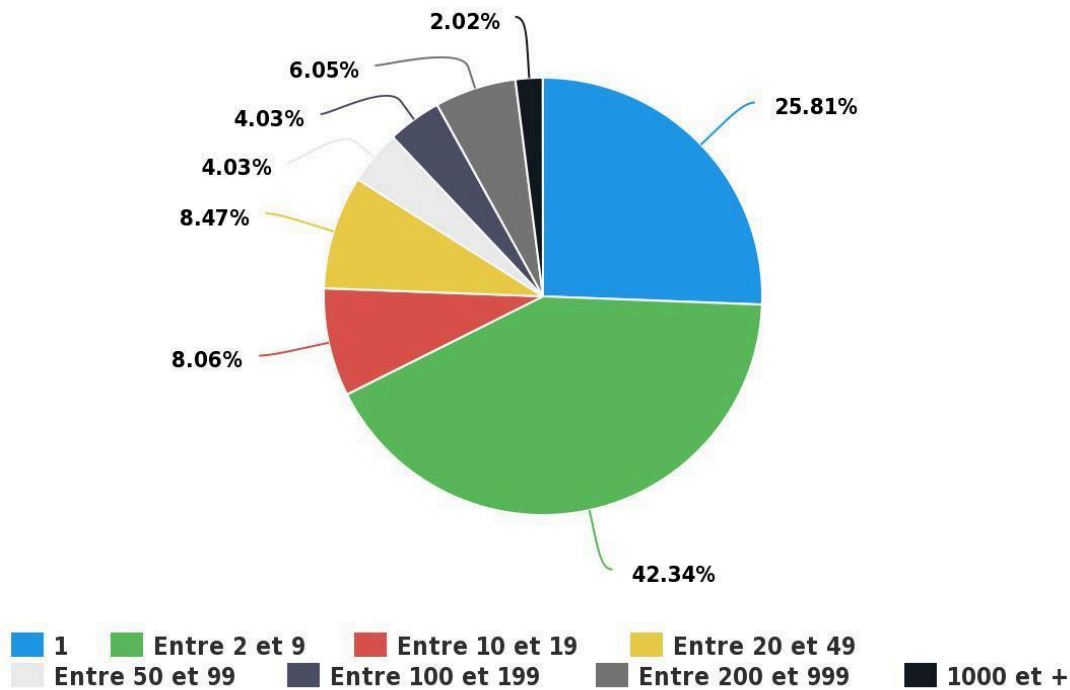


Figure 53. Réponses à la question "Combien de collections avez-vous déposées sur le site web du Spipoll au total ?" (sur 248 personnes)

On retrouve des résultats congruents si l'on s'intéresse au nombre de saisons de pollinisation auxquelles chaque répondant a participé (**Figure 54**). En effet, on remarque que

plus de la moitié (57,66%) des répondants n'ont participé au Spipoll qu'une seule année, ce qui est un bon indicateur du fait que cette part des Spipolliens n'est pas très fidèle au programme.

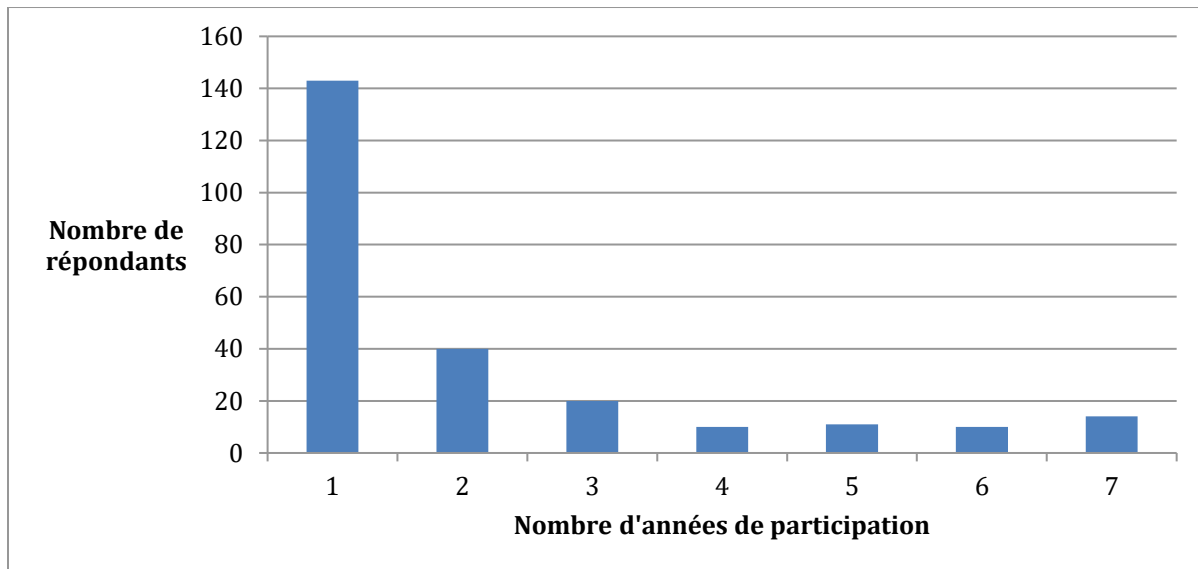


Figure 54. Proportion de répondants en fonction du nombre d'années de participation au programme Spipoll (sur 248 personnes)

❖ *Les motivations pour le programme Spipoll*

Qu'est-ce qui fait choisir un programme de sciences participatives plutôt qu'un autre ? Pour savoir ce qui plaît particulièrement aux volontaires dans le programme Spipoll, nous leur avons demandé quels aspects avaient constitué une motivation particulière pour ce programme. Les réponses récoltées sont présentées dans la **Figure 55**.

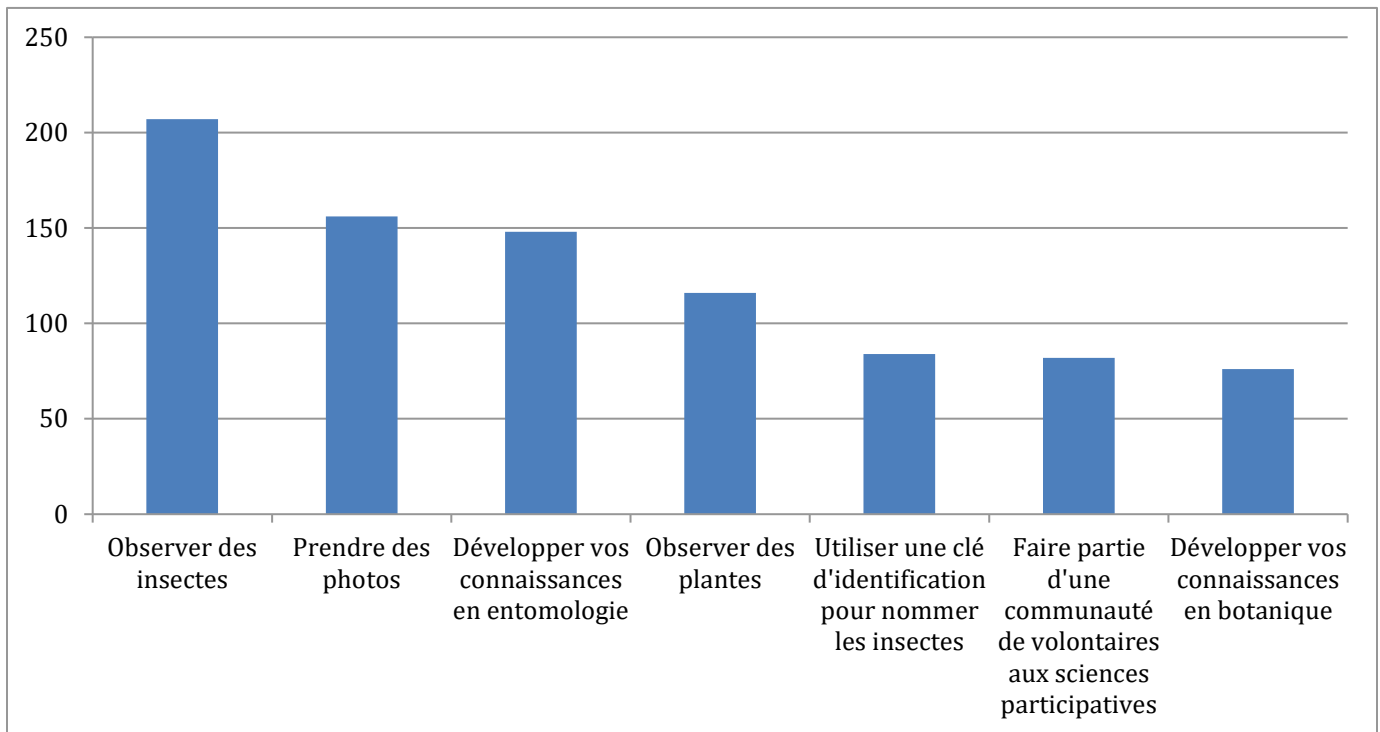


Figure 55. Réponses à la question “Parmi les facteurs suivants, quels sont ceux qui vous ont poussé(e) à commencer à participer au Spipoll ?” (sur 248 personnes).

Sans surprise, le fait que le programme se penche sur des **questions de pollinisation** est fondamental dans la motivation des participants. Même si le programme Spipoll cherche à mieux connaître les réseaux de pollinisation et donc, les interactions entre plantes et insectes, on peut observer que les volontaires sont plus intéressés par les aspects d’entomologie (83,47% d’entre eux souhaitent “observer les insectes” ; 59,68% “développer leurs connaissances en entomologie”) que par la botanique (pour 46,77%, “observer des plantes” est une motivation ; 30,64% souhaitent “développer leurs connaissances en botanique”).

Le fait de **prendre en photo des insectes et des plantes** attire particulièrement les Spipolliens (62,9% des répondants expriment ce souhait). L’originalité du Spipoll qui tient dans le fait de collecter les données *via* la prise de photographies sur le terrain, et pas de simplement noter des observations ou faire des comptages comme c’est le cas pour la plupart des programmes de sciences participatives, constitue donc une motivation importante.

❖ *Difficultés rencontrées avec le programme Spipoll*

Afin de mieux comprendre quelles difficultés rencontrées par les Spipolliens pourraient freiner leur participation au programme, nous leur avons posé la question suivante : “ Quel(s) facteur(s) pourrai(en)t vous faire stopper votre participation au Spipoll ou réduire la fréquence de votre participation ?”. Les possibilités de réponses étaient les suivantes :

1. “Le manque de temps”
2. “La difficulté à trouver des fleurs pour faire des sessions d'observation” (par exemple, si le participant habite une zone très urbanisée avec peu de parterres de fleurs)
3. “La faible quantité d'insectes pollinisateurs présents lors de vos précédentes sessions d'observation” ; selon les zones géographiques, le climat, la météo, l'environnement, etc. les insectes peuvent être observés à des densités très variables sur le territoire français
4. “La difficulté à identifier les insectes” ; ce point concerne en grande partie la qualité et la facilité d'utilisation de la clé d'identification du Spipoll
5. “Des difficultés techniques sur le site web du Spipoll (site lent, trop souvent en maintenance...)”
6. “La difficulté à comprendre le protocole et/ou les informations demandées lors de l'enregistrement des sessions d'observation”
7. “Le fait de ne pas avoir de ‘bon’ appareil photo” ; en effet, certains appareils (notamment les compacts) n'ont pas un assez bon zoom et/ou un déclencheur assez rapide pour pouvoir prendre en photo les insectes de façon optimale
8. “Le manque de retour auprès des Spipolliens sur les résultats scientifiques obtenus à partir des données récoltées” ; il est possible que certains bénévoles souhaitent plus de comptes rendus, de retour de la part des gestionnaires du Spipoll sur les résultats scientifiques obtenus grâce à leur travail
9. “La difficulté d'obtenir de l'aide de la part des responsables du Spipoll”
10. “Rien de tout cela”

Les résultats sont présentés sur la **Figure 56**.

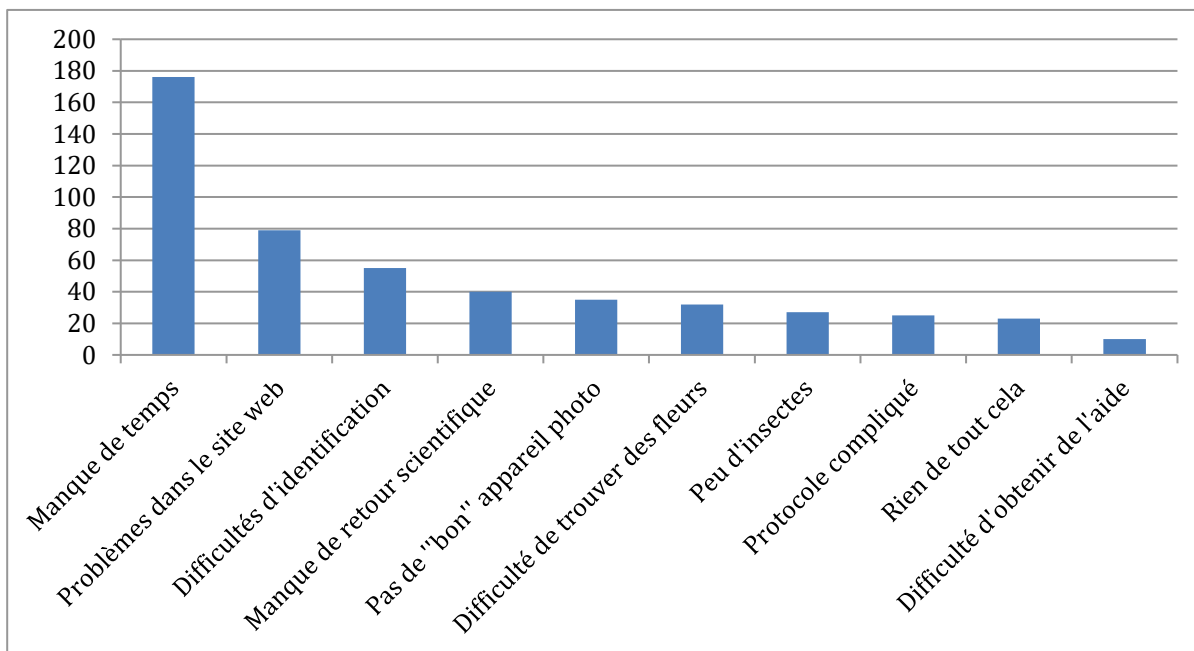


Figure 56. Réponses à la question " Quel(s) facteur(s) pourrai(en)t vous faire stopper votre participation au Spipoll ou réduire la fréquence de votre participation ?" (sur 248 personnes).

Le manque de temps est de loin la première raison invoquée pour l'arrêt et/ou la réduction de la fréquence de participation au programme (environ 70% des répondants). **Les difficultés d'identification des insectes d'après photo, donc potentiellement les problèmes dans l'utilisation de la clé "Insectes", arrivent en troisième position avec environ 21% des répondants exprimant ce souci.**

❖ *Expérience et satisfaction pour la clé d'identification*

La clé d'identification "Insectes" du programme Spipoll est un outil apprécié : presque deux tiers des volontaires l'utilisent pour plus de 50% des identifications qu'ils doivent effectuer sur le site web du programme (**Figure 57**). Plus étonnamment, ils sont quasiment la moitié à utiliser cet outil numérique pour décrire et nommer des insectes en dehors du contexte du Spipoll (**Figure 58**), c'est-à-dire pour identifier des spécimens qu'ils observeraient ou trouveraient dans la nature ou en ville lors de trajets ou promenades. Cela montre que cette clé d'identification est globalement appréciée et considérée comme un outil fiable, utile et pratique.

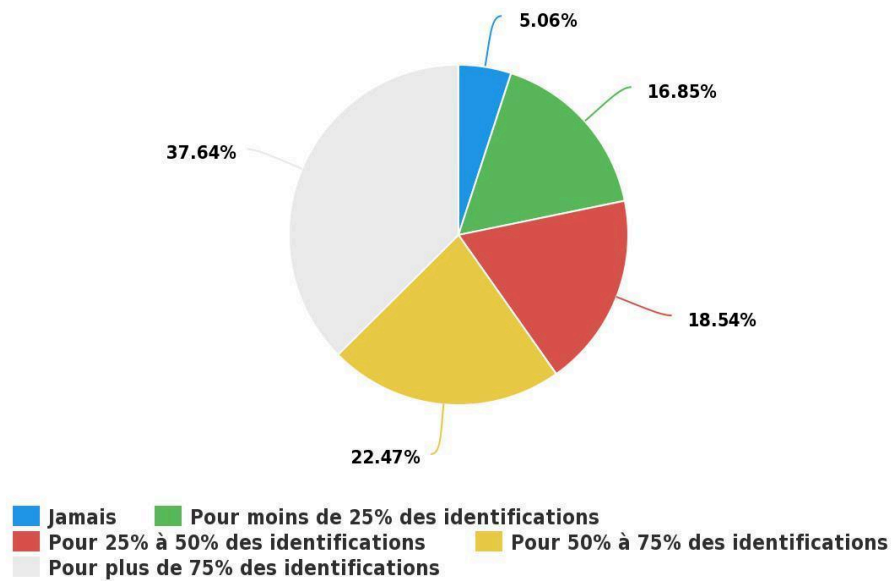


Figure 57. Réponses à la question “A quelle fréquence utilisez-vous la clé d'identification pour nommer les insectes présents sur vos photos du Spipoll ?” (sur 178 personnes).

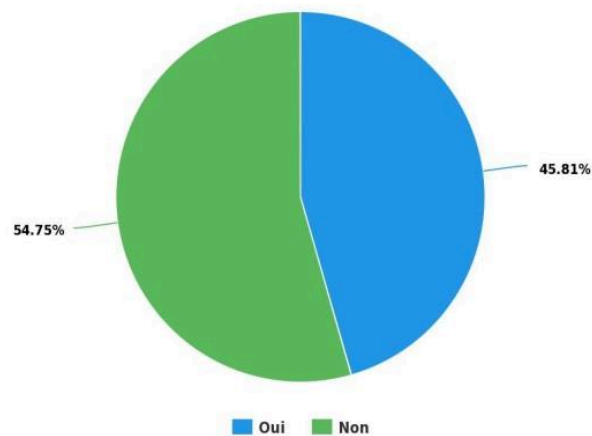


Figure 58. Réponses à la question “Vous arrive-t-il d'utiliser la clé d'identification "Insectes" du Spipoll pour un usage personnel en dehors du Spipoll ?” (sur 178 personnes).

La clé d'identification “Insectes” du Spipoll se veut très intuitive et abordable par tous les types de publics : enfants, retraités, amateurs, experts... Or l'analyse détaillée de cet outil a permis de mettre en évidence un certain nombre d'approximations, incohérences,

difficultés et blocages, allant parfois jusqu'à rendre impossible une assignation de taxon final correcte.

Les utilisateurs eux-mêmes font part de difficultés rencontrées dans l'utilisation de la clé d'identification (**Figure 59**) : bien qu'ils considèrent que l'outil est très utile (note moyenne de 4,27/5), ils trouvent globalement qu'il est **moyennement intuitif** (note de 3,64/5 pour la "facilité de prise en main") et que **les textes et explications pour les descripteurs et les états de descripteurs manquent de clarté** (note de 3,53/5).

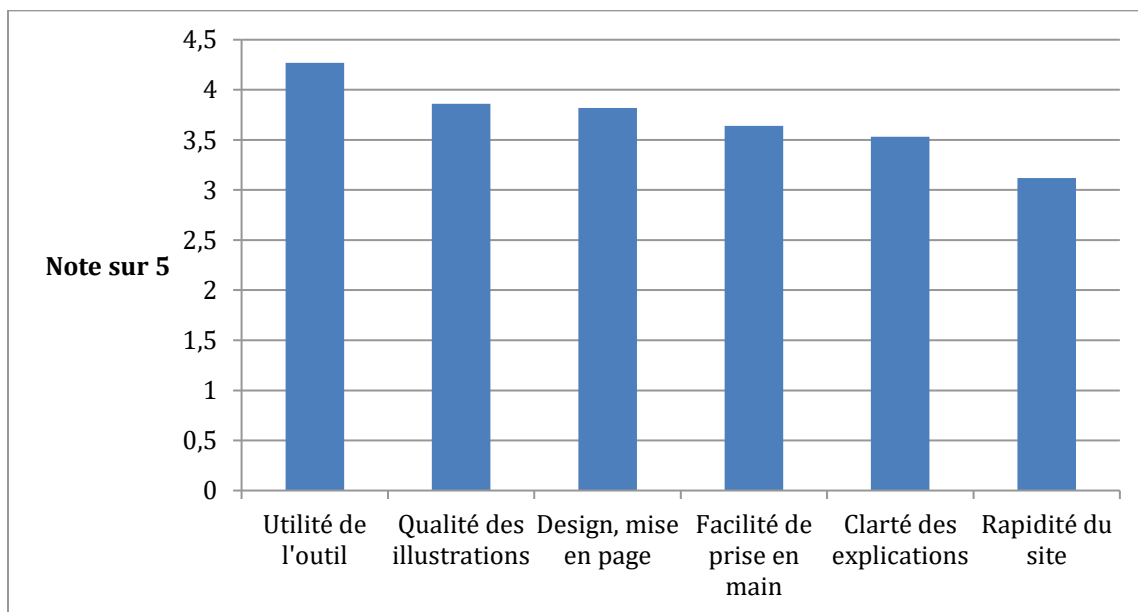


Figure 59. Résultats d'un sondage pour la question "Etes-vous satisfait(e) de la clé d'identification 'Insectes' du Spipoll ?". Les 165 répondants ont été invités à donner une note – de 1 à 5 – pour six composantes de la clé d'identification.

5.3.3d Motivations et expérience des sciences participatives

Nous avons demandé aux Spipolliens quelles étaient les **motivations initiales** qui les ont poussés à s'engager dans un programme de sciences participatives, et également, les bénéfices qu'ils ont pensé tirer de leur participation au programme Spipoll.

Les réponses proposées aux répondants dans le questionnaire étaient au nombre de 8 :

- Motivation 1 : Agir pour des valeurs qui vous paraissent importantes (humanisme, protection de l'environnement...). L'intitulé de cette motivation sera abrégé "Agir pour des valeurs".
- Motivation 2 : Prendre part à un projet scientifique (suivi d'un protocole, acquisition de données scientifiques, découverte des résultats...). Abrégé "Projet scientifique".
- Motivation 3 : Apprendre des choses (scientifiques, naturalistes, organisationnelles...) et/ou exercer des compétences (observer scientifiquement la nature, suivre un protocole précis...). Abrégé "Apprendre".
- Motivation 4 : Passer du temps avec vos proches (amis, famille...) en partageant une activité ensemble. Abrégé "Partager".
- Motivation 5 : Passer du temps à l'extérieur, au contact de la nature. Abrégé "Contact nature".
- Motivation 6 : Vous développer psychologiquement, vous sentir mieux dans votre vie. Abrégé "Développement".
- Motivation 7 : Gagner des compétences et/ou de l'expérience utiles pour votre carrière (pour trouver un travail, être plus performant(e) au travail...). Abrégé "Carrière".
- Motivation 8 : Passer le temps et/ou vous changer les idées (vous éloigner de vos propres problèmes). Abrégé "Passer le temps".

Les résultats sont présentés sur la **Figure 60**.

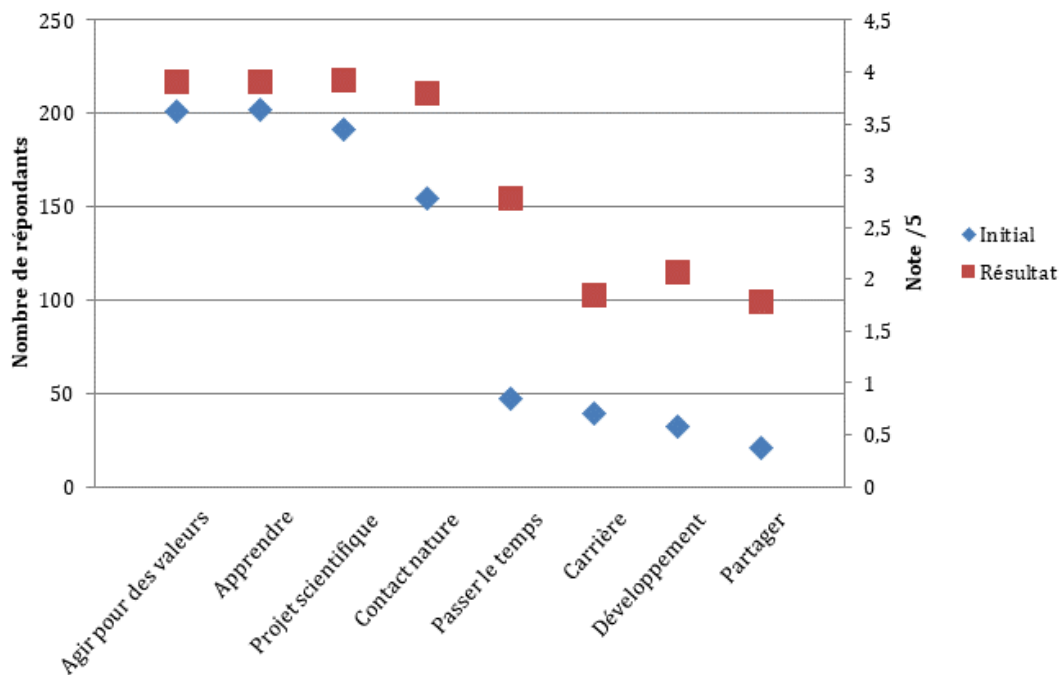


Figure 60. Adéquation entre motivations initiales et résultats après participation (sur 264 personnes).
 Losanges bleus : Réponses à la question à choix multiple “Quelles motivations vous ont poussé(e) à débiter votre implication dans un (des) programme(s) de sciences participatives ?” ; l’échelle est à gauche.
 Carrés rouges : Réponses à la question “Quels ont été, pour vous, les bénéfices de votre participation au Spipoll ?” ; l’échelle est à droite.

De façon globale, les résultats obtenus sont en accord avec la “*matching hypothesis*” de Clary & Snyder (1999) : **les bénéfices obtenus par les volontaires grâce à leur implication dans un programme de sciences participatives correspondent aux attentes qu’ils en avaient et aux motivations initiales qu’ils ont exprimées avant leur participation.** Le programme Spipoll apporte donc globalement satisfaction aux Spipolliens, qui ont de ce fait plus de chance de poursuivre leur engagement. On peut noter que l’écart entre la courbe des motivations initiales et celle des résultats est variable selon le type de motivation. Cet écart est **plus important pour les aspects “Passer le temps”, “Carrière”, “Développement” et “Partager”** ; cela signifie que les volontaires ont tiré dans ces domaines des bénéfices et de la satisfaction qu’ils n’attendaient pas forcément. C’est un point très positif, qui montre que les bénéfices de la participation au Spipoll sont plus larges que ceux qui étaient attendus.

5.3.3e Impact de l'implication dans le programme Spipoll

Pour achever cette étude sur les volontaires au programme Spipoll, j'ai cherché à savoir si la participation au programme avait un effet sur le mode de vie des Spipolliens, en particulier sur leur sensibilisation à la protection de l'environnement et de la biodiversité et sur les actions qu'ils entreprennent dans ce sens. Plus généralement, il est intéressant de se poser la question : **“Quelles répercussions l'engagement dans les sciences participatives peut-il avoir sur la façon dont les volontaires voient le monde et leur envie de s'impliquer plus fortement dans la protection de la nature ?”**

Pour obtenir des éléments de réponse, le sondage comportait une question sur les activités en rapport avec la conservation de la biodiversité effectuées régulièrement par les Spipolliens *avant* leur participation au programme, et *suite à* leur participation. Les activités proposées dans les réponses étaient les suivantes :

1. “Je suis attentif(ve) à la nature et à la biodiversité qui m'entourent”. Abrégé “Attentif à la nature”.
2. “Je me documente sur les problématiques environnementales”. Abrégé “Documentation”.
3. “Je m'implique dans des associations naturalistes ou de protection de la nature”. Abrégé “Associations”.
4. “J'assiste à des conférences dans le domaine de la biodiversité”. Abrégé “Conférences”.
5. “Je m'implique politiquement dans des actions de protection de la nature (manifestations, groupes de parole politiques, *sitting* dans les ZAD...)”. Abrégé “Politique”.
6. “Je cherche à passer plus de temps au contact de la nature”. Abrégé “Contact nature”.

Les résultats obtenus sont présentés sur la **Figure 61**.

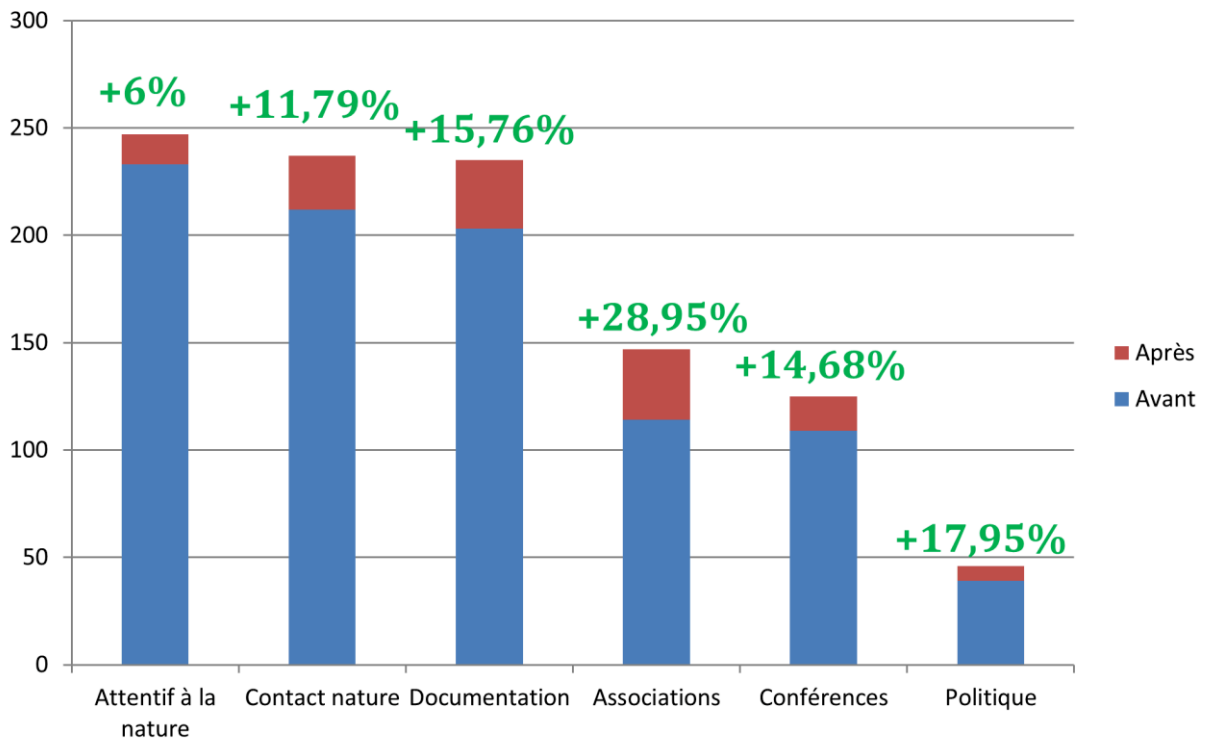


Figure 61. Réponses à la question “Les affirmations suivantes vous correspondent-elles, avant votre implication dans les sciences participatives, et après votre implication ?” (sur 248 personnes).

En bleu : nombre de répondants qui réalisaient déjà cette action avant leur participation au Spipoll

En rouge : nombre de répondant qui ne réalisaient pas cette action avant leur participation au Spipoll, mais qui la réalisent suite à leur participation.

En vert : pourcentage de progression du nombre de répondants qui ont commencé à réaliser chaque action après leur participation au Spipoll.

Il y a eu une **progression générale du nombre de personnes impliquées dans des actions en lien avec la protection de la nature** entre les chiffres avant participation au Spipoll et les chiffres obtenus après participation, comme observé de manière générale pour les programmes de sciences participatives (Lewandowski & Oberhauser 2017).

L’augmentation la plus forte s’est faite dans le domaine de l’engagement (bénévole) au sein d’associations naturalistes ou d’associations de protection de la nature telles que le WWF et GreenPeace. Il est possible que cette augmentation soit liée à l’existence de la communauté Spipoll et des échanges qui ont lieu entre Spipolliens, et entre Spipolliens et gestionnaires du programme. Les volontaires souhaiteraient ainsi poursuivre leur expérience de partage autour de questions sur la biodiversité.

La seconde augmentation la plus forte se situe au niveau des gens qui s'investissent politiquement pour la protection de la nature. Cette action est la plus "engagée" parmi les six actions citées ci-dessus, celle qui nécessite le plus de volonté et de persévérance ; c'est également un type d'actions qui a un effet particulièrement important auprès des décideurs politiques. Il est encourageant de constater que le programme Spipoll puisse donner l'envie et la motivation à des volontaires de sciences participatives de s'engager activement pour des causes importantes en faveur de l'environnement.

5.3.4 Pistes d'amélioration de la clé d'identification du Spipoll

En plus des éléments mis en évidence grâce à l'analyse de l'utilisation de la clé "Insectes" du programme Spipoll (voir le 5.2.1 pour les améliorations proposées), les réponses au questionnaire ont permis de mieux comprendre les difficultés rencontrées par les utilisateurs de la clé.

De manière générale, les Spipolliens ont à redire sur la clarté des explications associées aux descripteurs et états de descripteurs, et sur la facilité de prise en main de l'outil d'identification.

A la fin du sondage, une question ouverte permettait aux répondants de donner un avis libre sur la clé d'identification "Insectes". Leurs réponses peuvent se classer en plusieurs catégories :

- Problèmes généraux de **prise en main**
 - « Pas forcément évident de comprendre le fonctionnement au début »
 - « Cela reste très compliqué et on est parfois dans une impasse »
 - « Demande un peu de pratique pour ne pas faire fausse route »

- **Difficultés pour identifier correctement les spécimens** présents sur les photographies

« Je n'arrive pas à me servir correctement de la clé. [...] J'arrive souvent à une identification visiblement erronée. »

« Les critères auxquels il faut répondre [...] mènent souvent à des impasses »

« Je me trompe souvent et ça me déprime »

« Les réponses aux critères ne sont pas toujours aisées pour un novice. Ex : longueur des antennes, pilosité des insectes. »

- **Iconographie** à revoir

« Sur les photos qui illustrent chacun des critères de la clé, mettre le nom des taxons pourrait être utile »

« Que les photos mises soient le plus nettes possibles, avec un minimum de 5 photos systématiquement par taxon. Quelques fiches sont “en attente” et ne possèdent pas de photo. »

« Les photos proposées pour les taxons de la clé pourraient être revues et étoffées »

- **Textes explicatifs et “fiches taxons”** à étoffer

« Efforts de vulgarisation à réaliser sur certains termes comme “élytres” et “taxons”, qui ne facilitent pas la compréhension pour les néophytes »

« Compléter les descriptions de taxons. [...] C'est toujours intéressant d'avoir des renseignements sur les insectes que l'on photographie sur le terrain »

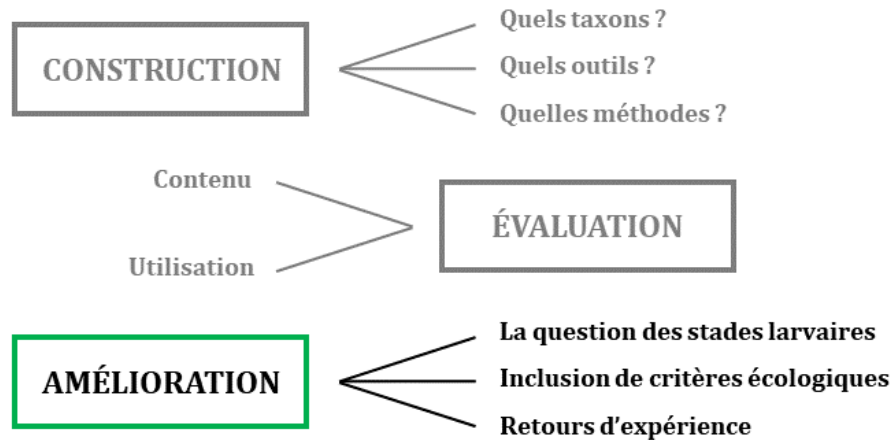
« La taille approximative de l'insecte pourrait être un plus »

« Vous êtes des pro de l'entomologie... pas nous. Soyez aimables d'être plus didactiques et vulgarisateurs en expliquant mieux les mots de votre langage. »

Ces remarques sont à prendre en considération pour l'amélioration de la clé d'identification “Insectes” du Spipoll. L'outil se veut intuitif mais ne l'est pas forcément, d'après les commentaires des Spipolliens. Comme souvent pour les clés d'identification numériques à accès libre, les textes et illustrations sont fondamentaux pour permettent des identifications correctes, et doivent être particulièrement réfléchis et travaillés.

Plusieurs Spipolliens ont fait des remarques positives sur l'utilité de la clé, en particulier lorsque leurs connaissances en entomologie sont très limitées. Ils ont également noté l'amélioration de l'outil d'identification au cours des années.

Le test de l'utilisation de la clé "Insectes" et les réponses au questionnaire envoyé directement aux Spipolliens ont permis de bien cibler les défauts et difficultés rencontrées lors de l'utilisation de l'outil. Les lignes directrices à suivre sont clairement identifiées et permettront une amélioration significative de la clé d'identification dans le futur.



Le travail présenté dans ce chapitre a permis d'évaluer et de confirmer la possibilité d'améliorer les clés d'identification numériques réalisées sous Xper3. L'intégration de critères autres que morpho-anatomiques (dans notre cas, relatifs à l'écologie), l'étude de données moléculaires pour préciser les déterminations et les critères écologiques, ainsi que le sondage des opinions des utilisateurs des clés, sont autant de moyens à exploiter afin d'obtenir des outils d'identification de la meilleure qualité possible.

Chapitre 6 – Discussion générale et conclusion

6.1 Bilan du travail de thèse

Ce travail de thèse a permis de construire des outils d'identification de l'entomofaune de France, sous une forme qui n'existait pas précédemment.

La **fiabilité** de ces clés d'identification a été d'une part vérifiée par des experts des différents groupes concernés (exactitude du codage des données dans les bases de connaissances) (cf. 3.2.2), et d'autre part testée par l'analyse de leur utilisation par différents types de publics, notamment pour mettre en évidence les points de blocage et les erreurs les plus couramment faites lors des identifications (cf. 4.2.3).

Pour rendre les clés les plus **intuitives** et faciles d'utilisation possible, l'importance de l'iconographie et des précisions textuelles a été soulevée à maintes reprises. En particulier, les utilisateurs sont demandeurs de textes explicatifs clairs et complets, notamment pour définir le vocabulaire relatif aux structures des insectes qui peuvent être difficiles à appréhender pour des non-spécialistes (cf. 5.3.4). De plus, les schémas et les photographies apportent beaucoup aux outils d'aide à l'identification numériques. Leur présence améliore fortement la facilité d'utilisation des clés et réduit les taux d'erreurs dans les déterminations – c'est le cas pour les clés d'identification à accès libre comme à accès simple. Chez les enfants, les clés d'identification basées sur des images entraînent moins d'erreurs de détermination que les clés d'identification comprenant uniquement du texte (Randler & Knappe 2007).

L'une des forces de la plateforme en ligne Xper³ est la possibilité de travail **collaboratif**: plusieurs co-créateurs de bases de connaissances peuvent y avoir accès librement pour les consulter et les modifier simultanément. Dans mon travail, la co-construction des clés avec des experts des groupes a été importante (cf. 3.2.2), mais nous aurions pu aller plus loin si les spécialistes avaient eu un accès direct aux modifications des bases de données. Pour des questions pratiques, j'ai préféré leur permettre de simplement consulter les clés d'identification, pour entamer un dialogue avec eux avant de réaliser des modifications que je pouvais suivre de façon certaine afin de ne pas "perdre le fil".

Xper³ rend les clés d'identification facilement **évolutives**. D'une part, les possibilités de modifications des clés permettent de corriger rapidement des erreurs, et de modifier facilement le contenu des clés pour suivre les connaissances scientifiques ou moduler la clé en fonction des paramètres souhaités – échantillonnage taxonomique, localisation, conditions environnementales... (van Spronsen *et al.* 2010).

D'autre part, dans un souci de **pédagogie**, nous prévoyons d'adapter les clés réalisées au cours de cette thèse à différents types de publics, notamment le grand public et les scolaires, comme cela a déjà été réalisé pour d'autres types de clés (par exemple, Dietz & von Helversen (2004) pour les chiroptères d'Europe). Les modifications incluront une simplification du vocabulaire employé, la pondération négative du poids ou la mise de côté des descripteurs les plus compliqués à observer, et une adaptation de l'iconographie – les images ont des effets importants chez les enfants (Randler & Knape 2007, Randler 2008). Du fait de la petite taille des larves d'insectes aquatiques et de la relative "difficulté" de détermination de ces taxons, les clés adaptées aux enfants ne permettront pas d'obtenir une identification jusqu'à l'espèce. L'objectif est principalement de mettre à disposition des outils ludiques et innovants, pour la découverte de l'entomofaune, la connaissance des principales structures morphologiques et des variabilités écologiques des insectes, et l'apprentissage de l'identification de l'entomofaune à l'aide d'un outil numérique spécialisé, dans une optique de **sensibilisation** à la diversité et à la fragilité de ces groupes d'animaux. En effet, l'apprentissage des méthodes d'identification des organismes vivants améliore la connaissance de la biodiversité et la compréhension des principes de l'écologie (Randler 2006, 2008).

Les bases de connaissances et clés d'identification Xper³ se veulent **accessibles** à tous. Leur présence en ligne et leur gratuité sont sans aucun doute de gros points forts. Cependant actuellement, pour accéder à une clé d'identification réalisée sous Xper³, il faut être en possession du lien URL de la clé fourni par les créateurs de cette clé. Pour le moment, seules quelques bases de connaissances sont accessibles *via* sur la page Wiki de Xper³. Les gestionnaires de la plateforme travaillent à la **création d'un portail listant tous les bases de connaissances réalisées sous Xper³** (si les créateurs des bases donnent leur accord pour la diffusion de leurs clés). Un tel portail augmentera fortement l'accessibilité des outils

d'identification Xper³ pour le plus grand nombre. Ce serait d'autant plus pertinent que beaucoup de clés d'identification sont difficilement accessibles car non publiées dans des revues scientifiques à comité de lecture ou dans des livres : elles peuvent être trouvées uniquement en ligne sur des sites web^{63,64} ou faire partie de la "littérature grise" (par exemple : Dietz & von Helversen 2004, Derolez *et al.* 2014). D'autre part, ce type d'initiative va dans le sens des principes FAIR recommandés par la Commission européenne⁶⁵, qui demande que les données, les codes et les autres résultats de recherche soient "Faciles à trouver, Accessibles, Interopérables et Réutilisables".

L'utilisation du système Xperience pour automatiser l'enregistrement des étapes de détermination de tous les utilisateurs d'une clé d'identification Xper³ donnée, est totalement nouvelle. Grâce à l'analyse des démarches d'identification des utilisateurs pour la clé du Spipoll et la clé des larves d'odonates réalisée au cours de cette thèse (cf. 4.2), j'ai pu mettre en évidence des tendances dans la façon dont ces derniers utilisent ces outils d'aide à la détermination des insectes :

- **Les utilisateurs doutent très peu dans leurs choix d'états de caractères**, malgré le fait que cette possibilité ait été clairement énoncée dans le cas de la clé des larves d'odonates (cf. 4.2.2c et 4.3.2b). Pour mieux comprendre ce comportement, il serait intéressant de questionner directement les utilisateurs sur les raisons qui les ont poussés à faire ces choix. Dans tous les cas, ce comportement majoritaire doit être pris en compte dans la création des clés d'identifications futures : les états de descripteurs doivent être revus afin que le "bon" choix unique soit fait le plus souvent, et les descripteurs entraînant peu ou pas de doute mais un fort pourcentage d'erreurs doivent être retirés ou leur "poids" doit être baissé.

⁶³ stroudcenter.org/macros/key

⁶⁴ www.discoverlife.org/20/q?guide=Ladybug

⁶⁵ Final report and action plan from the European Commission Expert Group on FAIR data (2018)

- **Certains descripteurs ou types de descripteurs sont plus souvent choisis par les utilisateurs et/ou entraînent moins d'erreurs.** Dans l'ensemble j'ai observé que les descripteurs relatifs à la taille des insectes, à leur forme et à leurs patrons de coloration, ainsi que les descripteurs offrant un choix binaire entre "présence" et "absence" d'une structure morphologique, étaient les plus faciles à appréhender pour les utilisateurs. À l'inverse, les descripteurs relatifs à des structures morphologiques de petite taille telles que les antennes et les pattes, entraînent plus d'erreurs (cf. 4.2.2c et 4.3.2b).

Parmi les animaux, les insectes n'ont pas la meilleure réputation auprès des sociétés occidentales, qui les considèrent souvent comme des nuisances (Howarth & Ramsay 1991, Kim 1993b). Pour cela et d'autres raisons telles que leur petite taille, il s'agit d'un groupe souvent ignoré du grand public (Kim 1993a). Beaucoup d'insectes sont considérés comme "laid", par exemple les larves d'insectes aquatiques. Seuls les espèces perçues de manière consensuelle comme "jolies" et/ou utiles (par exemple, les papillons et les abeilles mellifères) attirent une large sympathie de la part du grand public qui s'intéresse alors à leur conservation. En tant qu'humains, notre regard est plus facilement attiré par des formes de relative grande taille et par des couleurs vives. C'est l'une des raisons pour lesquelles les enfants sont plus sensibles aux espèces exotiques de couleurs vives et "inhabituelles" qu'à la biodiversité dite "ordinaire" qui les entoure (Ballouard *et al.* 2011).

Les analyses de l'utilisation de ces deux clés (Spipoll et odonates) étaient complémentaires. La base de données du Spipoll est de grande taille et ses utilisateurs sont nombreux. Le système Xperience permet de récolter rapidement des données sur l'utilisation de la clé d'identification "Insectes". Cependant l'étude de la clé des larves d'odonates, avec un panel de 24 testeurs, a permis de contrôler tous les paramètres de l'expérience, y compris le choix des spécimens à faire identifier. Ces spécimens avaient également l'avantage d'être réels et prélevés sur le terrain, contrairement aux photographies d'insectes que la clé du Spipoll permet d'identifier. Les conditions se rapprochaient donc plus de celles d'une étude impliquant des observations ou des relevés sur le terrain.

Malgré les nouvelles possibilités techniques pour l'identification des êtres vivants, la réalité dans la science aujourd'hui est qu'il y a toujours **besoin des clés d'identification à accès simple et des guides d'identification**. Beaucoup de publications actuelles en taxonomie contiennent des clés de détermination accompagnant la description ou la révision de taxons (Grand *et al.* 2018, Dos Santos *et al.* 2018, Razak *et al.* 2018, Urbano-Bonilla *et al.* 2018, Bauters *et al.* 2019). Ces outils d'aide à l'identification sont toujours largement utilisés par la communauté scientifique.

Les clés d'identification numériques peuvent être adaptées selon les besoins des différents publics :

- Le **public scolaire** recherche des outils **pédagogiques** si possible offrant une portabilité sur le terrain, pour l'**apprentissage** et la **sensibilisation** à la taxonomie et à l'écologie (Randler & Zehender 2006) ; une identification à l'espèce n'est pas requise.
- Le **grand public** et les naturalistes amateurs recherchent des outils **faciles à prendre en main**, offrant une détermination la plus précise possible (à l'espèce si c'est faisable en fonction des taxons). D'après Martellos & Nimis (2008) et leurs observations de l'utilisation du dispositif KeyToNature⁶⁶ (un ensemble d'outils numériques d'aide à l'identification utilisés dans les écoles et universités), la facilité d'utilisation des clés numériques et donc, la satisfaction des citoyens, augmentent lorsque les descripteurs considérés comme plus "faciles" par les utilisateurs sont placés en premier dans la liste des descripteurs proposés pour l'identification.
- Les **scientifiques entomologistes** et les **professionnels de terrain** tels que les techniciens de la DRIEE requièrent une **identification assez précise pour leurs travaux** et/ou le type de protocole qu'ils doivent suivre (l'identification au genre est suffisante pour les protocoles de suivis de populations dans les cours d'eau (IBG-DCE...)). Cependant, la détermination doit pouvoir se faire **efficacement et rapidement** ; s'il s'agit de clés d'identification, il est mieux que le contenu de leurs bases de connaissances soit réduit aux besoins des professionnels (échantillonnage

⁶⁶ www.keytonature.eu/wiki

taxonomique, localisation, type d'environnement...) pour un gain de temps. Les clés d'identification sous Xper³ peuvent être modulées pour suivre ces besoins. Dans le cas d'une utilisation professionnelle, ce sont des outils particulièrement intéressants par rapport aux ouvrages encore largement utilisés actuellement avec des clés d'identifications à large spectre et à accès simple tels que le Tachet *et al.* (2010).

De manière générale, le grand public apprécie les expériences interactives pour mieux comprendre la biodiversité. Lors de la participation à des programmes de sciences participatives, les volontaires sont particulièrement motivés par le fait de participer à une vraie expérience scientifique et de savoir leurs données utilisées pour faire avancer la science (Harnik & Ross 2003). De plus, l'utilisation d'outils pour l'identification des objets biologiques dans le cadre de programmes de sciences participatives peut modifier la perception de la nature pour les utilisateurs (Johnson *et al.* 2014, Bonney *et al.* 2016) (cf. 5.3.3e).

6.2 De nouvelles méthodes d'identification

6.2.1 Les logiciels d'identification automatique

Depuis le développement des technologies de reconnaissance visuelle, des logiciels ont vu le jour permettant d'indiquer à l'utilisateur directement l'identification finale du spécimen qu'il souhaite déterminer. Ces logiciels sont dits "**d'identification automatique**". Ils fonctionnent à partir de plusieurs systèmes d'identification automatisée, basée sur des caractères morphologiques (Gaston & O'Neill 2004). Il s'agit de *deep learning* (apprentissage profond), un ensemble de méthodes d'apprentissage automatique par les machines à l'aide de réseaux de neurones, qui rend très efficace la reconnaissance automatique des images ou des sons (Wäldchen & Mäder 2018).

Parmi ces systèmes utilisés pour l'identification d'organismes vivants, nous pouvons citer DAISY (Digital Automated Identification System) (Gauld *et al.* 2000), AIMS (Automatic

Identification and characterization of Microbial populationS) (Jonker *et al.* 2000), et ABIS (Automatic Bee Identification System) (Arbuckle *et al.* 2001).

Des logiciels ont été développés pour la reconnaissance automatique de certains groupes d'insectes, tels que des mites (Mayo & Watson 2006) ou des larves de plécoptères (Larios *et al.* 2008). Ces logiciels peuvent travailler à partir de photographies, en se focalisant par exemple sur des structures morphologiques particulières. Dans le cas de la reconnaissance automatique des larves de plécoptères, un logiciel "détecteur de courbes" analyse la photographie et recrée une image exploitable par le système de reconnaissance (**Figure 62**). Le programme DAIS (Developing an Automatic Insect Identification System), quant à lui, détecte la disposition des nervures sur les ailes des insectes (Yang *et al.* 2015).

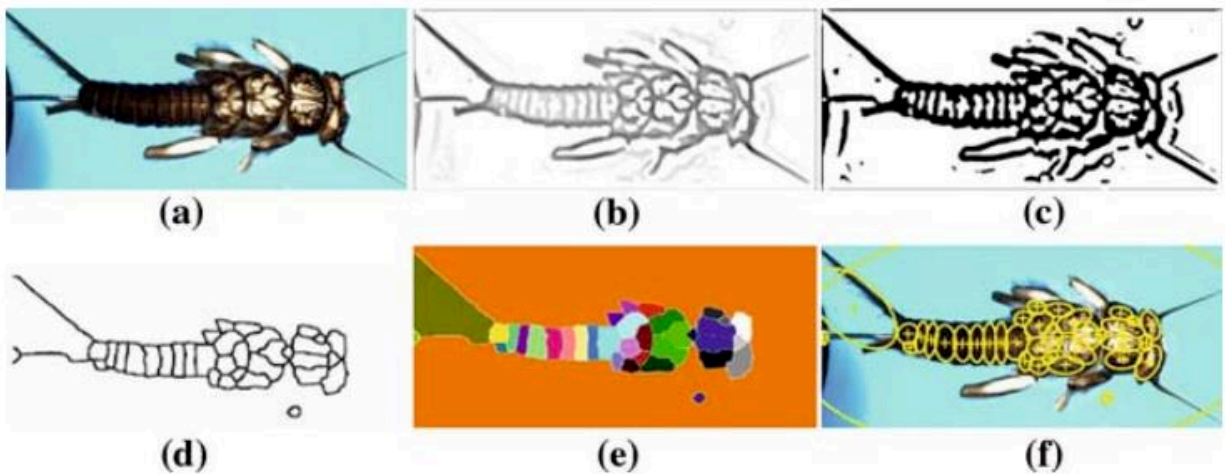


Figure 62. Régions définies par reconnaissance des courbes. (a) Image originale, (b) principales courbures, (c) image noir et blanc nettoyée. (d) Délimitations et (e) régions formées après transformation par le logiciel. (f) Régions finales créées par l'attribution d'une ellipse à chaque région. Figure issue de Larios *et al.* (2008).

En ce qui concerne l'identification automatique des végétaux, l'un des précurseurs a été l'application Leafsnap⁶⁷ qui permet, à partir de photographies, d'identifier 184 espèces américaines communes grâce à la forme de leurs feuilles (Kumar *et al.* 2012). Plus

⁶⁷ leafsnap.com

récemment, des chercheurs français ont développé **Pl@ntNet**⁶⁸ (Joly *et al.* 2016), dont l'utilisation des réseaux de neurones permet d'analyser plusieurs types de structures végétales sur la photographie d'un spécimen vivant ou mort. Pl@ntNet est une application disponible sur iOS et Android, et elle possède également une version web. Elle a pour ambition de mettre en place une base de données pour des dizaines de milliers d'espèces végétales du monde entier, notamment concernant la localité et l'écologie des plantes. Pl@ntNet est également un programme de sciences participatives. Il fait appel à un système participatif de vérification des identifications des spécimens présents sur les photos. Les volontaires ont également la possibilité de noter la qualité de chaque photo et d'écrire des commentaires quant à la détermination, la localité, l'écologie...

L'intelligence artificielle de ces systèmes peut apprendre par elle-même sans qu'il faille la coder de nouveau. Grâce à cette propriété, la fiabilité des identifications s'améliore au cours du temps. Dorénavant, pour certains programmes le taux d'erreurs de la machine dans les identifications est plus bas que celui des humains spécialistes des taxons concernés (Wäldchen & Mäder 2018). Le programme DAIS par exemple, permet d'obtenir 90% à 98% d'identifications correctes à l'espèce après seulement cinq répétitions pour un même spécimen (Yang *et al.* 2015). Les machines fonctionnant sur du *deep learning* améliorent rapidement leurs performances au cours des années, et les identifications par l'image deviennent vraiment fiables.

La fiabilité des identifications par le *deep learning* fait se poser la question de l'intérêt de continuer à créer et utiliser des clés de détermination, qui peuvent entraîner plus d'erreurs que les logiciels d'identification automatique. En effet ce type de systèmes peut faire gagner beaucoup de temps pendant les processus d'identification, ce qui est en particulier très utile lorsque l'utilisateur cherche à déterminer un grand nombre de spécimens. C'est le cas par exemple dans les suivis de populations sur le terrain, où des centaines d'individus sont récoltés et doivent être correctement identifiés pour alimenter les rapports sur la biodiversité dans divers localités et types d'habitats. Par exemple, les

⁶⁸ plantnet.org

techniciens de terrain de la DRIEE gagneraient beaucoup à pouvoir utiliser de tels systèmes, quand aujourd'hui encore ils utilisent des clés d'identification et guides sur papier.

Cependant ces systèmes d'identification automatiques ont des défauts à la fois techniques et scientifiques. L'un des grands défauts des systèmes actuels de réseaux de neurones est qu'**ils ne sont pas du tout explicatifs**. Le système va toujours proposer un résultat (nom de taxon dans la plupart des cas). Le système même d'identification automatique peut alors être remis en cause. Le premier problème est qu'en cas d'erreur du logiciel, il est difficile pour l'utilisateur de comprendre où se situe l'erreur et quelles sont les différences morphologiques entre le spécimen à identifier et la détermination effectivement indiquée par le logiciel. L'utilisateur n'a pas de référentiel, ni d'indications pour comprendre pourquoi la détermination proposée ne serait pas correcte. Le second problème est que malgré le fait qu'il soit très pratique et agréable d'obtenir une identification claire associée à un nom scientifique, même si elle est correcte, **l'utilisateur n'apprend rien** sur la morphologie et l'anatomie de son spécimen d'intérêt. La grande force des clés d'identification numériques à accès multiple est la **participation active** de l'utilisateur. Au cours de sa démarche d'identification l'utilisateur, tout en étant guidé, apprend à reconnaître les structures morfo-anatomiques et améliore sa compréhension de la biologie de son spécimen et des taxons qui lui sont proches. Cette **démarche pédagogique** est particulièrement intéressante à utiliser dans le cas d'apprentissage au grand public, par exemple au cours d'ateliers naturalistes, de formations, et dans les cursus scolaires pour les enseignements en sciences naturelles. Les logiciels d'identification automatique peuvent donc se révéler un considérable gain de temps dans les cas où l'utilisateur possède déjà de bonnes connaissances du groupe auquel appartient le spécimen à identifier. Ce type d'utilisateur (par exemple, un professionnel de terrain effectuant des suivis de populations ou des inventaires d'espèces) n'a pas particulièrement besoin de pédagogie, et il est souvent en mesure de détecter une erreur évidente de la part du logiciel. Dans la plupart des autres cas, les clés d'identification numériques à accès multiples telles que celles réalisées à l'aide de Xper³ sont plus pertinentes et utiles du fait de l'aspect pédagogique de leur "guidage pas à pas".

6.2.2 L'identification par séquençage rapide d'ADN

Il existe également depuis quelques années des systèmes portables de séquençage d'ADN, utilisables en laboratoire ou sur le terrain. C'est le cas du dispositif MinION⁶⁹ (Jain *et al.* 2015, Benítez-Páez *et al.* 2016), précurseur dans le domaine. Ces dispositifs sont petits et se connectent simplement en USB à un ordinateur. Pour obtenir une séquence, il suffit de déposer une goutte d'échantillon dans le système, et le séquençage se fait par nanopores sans besoin de manipulations supplémentaires. L'identification peut ensuite se faire par barcoding. Le dispositif MinION permet même de séquencer des petits génomes comme ceux des bactéries ou des virus (Check Hayden 2015). Ces machines permettent un **gain de temps considérable** pour les identifications à partir d'ADN, et peuvent être facilement utilisés dans le cadre de protocoles utilisant des techniques d'ADN environnemental (eDNA). Cependant, avec des taux d'erreur pouvant avoisiner les 40%, ce dispositif ne peut encore se substituer aux technologies de séquençage existantes (Laver *et al.* 2015).

D'autre part, à l'instar des outils d'identification automatique, **le séquençage automatique n'offre aucune pédagogie et l'identification est conditionnée à l'existence d'une base de données** conséquente, afin de relier les séquences ADN obtenues à des données existantes (barcoding). L'existence d'un spécimen type décrit et nommé est nécessaire à la base de toute détermination de taxon. Les collections d'histoire naturelle jouent pour cela un rôle majeur en tant que référentiels. Les collections sont plus accessibles que jamais grâce à leur numérisation massive partout dans le monde. 5 400 000 spécimens de plantes vasculaires issus de l'Herbier national situé au MNHN ont par exemple été numérisés, et les données ont été mises à disposition de la communauté internationale *via* le portail du GBIF (Le Bras *et al.* 2017). De même, il est impossible de se passer des experts de la taxonomie des différents groupes de plantes et d'animaux pour la description de nouveaux taxons et la vérification des identifications (Drew 2011).

⁶⁹ nanoporetech.com/products/minion

6.3 L'utilisation de critères non morphologiques dans les clés d'identification

Les bases de données sont incomplètes en ce qui concerne l'écologie des insectes aquatiques (Erman 1996, Vieira *et al.* 2006, Copplestone *et al.* 2008, Brown *et al.* 2009). Ces manquements réduisent le nombre d'informations autres que les caractères morpho-anatomiques que l'on peut intégrer dans les bases de connaissance et qui peuvent augmenter la puissance de détermination pour ces groupes.

S'il est possible de trouver des listes d'espèces le plus souvent exhaustives à l'échelle nationale ou régionale pour les ordres présents en Europe, les préférences écologiques "fines" des espèces sont mal connues. En particulier, les connaissances sont lacunaires pour les larves en ce qui concerne la localisation précise des populations, les préférences écologiques des espèces sur de petites échelles géographiques et les paramètres physico-chimiques et hydrobiologiques des microhabitats. La base de données en ligne "Freshwater Ecology" est une excellente initiative mais caractéristique de la difficulté à regrouper toutes les données existantes, y compris à très petite échelle. La difficulté d'accès aux individus liée à la vie aquatique, peut sans doute en partie expliquer ce manque de données. Les insectes aquatiques pâtissent à la fois du *taxonomic gap* – la différence entre le nombre d'espèces décrites et le nombre estimé d'espèces présentes sur la planète – et du *taxonomic impediment* – le manque de ressources financières et humaines pour l'étude de certains groupes d'êtres vivants (Dubois 2010).

Au cours de notre étude sur l'ADN des jeunes larves d'éphémères dans le but d'obtenir des données écologiques plus précises pour nos taxons d'intérêts en Ile-de-France (cf. 5.1.3), nous avons justement cherché à combler une partie des lacunes des bases de données. En dehors de l'utilisation de critères relatifs à l'écologie dans les clés d'identification, les larves d'insectes aquatiques, comme beaucoup d'autres groupes d'animaux et de végétaux, mériteraient d'être mieux connus en ce qui concerne leur distribution précise sur les territoires, leurs préférences écologiques, ou encore leur comportement. Plus les

connaissances sont fournies sur la biodiversité, plus nous sommes en mesure de la protéger efficacement (Watson 2005).

Le développement de programmes de **sciences participatives** associés à des **outils d'aide à l'identification performants**, représente un moyen efficace de combler les informations manquant dans les bases de données en ce qui concerne les "invertébrés" aquatiques, grâce à l'obtention de larges quantités de données sur des échelles de temps longues.

En France, il manque un programme entièrement dédié aux "macro-invertébrés" des eaux douces. Le Steli (Suivi temporel des libellules) de Vigie-Nature est limité aux odonates adultes et s'adresse aux professionnels de terrain uniquement. En parallèle des efforts collectifs fournis dans le cadre de la Directive européenne Cadre sur l'Eau (DCE) pour améliorer l'état physico-chimique et écologique des cours d'eau, il serait pertinent de mettre en place un programme accessible au grand public, à l'instar de l'*Anglers' Riverfly Monitoring Initiative*⁷⁰ ou du *Freshwater Watch*^{71,72} au Royaume-Uni. Ce programme gagnerait beaucoup à s'appuyer sur des clés d'identification telles que celles réalisées pendant ce doctorat, à l'image de la clé d'identification "Insectes" qui apporte globalement satisfaction aux volontaires du Spipoll. Un portage des clés sous Android *via* l'application BiodiversiClés⁷³, développée par des chercheurs et ingénieurs du MNHN, permettrait en plus d'avoir accès aux outils d'identification directement sur smartphone ou tablette et sans nécessité d'avoir une connexion internet, ce qui est idéal pour une utilisation sur le terrain.

De nombreuses données précises et locales sont récoltées chaque année par des bureaux d'études ou des structures publiques et privées d'étude et de protection de la biodiversité par exemple. Le problème est que ces bases de données de relative petite taille ne sont pas forcément mises en commun. Cela s'observe pour beaucoup d'autres groupes

⁷⁰ www.riverflies.org/rp-riverfly-monitoring-initiative

⁷¹ earthwatch.org.uk/get-involved/projects-activities/freshwater-watch

⁷² freshwaterwatch.thewaterhub.org/content/citizen-science

⁷³ play.google.com/store/apps/details?id=mnhn.inpn.biodiversicles&hl=fr

d'organismes. Des initiatives telles que le Cettia⁷⁴, un portail de saisie de données naturalistes à l'échelle de l'Ile-de-France, tentent de pallier ce problème, mais la quantité de données reste marginale et cantonnée à une aire géographique restreinte. En France, l'Inventaire national du patrimoine naturel (INPN)⁷⁵ est mis à jour très régulièrement pour inventorier les espèces présentes sur leur territoire, leur taxonomie et leur distribution. Au niveau mondial, il existe des exemples d'immenses bases de données mises à jour très régulièrement, regroupant taxonomie, localités, informations relatives à l'écologie et photographies, et accessibles à tous directement en ligne. Nous pouvons citer notamment le GBIF (Global Biodiversity Information Facility)⁷⁶ pour l'ensemble de la biodiversité connue, et en ce qui concerne des groupes plus précis, Mammal Species of the World⁷⁷ pour les mammifères, AviBase⁷⁸ pour les oiseaux, FishBase⁷⁹ pour les vertébrés non-tétrapodes ou encore Tropicos⁸⁰ pour les végétaux. Des applications mobiles ont également été créées pour permettre l'accès à de grandes bases de données et le partage entre citoyens d'observations naturalistes. C'est le cas de CarNat⁸¹ dont les données contribuent à l'INPN, ou encore de iNaturalist⁸² (van Horn *et al.* 2018) qui regroupe actuellement plus de 17 millions d'observations de plus de 200 000 espèces.

⁷⁴ cettia-idf.fr

⁷⁵ inpn.mnhn.fr

⁷⁶ www.gbif.org

⁷⁷ www.departments.bucknell.edu/biology/resources/msw3

⁷⁸ avibase.bsc-eoc.org

⁷⁹ www.fishbase.de

⁸⁰ www.tropicos.org

⁸¹ carnet-terrain-electronique.fr/carnat

⁸² www.inaturalist.org

6.4 Conclusion

Le travail réalisé au cours de ce doctorat a apporté de nouvelles possibilités pour l'identification de l'entomofaune :

- Les clés d'identification numériques interactives à accès libre réalisées sous Xper³ sont accessibles en ligne, fiables, faciles à prendre en main, collaboratives et évolutives. Elles présentent de nombreux avantages par rapport aux clés à accès simple. Ces clés facilitent l'identification de certains groupes de l'entomofaune de France.
- Les clés des insectes sous Xper³ peuvent répondre aux besoins de différents types de publics – professionnels de terrain, chercheurs entomologistes, volontaires des sciences participatives, grand public, public scolaire (élèves et enseignants) – grâce au fait que ce logiciel rend très simple la modification des bases de connaissances et des clés d'identification.
- Mon travail a souligné la place importante des citoyens dans l'identification de l'entomofaune, notamment *via* l'utilisation de clés d'identification numériques dans le cadre de programmes de sciences participatives.
- Le nouveau système Xperience a permis pour la première fois d'enregistrer et d'analyser en détail les démarches d'identification des utilisateurs de ce type de clés, et nous avons pu en déduire des tendances sur leur façon d'utiliser ces outils. Ces observations peuvent être prises en compte pour l'amélioration des clés d'identification afin de les rendre plus intuitives et efficaces pour la détermination des organismes.

Mon travail a utilisé une **approche intégrative** :

- Un protocole utilisé en **sciences humaines** (création et analyse de réponses à des questionnaires) a permis de compléter l'étude des démarches d'identification grâce à Xperience, en récoltant des informations sur la perception et la satisfaction des volontaires de sciences participatives en ce qui concernait leur utilisation de la clé "Insectes" du Spipoll.

- Des spécimens de **collections** ont été utilisés pour la description des taxons dans les clés (observation des structures morpho-anatomiques), mais également pour une étude moléculaire dans le but d'identifier des groupes d'espèces ne pouvant pas forcément être déterminés à l'aide de caractères morpho-anatomiques seuls.
- Le contenu de bases de données relatives à l'**écologie** des insectes aquatiques a été utilisé afin d'affiner la capacité de discrimination de clés d'identifications, en intégrant des critères écologiques dans les bases de connaissances Xper³ en plus des caractères morpho-anatomiques utilisés le plus couramment dans les clés d'identification.

L'utilisation des nouvelles technologies permettrait des **perspectives** intéressantes à ce travail :

L'un des principaux avantages des clés d'identification à accès simple par rapport aux clés d'identification numériques à accès libre est qu'elles peuvent être très précises et ciblées pour des conditions particulières (type de public, nombre d'espèces d'intérêt, type d'environnement...). Malgré la possibilité actuelle de sélectionner *a posteriori* des items, descripteurs et paramètres particuliers au sein d'une grande base de données utilisée pour la construction d'une clé d'identification numérique, le processus reste chronophage.

Nous pourrions imaginer le développement d'un système pour Xper³ qui ferait gagner du temps sur ce processus, notamment pour l'adaptation des clés à différents types de publics ou de circonstances. Un système utilisant le *deep learning* pourrait réaliser une recherche d'informations automatisée en ligne, par exemple depuis les bases de données du GBIF ou de iNaturalist, concernant les taxons intégrés dans la base de connaissances Xper³. Ce système pourrait ensuite trier les informations (distribution, préférences écologiques...) et les intégrer à la base de connaissances. Enfin, une fonctionnalité intégrée au système donnerait la possibilité à l'utilisateur de sélectionner facilement les critères qui l'intéressent, par exemple une liste réduite de taxons et/ou de descripteurs, une localité précise, ou encore des paramètres écologiques particuliers.

RÉFÉRENCES

- Alba V, Montemurro C, Sabetta W, Pasqualone A, Blanco A (2009) SSR-based identification key of cultivars of *Olea europaea* L. diffused in Southern-Italy. *Scientia Horticulturae* 123: 11–16.
- Ali A, Hussain A, Baloch M, Ahmed S (2014) Population trend of canola aphid, *Lipaphis erysimi* (Kalt) (Homoptera: Aphididae) and its associated bio-control agent, *Coccinella septempunctata* (Linnaeus) (Coleoptera: Coccinellidae) in different brassica lines. *Journal of Biology, Agriculture and Healthcare* 4: 19–22.
- Allsopp MH, de Lange WJ, Veldtman R (2008) Valuing insect pollination services with cost of replacement. Hector A (Ed). *PLoS ONE* 3: e3128.
- Arbuckle T, Schröder S, Steinhage V, Wittmann D (2001) Biodiversity informatics in action: identification and monitoring of bee species using ABIS. *EnviroInfo* 1: 425–430.
- Ballouard J-M, Brischoux F, Bonnet X (2011) Children prioritize virtual exotic biodiversity over local biodiversity. Somers M (Ed). *PLoS ONE* 6: e23152.
- Barbier EB (2011) Wetlands as natural assets. *Hydrological Sciences Journal* 56: 1360–1373.
- Barnaud G, Coïc B (2011) Mesures compensatoires et correctives liées à la destruction des zones humides – Revue bibliographique et analyse critique des méthodes. *Convention ONEMA Partenariat 2010 – Domaine Milieux Physiques – Action n°9*: 1–36.
- Bauernfeind E, Soldán T (2012) *The Mayflies of Europe (Ephemeroptera)*. Apollo Books. 781 pp.
- Baugnée J, Branquart E (2000) Clef de terrain pour la reconnaissance des principales coccinelles de Wallonie (Chilocorinae, Coccinellinae & Epilachninae). *Jeunes & Nature Association et Faculté Universitaire des Sciences Agronomiques de Gembloux*, 43 pp.
- Baugnée J, Branquart E, Maes D, Segers S (2011) Veldbepalingslijst voor de lieveheersbeestjes van België en Nederland (Chilocorinae, Coccinellinae, Epilachninae & Coccidulinae): herziene druk met larventabel. *Jeugdbond voor Natuur en Milieu (Gent), Jeunes & Nature Association (Wavre) i.s.m. het Instituut voor Natuur- en Bosonderzoek*, Bruxelles, 79 pp.
- Bauters K, Larridon I, Goetghebeur P (2019) A taxonomic study of *Scleria* subgenus *Hypoporum*: synonymy, typification and a new identification key. *Phytotaxa* 394: 1–49.
- Benítez-Páez A, Portune KJ, Sanz Y (2016) Species-level resolution of 16S rRNA gene amplicons sequenced through the MinION™ portable nanopore sequencer. *GigaScience* 5: 4.
- Berland L, Perrier R (1981) *La Faune de la France en tableaux synoptiques illustrés. 7, Hyménoptères*. Delagrave, Paris, 211 pp.
- Bertrand H (1928) *Les larves et nymphes des Dytiscides, Hygrobiides, Haliplides*. *Encyclopédie entomologique*. Lechevalier P (Ed). Paris, 360 pp.
- Blank M, Laine AO, Jürss K, Bastrop R (2008) Molecular identification key based on PCR/RFLP for three polychaete sibling species of the genus *Marenzelleria*, and the species' current distribution in the Baltic Sea. *Helgoland Marine Research* 62: 129–141.
- Blaxter ML (2004) The promise of a DNA taxonomy. *Philosophical Transactions of the Royal Society B: Biological Sciences* 359: 669–679.

- Bohmann K, Evans A, Gilbert MTP, Carvalho GR, Creer S, Knapp M, Yu DW, de Bruyn M (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution* 29: 358–367.
- Bonney R, Cooper CB, Dickinson J, Kelling S, Phillips T, Rosenberg K V., Shirk J (2009) Citizen science: A developing tool for expanding science knowledge and scientific literacy. *BioScience* 59: 977–984.
- Bonney R, Phillips TB, Ballard HL, Enck JW (2016) Can citizen science enhance public understanding of science? *Public Understanding of Science* 25: 2–16.
- Bonter DN, Cooper CB (2012) Data validation in citizen science: a case study from Project FeederWatch. *Frontiers in Ecology and the Environment* 10: 305–307.
- Le Bras G, Pignal M, Jeanson ML, Muller S, Aupic C, Carré B, Flament G, Gaudeul M, Gonçalves C, Invernón VR, Jabbour F, Lerat E, Lowry PP, Offroy B, Pimparé EP, Poncy O, Rouhan G, Haevermans T (2017) The French Muséum national d'histoire naturelle vascular plant herbarium collection dataset. *Scientific Data* 4: 170016.
- Brochard C, Groenendijk D, van der Ploeg E, Termaat T (2012) *Fotogids Larvenhuidjes van Libellen*. KNNV Uitgeverij, 323 pp.
- Brochard C, van der Ploeg E (2014) *Fotogids larven van libellen*. KNNV Uitgeverij, 243 pp.
- Brown LE, Céréghino R, Compin A (2009) Endemic freshwater invertebrates from southern France: Diversity, distribution and conservation implications. *Biological Conservation* 142: 2613–2619. doi: 10.1016/j.biocon.2009.06.009
- Buffagni A, Cazzola M, López-Rodríguez M, Alba-Tercedor J, Armanini D (2009) *Distribution and ecological preferences of European freshwater organisms. Volume 3. Ephemeroptera*. Pensoft Publishers, 254 pp.
- Bukkens S (1997) The nutritional value of edible insects. *Ecology of Food and Nutrition* 36: 287–319.
- Cadez N, Raspor P, Cock A, Boekhout T, Smith MT (2002) Molecular identification and genetic diversity within species of the genera *Hanseniaspora* and *Kloeckera*. *FEMS Yeast Research* 1: 279–289.
- Cane J, Tepedino V (2001) Causes and extent of declines among native North American invertebrate pollinators: detection, evidence, and consequences. *Conservation Ecology* 5.
- Chambers LE, Barnard P, Poloczanska ES, Hobday AJ, Keatley MR, Allsopp N, Underhill LG (2017) Southern hemisphere biodiversity and global change: Data gaps and strategies. *Austral Ecology* 42: 20–30.
- Chapin J, Brou V (1991) *Harmonia axyridis* (Pallas), the third species of the genus to be found in the United States (Coleoptera: Coccinellidae). *Proceedings of the Entomological Society of Washington* 93: 630–635.
- Chapman AD (2009) Numbers of living species in Australia and the world. *Australian Government: Department of the Environment, Water, Heritage and the Arts* 2: 84.
- Check Hayden E (2015) Pint-sized DNA sequencer impresses first users. *Nature News* 521: 15.
- Clary EG, Snyder M (1999) The motivations to volunteer: theoretical and practical considerations. *Current Directions in Psychological Science* 8: 156–159.
- Cohn J (2008) Citizen science: Can volunteers do real research? *BioScience* 58: 192–197.

- Comont R, Roy H, Harrington R, Shortall C, Purse B (2014) Ecological correlates of local extinction and colonisation in the British ladybird beetles (Coleoptera: Coccinellidae). *Biological Invasions* 16: 1805–1817.
- Copplestone D, Hingston J, Real A (2008) The development and purpose of the FREDERICA radiation effects database. *Journal of Environmental Radioactivity* 99: 1456–1463.
- Corvez A, Grand A (2014) Enabling comparisons of characters using an Xper2 based knowledge-base of fern morphology. *Phytotaxa* 183: 145–158.
- Costanza R, Arge R, Groot R De, Farber S, Hannon B, Limburg K, Naeem S, Neill RV (1997) The value of the world's ecosystem services and natural capital. *Nature* 387: 253–260.
- Costello MJ, Bouchet P, Emblow CS, Legakis A (2006) European marine biodiversity inventory and taxonomic resources: State of the art and gaps in knowledge. *Marine Ecology Progress Series* 316: 257–268.
- Couvet D, Jiguet F, Julliard R, Levrel H, Teyssedre A (2008) Enhancing citizen contributions to biodiversity science and public policy. *Interdisciplinary Science Reviews* 33: 95–103.
- Cowart DA, Pinheiro M, Mouchel O, Maguer M, Grall J, Miné J, Arnaud-Haond S (2015) Metabarcoding is powerful yet still blind: A comparative analysis of morphological and molecular surveys of seagrass communities. *PLoS ONE* 10: 1–26.
- Crall AW, Jordan R, Holfelder K, Newman GJ, Graham J, Waller DM (2013) The impacts of an invasive species citizen science training program on participant attitudes, behavior, and science literacy. *Public Understanding of Science* 22: 745–764.
- Daily G (1997) *Nature's services*. Island Press, Washington DC.
- Dallwitz M (1993) DELTA and INTKEY. In: Fortuner R (Ed), *Advances in computer methods for systematic biology: artificial intelligence, databases, computer vision*. The Johns Hopkins University Press, Baltimore, Maryland, 287–296.
- Dallwitz MJ, Paine TA, Zurcher EJ (2013) *Principles of interactive keys*. delta-intkey.com/www/interactivekeys.pdf.
- Dauguet P (1949) *Les Coccinellini de France*. Editions de l'entomologiste, Paris, 46 pp.
- Dayrat B (2005) Towards integrative taxonomy. *Biological Journal of the Linnean Society*: 407–415.
- Degrange C (1960) *Recherches sur la reproduction des Ephéméroptères*. Travaux du Laboratoire d'hydrobiologie et de pisciculture de l'Université de Grenoble 51: 7–193.
- Deguines N, Julliard R, de Flores M, Fontaine C (2012) The whereabouts of flower visitors: contrasting land-use preferences revealed by a country-wide survey based on citizen science. *PLoS ONE* 7: e45822.
- Deguines N, Julliard R, de Flores M, Fontaine C (2016) Functional homogenization of flower visitor communities with urbanization. *Ecology and Evolution* 6: 1967–1976.
- Derolez B, Orczyk N, Declercq S (2014) *Clé de détermination d'identification des coccinelles du Nord-Pas-de-Calais, version 4.1*. 84 pp.
- Desaegher J, Nadot S, Fontaine C, Colas B (2018) Floral morphology as the main driver of flower-feeding insect occurrences in the Paris region. *Urban Ecosystems* 21: 585–598.
- Dickinson JL, Shirk J, Bonter D, Bonney R, Crain RL, Martin J, Phillips T, Purcell K (2012) The current state of citizen science as a tool for ecological research and public engagement. *Frontiers in Ecology and the Environment* 10: 291–297.

- Dickinson JL, Zuckerberg B, Bonter DN (2010) Citizen science as an ecological research tool: challenges and benefits. *Annual review of ecology, evolution, and systematics* 41: 149–172.
- Dietz C, von Helversen O (2004) *Illustrated identification key to the bats of Europe*.
- Dijkstra KDB, Lewington R (2006) *Field guide to the dragonflies of Britain and Europe: including western Turkey and north-western Africa*. British Wildlife Publishing, Dorset, 320 pp.
- Dommanget J-L (1998) *Les libellules et leurs habitats: caractéristiques générales, éléments de gestion et de restauration*. Société Française d'Odonatologie.
- Domroese MC, Johnson EA (2016) Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project. *Biological Conservation* 208: 40–47.
- Doucet G (2010) *Clé de détermination des exuvies des odonates de France*. Société Française d'Odonatologie, 64 pp.
- Drew LW (2011) Are we losing the science of taxonomy? *BioScience* 61: 942–946.
- Dubois A (2010) Zoological nomenclature in the century of extinctions: priority vs 'usage'. *Organisms Diversity & Evolution* 10: 259–274.
- Dubuisson J-Y, Bouquin S, Vignes-Lebbe R (2018) Outil d'identification informatisé : Xper et orchidées de France métropolitaine. *L'Orchidophile* 219: 311–318.
- Dudgeon D, Arthington AH, Gessner MO, Kawabata Z, Knowler DJ, Lévêque C, Naiman RJ, Prieur-Richard A, Soto D, Stiassny MLJ, Sullivan C (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological reviews of the Cambridge Philosophical Society* 81: 163–182.
- Dupont P (2010) *Plan national d'actions en faveur des Odonates*. Office pour les insectes et leur environnement, Société Française d'Odonatologie, Ministère de l'Ecologie, de l'Energie, du Développement durable et de la Mer, 170 pp.
- Edwards M, Morse DR (1995) The potential for computer-aided identification in biodiversity research. *Trends in Ecology & Evolution Reviews* 10: 153–158.
- Erman N (1996) Status of aquatic invertebrates. *Sierra Nevada ecosystem project: final report to Congress* 2: 987–1008
- Farnsworth EJ, Chu M, Kress WJ, Neill AK, Best JH, Pickering J, Stevenson RD, Courtney GW, Dyk JK, Ellison AM (2013) Next-Generation Field Guides. *BioScience* 63: 891–899.
- Follett R, Strezov V (2015) An analysis of citizen science based research: Usage and publication patterns. *PLoS ONE* 10: e0143687.
- Folmer O, Black M, Hoeh W, Lutz R, Vrijenhoek R (1994) DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. *Molecular Marine Biology and Biotechnology* 3: 294–299.
- Fontaine B, van Achtenberg K, Alonso-Zarazaga MA, Araujo R, Asche M, Aspöck H, Aspöck U, Audisio P, Aukema B, Bailly N, Balsamo M, Bank RA, Belfiore C, Bogdanowicz W, Boxshall G, Burckhardt D, Chylarecki P, Deharveng L, Dubois A, Enghoff H, Fochetti R, Fontaine C, Gargominy O, Lopez MSG, Goujet D, Harvey MS, Heller KG, van Helsdingen P, Hoch H, de Jong Y, Karsholt O, Los W, Magowski W, Massard JA, McInnes SJ, Mendes LF, Mey E, Michelsen V, Minelli A, Nafria JMN, van Nieukerken EJ, Pape T, de Prins W, Ramos M, Ricci C, Roselaar C, Rota E, Segers H, Timm T, van Tol J, Bouchet P (2012b) New species in the old world: Europe as a frontier in biodiversity exploration, a Test Bed for 21st century taxonomy. *PLoS ONE* 7: e36881.

- Fontaine B, Perrard A, Bouchet P (2012a) 21 Years of Shelf Life Between Discovery and Description of New Species. *Current Biology* 22: R943–R944.
- Fox R (2013) The decline of moths in Great Britain: a review of possible causes. *Insect Conservation and Diversity* 6: 5–19.
- Freitag A, Pfeffer MJ (2013) Process, not product: investigating recommendations for improving citizen science “success.” *PLoS ONE* 8: e64079.
- Gardiner MM, Allee LL, Brown PM, Losey JE, Roy HE, Smyth RR (2012) Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. *Frontiers in Ecology and the Environment* 10: 471–476.
- Gargominy O, Ripken TEJ (2011) Une collection de référence pour la malacofaune terrestre de France. *MalaCo Hors Série*: 1–108.
- Gaston KJ, O’Neill MA (2004) Automated species identification: why not? *Philosophical Transactions of the Royal Society of London Series B: Biological Sciences* 359: 655–667.
- Gauld I, O’Neill M, Gaston K (2000) Driving Miss DAISY: the performance of an automated insect identification system. In: Austin A, Dowton M (Eds), *Hymenoptera: evolution, biodiversity and biological control*. Collingwood, 303–312.
- Gerke N, Tiedemann R (2001) A PCR-based molecular identification key to the glochidia of European freshwater mussels (Unionidae). *Conservation Genetics* 2: 285–287.
- Geslin B, Gauzens B, Baude M, Dajoz I, Fontaine C, Henry M, Ropars L, Rollin O, Thébault E, Vereecken N (2017) Massively introduced managed species and their consequences for plant–pollinator interactions. In: *Advances in Ecological Research*. Academic Press, 147–199.
- Gordon R (1985) The Coccinellidae (Coleoptera) of America north of Mexico. *Journal of the New York Entomological Society* 93: 1–912.
- Goulson D, Lye G, Darvill B (2008) Decline and conservation of bumble bees. *Annual Review of Entomology* 53: 191–208.
- Goulson D, Nicholls E, Botías C, Rotheray EL (2015) Bee declines driven by combined stress from parasites, pesticides, and lack of flowers. *SciencExpress*: 1–16.
- Graf W, Lorenz A, de Figueroa J, Lücke S, López-Rodríguez M, Davies C (2009) *Distribution and ecological preferences of European freshwater organisms. Volume 2. Plecoptera*. Schmidt-Kloiber A, Hering D (Eds). Pensoft Publishing.
- Graf W, Murphy J, Dahl J, Zamora-Munoz C, Lopez-Rodriguez M (2008) *Distribution and ecological preferences of European freshwater organisms. Volume 1. Trichoptera*. Schmidt-Kloiber A, Hering D (Eds). Pensoft Publishing, 262 pp.
- Grand D, Boudot J-P (2006) *Les libellules de France, Belgique, et Luxembourg*. Biotope, collection Parthénope, Mèze, 480 pp.
- Grand D, Boudot J-P, Doucet G (2014) *Cahier d’identification des Libellules de France, Belgique, Luxembourg et Suisse*. Biotope Ed. France, 136 pp.
- Grand D, Marinov M, Cook C, Jourdan H, Rouys S, Theuerkauf J (2018) Identification key to adult Odonata of New Caledonia and Wallis and Futuna. *Odonatologica* 47: 179–180.
- Grattepanche JD, Walker LM, Ott BM, Paim Pinto DL, Delwiche CF, Lane CE, Katz LA (2018) Microbial diversity in the eukaryotic SAR clade: Illuminating the darkness between morphology and molecular data. *BioEssays* 40: 1–12.

- Griffing LR (2011) Who invented the dichotomous key? Richard Waller's watercolors of the herbs of Britain. *American Journal of Botany* 98: 1911–1923.
- Guiney MS, Oberhauser KS (2009) Conservation volunteers' connection to nature. *Ecopsychology* 1: 187–197.
- Hagedorn G (2007) *Structuring descriptive data of organisms — requirement analysis and information models*. PhD thesis, Universität Bayreuth.
- Hagedorn G, Rambold G, Martellos S (2010) Types of identification keys. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 59–64.
- Hagedorn G, Thiele K, Morris R, Heidorn P (2005) The Structured Descriptive Data (SDD) w3c-xml-schema, version 1.0. *Biodiversity Information Standards (TDWG)*.
- Hájek J (2009) Dytiscidae-Icones Insectorum Europae Centralis. *Folia Heyrovskyana (B)* 11: 1–32.
- Hallmann CA, Sorg M, Jongejans E, Siepel H, Hofland N, Schwan H, Stenmans W, Müller A, Sumser H, Hörrén T, Goulson D, De Kroon H (2017) More than 75 percent decline over 27 years in total flying insect biomass in protected areas. *PLoS ONE* 12: e0185809.
- Harnik PG, Ross RM (2003) Assessing data accuracy when involving students in authentic paleontological research. *Journal of Geoscience Education* 51: 76–84.
- Hebert PDN, Cywinska A, Ball SL, DeWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences* 270: 313–321.
- Hebert PDN, Penton EH, Burns JM, Janzen DH, Hallwachs W (2004) Ten species in one: DNA barcoding reveals cryptic species in the neotropical skipper butterfly *Astrartes fulgerator*. *Proceedings of the National Academy of Sciences of the United States of America* 101: 14812–14817.
- Hemptinne J, Magro A, Majerus M (2005) *Les coccinelles*. Delachaux et Niestlé, Paris, 190 pp.
- Hillis DM (1987) Molecular versus morphological approaches to systematics. *Annual Review of Ecology and Systematics* 18: 23–42.
- Hodek I, Honěk A (2009) Scale insects, mealybugs, whiteflies and psyllids (Hemiptera, Sternorrhyncha) as prey of ladybirds. *Biological Control* 51: 232–243.
- Holland J, Hutchison M, Smith B, Aebischer N (2006) A review of invertebrates and seed-bearing plants as food for farmland birds in Europe. *Annals of Applied Biology* 148: 49–71.
- Horn G Van, Aodha O Mac, Song Y, Cui Y, Sun C, Shepard A, Adam H, Perona P, Belongie S (2018) The iNaturalist species classification and detection dataset. In: *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 8769–8778.
- Houard X, Merlet F (2014) *Liste rouge régionale des libellules d'Ile-de-France*. Natureparif, Office pour les insectes et leur environnement, Société Française d'Odonatologie, Paris, 80 pp.
- Houllier F, Merilhou-Goudard J (2016) *Les sciences participatives en France. États des lieux, bonnes pratiques et recommandations*. 63 pp.
- Howarth F, Ramsay G (1991) The conservation of island insects and their habitats. In: Collins N, Thomas J (Eds), *The Conservation of Insects and Their Habitats*. Academic Press, New York, 71–108.
- Hynes HBN (1977) *A key to the adults and nymphs of the British stoneflies (Plecoptera)*. Scientific Association FB (Ed). 90 pp.

- Iablokoff-Khinzorian S (1982) *Les coccinelles, coléoptères Coccinellidae. Tribu Coccinellini des régions Paléarctique et Orientale*. Boubée, Paris, 568 pp.
- Ide FP (1935) Post embryological development of Ephemeroptera (mayflies). External characters only. *Canadian Journal of Research*: 433–478.
- Iorgu IŞ, Iorgu EI, Szövényi G, Orci KM (2017) A new, morphologically cryptic bush-cricket discovered on the basis of its song in the Carpathian Mountains (Insecta, Orthoptera, Tettigoniidae). *ZooKeys* 680: 57–72.
- Irwin A (1995) *Citizen science: A study of people, expertise and sustainable development*. London, 216 pp.
- Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science* 33: 312–324.
- Jain M, Fiddes IT, Miga KH, Olsen HE, Paten B, Akeson M (2015) Improved data analysis for the MinION nanopore sequencer. *Nature Methods* 12: 351–356.
- Jetz W, McPherson JM, Guralnick RP (2012) Integrating biodiversity distribution knowledge: Toward a global map of life. *Trends in Ecology and Evolution* 27: 151–159.
- Johnson MF, Hannah C, Acton L, Popovici R, Karanth KK, Weinthal E (2014) Network environmentalism: Citizen scientists as agents for environmental advocacy. *Global Environmental Change* 29: 235–245.
- Joly A, Bonnet P, Goëau H, Barbe J, Selmi S, Champ J, Dufour-Kowalski S, Affouard A, Carré J, Molino J-F, Boujemaa N, Barthélémy D (2016) A look inside the Pl@ntNet experience. *Multimedia Systems* 22: 751–766.
- Jonker R, Groben R, Tarran G, Medlin L, Wilkins M, García L, Zabala L, Boddy L (2000) Automated identification and characterisation of microbial populations using flow cytometry: the AIMS project. *Scientia Marina* 64: 225–234.
- Jouveau S, Delaunay M, Vignes-Lebbe R, Nattier R (2018) A multi-access identification key based on colour patterns in ladybirds (Coleoptera, coccinellidae). *ZooKeys*: 55–73.
- Julliard R (2017) Science participative et suivi de la biodiversité : l'expérience Vigie-Nature. *Natures Sciences Sociétés* 25: 412–417.
- Kerner A, Debrenne F, Vignes-Lebbe R (2011) Cambrian archaeocyathan metazoans: Revision of morphological characters and standardization of genus descriptions to establish an online identification tool. *ZooKeys* 150: 381–395.
- Kim K (1993a) Biodiversity, conservation and inventory: why insects matter. *Biodiversity and Conservation* 2: 191–214.
- Kim K (1993b) Insect pests and evolution. In: Kim K, McPherson B (Eds), *Evolution of insect pests*. John Wiley and Sons, Inc., New York, 3–26.
- Klein A-M, Vaissière BE, Cane JH, Steffan-Dewenter I, Cunningham SA, Kremen C, Tscharntke T (2007) Importance of pollinators in changing landscapes for world crops. *Proceedings of the Royal Society B: Biological Sciences* 274: 303–313.
- Klimmek F, Baur H (2018) An interactive key to Central European species of the *Pteromalus albipennis* species group and other species of the genus (Hymenoptera: Chalcidoidea: Pteromalidae), with the description of a new species. *Biodiversity Data Journal* 6: e27722.

- Kluser S, Peduzzi P (2007) *Global pollinator decline: A literature review*. United Nations Environment Programme.
- Ko TWK, Stephenson SL, Bahkali AH, Hyde K (2011) From morphology to molecular biology: can we use sequence data to identify fungal endophytes? *Fungal Diversity* 50: 113.
- Kollmuss A, Agyeman J (2002) Mind the gap: why do people act environmentally and what are the barriers to pro-environmental behavior? *Environmental Education Research* 8: 239–260.
- Kullenberg C, Kasperowski D (2016) What is citizen science? – A scientometric meta-analysis. *PLoS ONE* 11: e0147152.
- Kumar N, Belhumeur P, Biswas A, Jacobs D, Kress W, Lopez I, Soares J (2012) Leafsnap: A computer vision system for automatic plant species identification. In: *European Conference on Computer Vision*, 502–516.
- Lamarck J-B (1778) *Flore Française, ou Description succincte de toutes les plantes qui croissent naturellement en France*. Paris, l’Imprimerie Royale.
- Larios N, Deng H, Zhang W, Sarpola M, Yuen J, Paasch R, Moldenke A, Lytle DA, Correa SR, Mortensen EN, Shapiro LG, Dietterich TG (2008) Automated insect identification through concatenated histograms of local appearance features: feature vector generation and region detection for deformable objects. *Machine Vision and Applications* 19: 105–123.
- Larsen BB, Miller EC, Rhodes MK, Wiens JJ (2017) Inordinate fondness multiplied and redistributed: the number of species on Earth and the new pie of life. *The Quarterly Review of Biology* 92: 229–265.
- Laver T, Harrison J, O’Neill PA, Moore K, Farbos A, Paszkiewicz K, Studholme DJ (2015) Assessing the performance of the Oxford Nanopore Technologies MinION. *Biomolecular Detection and Quantification* 3: 1–8.
- Law J, Lynch M (1988) Lists, field guides, and the descriptive organization of seeing: Birdwatching as an exemplary observational activity. *Human Studies* 11: 271–303.
- Leal MC, Hilário A, Munro MHG, Blunt JW, Calado R (2016) Natural products discovery needs improved taxonomic and geographic information. *Natural Product Report* 33: 747–750.
- Lebbe J (1984) *Manuel d’utilisation du logiciel XPER*. Micro Application, Rueil Malmaison.
- Lebbe J (1991) *Représentation des concepts en biologie et en médecine: introduction à l’analyse des connaissances et à l’identification assistée par ordinateur*. Université Pierre et Marie Curie, Paris, France
- Lecointre G, Le Guyader H (2017) *Classification phylogénétique du vivant - Tome 2. 4^e éd.* Belin (Ed). 832 pp.
- Legrand M (2013) Vigie-Nature : sciences participatives et biodiversité à grande échelle. *Cahiers des Amériques latines* 72–73: 65–84.
- Levé M, Baudry E, Bessa-Gomes C (2019) Domestic gardens as favorable pollinator habitats in impervious landscapes. *Science of the Total Environment* 647: 420–430.
- Lewandowski E, Oberhauser K (2017) Butterfly citizen scientists in the United States increase their engagement in conservation. *Biological Conservation* 208: 106–112.
- Linnæus C (1753) *Species Plantarum, Tomus 1*. Salvius, Stockholm.

- Lister AM, Brooks SJ, Fenberg PB, Glover AG, James KE, Johnson KG, Michel E, Okamura B, Spencer M, Stewart JR, Todd JA, Valsami-Jones E, Young J (2011) Natural history collections as sources of long-term datasets. *Trends in Ecology and Evolution* 26: 153–154.
- Losey J, Vaughan M (2006) The economic value of ecological services provided by insects. *AIBS Bulletin* 56: 311–323.
- Macadam C, Bennett C (2010) *A pictorial guide to British Ephemeroptera*. AIDGAP. Field Studies Council Publications, 128 pp.
- Majerus M, Strawson V, Roy H (2006) The potential impacts of the arrival of the harlequin ladybird, *Harmonia axyridis* (Pallas) (Coleoptera: Coccinellidae), in Britain. *Ecological Entomology* 31: 207–215.
- Malausa T, Delaunay M, Fleisch A, Groussier-Bout G, Warot S, Crochard D, Guerrieri E, Delvare G, Pellizzari G, Kaydan MB, Al-Khateeb N, Germain J-F, Brancaccio L, Le Goff I, Bessac M, Ris N, Kreiter P (2016) Investigating biological control agents for controlling invasive populations of the mealybug *Pseudococcus comstocki* in France. *PLoS ONE* 11: e0157965.
- Maloney MP, Ward MP (1973) Let's hear from the people: An objective scale for the measurement of ecological attitudes and knowledge. *American Psychologist* 28: 583.
- Manach J (1984) *Traité pratique d'identification des éphémères*. S.A. PEL (Ed). Paris, 160 pp.
- Marshall S (2000) Comments on error rates in insect identifications. *Newsletter of the Biological Survey of Canada (Terrestrial Arthropods)* 19.
- Martin P, Schmelz RM, Dole-Olivier MJ (2015) Groundwater oligochaetes (Annelida, Clitellata) from the Mercantour National Park (France), with the descriptions of one new genus and two new stygobiont species. *Zoosystema* 37: 551–569.
- Mayo M, Watson A (2006) Automatic species identification of live moths. In: *International Conference on Innovative Techniques and Applications of Artificial Intelligence*. Springer, London, 58–71.
- Mayr E (1969) *Principles of systematic zoology*. 428 pp.
- Meusnier I, Singer G, Landry J-F, Hickey D, Hebert P, Hajibabaei M (2008) A universal DNA mini-barcode for biodiversity analysis. *BMC Genomics* 9: 4.
- Meyer J, Sale M, Mulholland P, Poff N (1999) Impacts of climate change on aquatic ecosystem functioning and health. *Journal of the American Water Resources Association* 35: 1373–1386.
- Mondry CP, Villeneuve B, Archaimbault V, Usseglio-Polatera P (2012) A new macroinvertebrate-based multimetric index (I2M2) to evaluate ecological quality of French wadeable streams fulfilling the WFD demands: A taxonomical and trait approach. *Ecological Indicators* 18: 452–467.
- Le Monnier Y, Livory A (2003) *Atlas des coccinelles de la Manche : une enquête Manche-Nature 1998-2001*. Manche-Nature, Coutances, 208 pp.
- Mora C, Tittensor DP, Adl S, Simpson AGB, Worm B (2011) How many species are there on earth and in the ocean? *PLoS Biology* 9: 1–8.
- Nilsson A (1996a) *Aquatic Insects of North Europe: A Taxonomic Handbook. Volume 1: Ephemeroptera-Plecoptera-Heteroptera-Neuroptera-Megaloptera-Coleoptera-Trichoptera-Lepidoptera*. Apollo Books, 274 pp.
- Nilsson A (1996b) *Aquatic Insects of North Europe: A Taxonomic Handbook. Volume 2: Odonata - Diptera*. Apollo Books, 440 pp.

- Ong W (1958) *Ramus, method, and the decay of dialogue: From the art of discourse to the art of reason*. Harvard University Press, Cambridge, Massachusetts, USA.
- Osmundson TW, Robert VA, Schoch CL, Baker LJ, Smith A, Robich G, Mizzan L, Garbelotto MM (2013) Filling gaps in biodiversity knowledge for macrofungi: Contributions and assessment of an Herbarium collection DNA barcode sequencing project. *PLoS ONE* 8: e62419.
- Padial JM, Miralles A, De la Riva I, Vences M (2010) The integrative future of taxonomy. *Frontiers in Zoology* 7: 1–14.
- Page TJ, Choy SC, Hughes JM (2005) The taxonomic feedback loop: Symbiosis of morphology and molecules. *Biology Letters* 1: 139–142.
- Pankhurst R, Aitchinson R (1975) An on-line identification program. In: Pankhurst R (Ed), *Biological identification with computers*. Academic Press, 181–194.
- Postel S, Carpenter S (1997) Freshwater ecosystem services. In: *Nature's services: Societal dependence on natural ecosystems*. Island Press, 412 pp.
- Potts S, Biesmeijer J, Kremen C, Neumann P, Schweiger O, Kunin W (2010) Global pollinator declines: trends, impacts and drivers. *Trends in Ecology & Evolution* 25: 345–353.
- Power A (2010) Ecosystem services and agriculture: tradeoffs and synergies. *Philosophical Transactions of the Royal Society B: Biological Sciences* 365: 2959–2971.
- Pryce D, Macadam C, Brooks S (2007) *Guide to the British stonefly (Plecoptera) families: adults and larvae*. AIDGAP. Field Studies Council Publications.
- Pyle R, Bentzien M, Opler P (1981) Insect conservation. *Annual Review of Entomology* 26: 233–258.
- Rader R, Bartomeus I, Garibaldi LA, Garratt MPD, Howlett BG, Winfree R, Cunningham SA, Mayfield MM, Arthur AD, Andersson GKS, Bommarco R, Brittain C, Carvalheiro LG, Chacoff NP, Entling MH, Foully B, Freitas BM, Gemmill-Herren B, Ghazoul J, Griffin SR, Gross CL, Herbertsson L, Herzog F, Hipólito J, Jaggard S, Jauker F, Klein A-M, Kleijn D, Krishnan S, Lemos CQ, Lindström SAM, Mandelik Y, Monteiro VM, Nelson W, Nilsson L, Pattemore DE, de O. Pereira N, Pisanty G, Potts SG, Reemer M, Rundlöf M, Sheffield CS, Scheper J, Schüepp C, Smith HG, Stanley DA, Stout JC, Szentgyörgyi H, Taki H, Vergara CH, Viana BF, Woyciechowski M (2016) Non-bee insects are important contributors to global crop pollination. *Proceedings of the National Academy of Sciences* 113: 146–151.
- Ramus P (1574) *The logicke of Peter Ramus (Translated by Roland MacIlmaine)*. Dunn CM (Ed). Published in 1969 by San Fernando Valley State College, Northridge, California, USA.
- Ramus P (1592) *The art of arithmeticke in whole numbers and fractions: in a more readie and easie method then hitherto hath bene published (Translated into English by William Kempe)*. London.
- Randler C (2006) Empirical evaluation of a dichotomous key for amphibian identification in pupils and students. *Journal of Science Education* 7: 34.
- Randler C (2008) Teaching species identification - A prerequisite for learning biodiversity and understanding ecology. *Eurasia Journal of Mathematics, Science and Technology Education* 4: 223–231.
- Randler C, Knappe B (2007) Comparison of a dichotomous, language-based with an illustrated identification key for animal tracks and signs. *Journal of Science Education* 8: 32–35.

- Randler C, Zehender I (2006) Effectiveness of reptile species identification – A comparison of a dichotomous key with an identification book. *Eurasia Journal of Mathematics, Science and Technology Education* 2: 55–65.
- Ray J (1686) *Historia generalis plantarum*. Londra, 1188 pp.
- Ray J (1690) *Synopsis methodica stirpium Britannicarum*. Impensis Gulielmi & Joannis Innys.
- Razak NFA, Supramaniam C V., Zieritz A (2018) A dichotomous PCR–RFLP identification key for the freshwater mussels (Bivalvia: Unionida) of Peninsular Malaysia. *Conservation Genetics Resources*, 1–8
- Reinhold L, Law B, Ford G, Pennay M (2001) Key to the bat calls of south-east Queensland and north-east New South Wales. *Forest Ecosystem Research and Assessment Technical Paper* 7.
- Riccamboni R, Mereu A, Boscarol C (2010) Keys to Nature: a test on the iPhone market. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 445–450.
- Richoux P (1982) Introduction pratique à la systématique des organismes des eaux continentales françaises. 2. Coléoptères aquatiques genres: adultes et larves (première partie). *Bulletin mensuel de la Société linnéenne de Lyon* 4: 105–128.
- Roubik D (1995) Pollination of cultivated plants in the tropics. No. 118. *Food & Agriculture Org.*
- Roy H, Adriaens T, Isaac N, Kenis M, Onkelinx T, San Martin G, Brown P, Hautier L, Poland R, Roy D, Comont R, Eschen R, Frost R, Zindel R, Van Vlaenderen J, Nedvěd O, Ravn H, Grégoire J-C, de Biseau J-C, Maes D (2012) Invasive alien predator causes rapid declines of native European ladybirds. *Diversity and Distributions* 18: 717–725.
- Roy H, Brown P, Comont R, Poland R, Sloggett J, Majerus M, Kearns P (2013) *Naturalists' Handbook 10: Ladybirds*. Pelagic Publishing, Exeter, 142 pp.
- Rugman-Jones PF, Hoddle MS, Mound LA, Stouthamer R (2006) Molecular identification key for pest species of *Scirtothrips* (Thysanoptera: Thripidae). *Journal of Economic Entomology* 99: 1813–1819.
- Ruse M (1969) Definitions of species in biology. *The British Society for the Philosophy of Science* 20: 97–119.
- dos Santos J, Silveira O, Carpenter J (2018) Taxonomic revision of the *Protopolybia sedula* species-group (Hymenoptera, Vespidae, Polistinae), with an new identification key to species. *Zootaxa* 4403: 87.
- SCBD (2010) *What is the Problem? The Taxonomic Impediment*.
- Schemske D, Horvitz C (1984) Variation among floral visitors in pollination ability: a precondition for mutualism specialization. *Science* 225: 519–521.
- Schmidt-Kloiber A, Hering D (2015) www.freshwaterecology.info - An online tool that unifies, standardises and codifies more than 20,000 European freshwater organisms and their ecological preferences. *Ecological Indicators* 53: 271–282.
- Seago A, Giorgi J, Li J, Ślipiński A (2011) Phylogeny, classification and evolution of ladybird beetles (Coleoptera: Coccinellidae) based on simultaneous analysis of molecular and morphological data. *Molecular Phylogenetics and Evolution* 60: 137–151.
- Shaffer HB, Fisher RN, Davidson C (1998) The role of natural history collections in documenting species declines. *Trends in Ecology & Evolution* 13: 27–30.
- Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nature Biotechnology* 26: 1135–1145.

- Shwartz A, Muratet A, Simon L, Julliard R (2013) Local and management variables outweigh landscape effects in enhancing the diversity of different taxa in a big metropolis. *Biological Conservation* 157: 285–292.
- Simpson GG (1961) *Principles of animal taxonomy*.
- Snodgrass R (1954) The dragonfly larva. *Smithsonian Miscellaneous Collections* 123: 1–44.
- SNPN (2012) *Bilan de l'inventaire des mares 2011*.
- van Spronsen E, Martellos S, Seijts D, Schalk P, Nimis PL (2010) Modifiable digital identification keys. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 127–131.
- Stefano M, Nimis PL (2008) *KeyToNature*: teaching and learning biodiversity. Dryades, the Italian experience. In *Proceedings of the IASK International Conference Teaching and Learning*, 863–868
- Suarez AV, Tsutsui ND (2004) The value of museum collections for research and society. *BioScience* 54: 393–402.
- Szövényi G, Puskás G, Orci KM (2012) *Isophya nagyii*, a new phaneropterid bush-cricket (Orthoptera: Tettigonioidea) from the Eastern Carpathians (Caliman Mountains, North Romania). *Zootaxa* 79: 67–79.
- Tachet H, Richoux P, Bournaud M, Usseglio-Polatera P (2010) *Invertébrés d'eau douce: systématique, biologie, écologie - Nouvelle édition revue et augmentée*. CNRS Editions, Paris, 608 pp.
- Tan H, Fischer A, Tinchin P, Stieger M, Steenbekkers L, van Trijp H (2015) Insects as food: exploring cultural exposure and individual experience as determinants of acceptance. *Food Quality and Preference* 42: 78–89.
- Thapa R (2006) Honeybees and other insect pollinators of cultivated plants: A review. *Journal of the Institute of Agriculture and Animal Science* 27: 1–23.
- Thompson SCG, Barton MA (1994) Ecocentric and anthropocentric attitudes toward the environment. *Journal of environmental Psychology* 14: 149–157.
- Tilman D (2000) Causes, consequences and ethics of biodiversity. *Nature* 405: 208.
- Trilar T (2010) A Wiki-based key to garden and village birds. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 95–98.
- Trombulak S, Frissell C (2000) Review of ecological effects of roads on terrestrial and aquatic communities. *Conservation Biology* 14: 18–30.
- Tronquet M (2014) *Catalogue des Coléoptères de France*. Association Roussillonnaise d'Entomologie, 1052 pp.
- Troudet J, Grandcolas P, Blin A, Vignes-Lebbe R, Legendre F (2017) Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports* 7: 9132.
- Tyler EHM, Somerfield PJ, Berghe E Vanden, Bremner J, Jackson E, Langmead O, Palomares MLD, Webb TJ (2012) Extensive gaps and biases in our knowledge of a well-known fauna: Implications for integrating biological traits into macroecology. *Global Ecology and Biogeography* 21: 922–934.
- Ung V, Dubus G, Zaragueta-Bagils R, Vignes-Lebbe R (2010) Xper2: introducing e-taxonomy. *Bioinformatics* 26: 703–704.

- Urbano-Bonilla A, Ballen GA, Herrera-R GA, Zamudio J, Herrera-Collazos EE, DoNascimento C, Prada-Pedrerros S, Maldonado-Ocampo JA (2018) Fishes of the Cusiana River (Meta River basin, Colombia), with an identification key to its species. *ZooKeys* 733: 65–97.
- Varese G, Anastasi A, Voyron S, Marchisio V (2010) An interactive tool for the identification of airborne and food fungi. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 183–187.
- Vieira BN, Poff NL, Carlisle DM, Moulton SR, Koski ML, Kondratieff BC (2006) A Database of lotic invertebrate traits for North America. *US Geological Survey Data Series* 187: 1–15.
- Vignes-Lebbe R, Chesselet P, Diep Thi M-H (2016) Xper³: new tools for collaborating, training and transmitting knowledge on botanical phenotypes. In: Rakotoarisoa N, Blackmore S, Riera B (Eds), *Botanists of the twenty-first century: roles, challenges and opportunities*. Paris, 228–239.
- Wäldchen J, Mäder P (2018) Machine learning for image based species identification. *Methods in Ecology and Evolution* 9: 2216–2225.
- Waller R (1689) *Tables of the English herbs*. Royal Society Archives. Journal Book Original 8: 252.
- Warren PS (2003) Winter dialects in the bronzed cowbird and their relationship to breeding-season dialects. *Animal Behaviour* 65: 1169–1178.
- Watson RT (2005) Turning science into policy: challenges and experiences from the science–policy interface. *Philosophical Transactions of the Royal Society B: Biological Sciences* 360: 471–477.
- Weber G, Hagedorn G (2010) Wiki keys on mobile devices. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*, 89–93.
- Will KW, Rubinoff D (2004) Myth of the molecule: DNA barcodes for species cannot replace morphology for identification and classification. *Cladistics* 20: 47–55.
- Williams HC, Ormerod SJ, Bruford MW (2006) Molecular systematics and phylogeography of the cryptic species complex *Baetis rhodani* (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution* 40: 370–382.
- Williams I (1994) The dependence of crop production within the European Union on pollination by honey bees. *Agricultural Zoology Reviews* (United Kingdom).
- Worm B, Barbier E, Beaumont N, Duffy J, Folke C, Halpern B, Jackson J, Lotze H, Micheli F, Palumbi S, Sala E, Selkoe K, Stachowicz J, Watson R (2006) Impacts of biodiversity loss on ocean ecosystem services. *Science* 314: 787–790.
- Wright DR, Underhill LG, Keene M, Knight AT (2015) Understanding the motivations and satisfactions of volunteers to improve the effectiveness of citizen science programs. *Society & Natural Resources* 28: 1013–1029.
- Yamin F (1995) Biodiversity, ethics and international law. *International Affairs* 71: 529–546.
- Yang H-P, Ma C-S, Wen H, Zhan Q-B, Wang X-L (2015) A tool for developing an automatic insect identification system based on wing outlines. *Scientific Reports* 5: 12786.
- Vander Zanden MJ, Hansen GJA, Higgins SN, Kornis MS (2010) A pound of prevention, plus a pound of cure: Early detection and eradication of invasive species in the Laurentian Great Lakes. *Journal of Great Lakes Research* 36: 199–205.
- Zervakis GI, Ntougias S, Gargano ML, Besi MI, Polemis E, Typas MA, Venturella G (2014) A reappraisal of the *Pleurotus eryngii* complex – New species and taxonomic combinations based on the

application of a polyphasic approach, and an identification key to *Pleurotus* taxa associated with Apiaceae plants. *Fungal Biology* 118: 814–834.

Zielińska E, Karaś M, Jakubczyk A, Zieliński D, Baraniak B (2019) Edible insects as source of proteins. *Bioactive Molecules in Food*, 1–53.

ANNEXES

Annexe 1 : Article publié dans *ZooKeys*

A multi-access identification key based on colour patterns in ladybirds (Coleoptera, Coccinellidae)

Séverin Jouveau^{1,2,*}, Mathilde Delaunay^{1,*}, Régine Vignes-Lebbe¹, Romain Nattier¹

1 Institut Systématique Evolution Biodiversité (ISYEB), Muséum national d'Histoire naturelle, CNRS, Sorbonne Université, EPHE, 57 rue Cuvier, CP 50, 75005 Paris, France **2** INRA - UMR 1202 Biodiversité Gènes & Communautés, 69 route d'Arcachon, 33612 Cestas, France

Corresponding author: Romain Nattier (nattier@mnhn.fr)

Academic editor: M. Thomas | Received 9 November 2017 | Accepted 2 April 2018 | Published 14 May 2018

<http://zoobank.org/9B684185-787B-4F79-81A4-32354633786F>

Citation: Jouveau S, Delaunay M, Vignes-Lebbe R, Nattier R (2018) A multi-access identification key based on colour patterns in ladybirds (Coleoptera, Coccinellidae). ZooKeys 758: 55–73. <https://doi.org/10.3897/zookeys.758.22171>

Abstract

An identification key based on French ladybird colouration is proposed for the tribes Chilocorini, Coccinellini, and Epilachnini. These tribes were chosen based on their relatively limited species diversity, as well as their large size and high colour diversity, making them easy to observe and collect. The identification key runs on Xper³ software, which allows the building of structured knowledge bases and online free-access keys. The online interactive Xper key is available at <http://french-ladybird.identificationkey.fr>.

Keywords

Citizen science, Coccinellidae, Coccinellini, Chilocorini, Epilachnini, France, interactive identification keys, ladybirds, Xper

Introduction

The identification of species is central in ecology, conservation biology, systematics, and related disciplines (species inventories and community studies, ecosystem management, establishment and improvement of environmental public policies, taxonomic

* These authors have equally contributed

reviews, and management of natural history collections) (Oliver 1988, Hebert et al. 2003, Smith et al. 2008, Vander Zanden et al. 2010). Europe is one of the best-known parts in the world in terms of biodiversity (Fontaine et al. 2012), especially concerning distribution patterns at the country scale. This has been highlighted by the release of the Fauna Europaea database since 2004 (de Jong et al. 2014), which gathers the scientific names and distributions of all living European animal species and is assembled by a large network of specialists. However, most new species are described by non-professional taxonomists (Fontaine et al. 2012) and the distribution of a majority of organisms remains poorly known. Citizen science programs aim to fill that gap (Silvertown 2009), thanks to the participation of amateurs and the general public to the inventory and description of life (e.g., National Biodiversity Network in the UK, Swedish Species Gateway in Sweden, Chicago Wilderness Project in the USA, Vigie-Nature in France; see Silvertown 2009). From this perspective, visual and interactive identification of species offers tremendous potential for the general public.

If the identification of large and charismatic animals may be easy, the majority of organisms require expert skills for accurate identification and the inability to identify species represents a major challenge known as the Taxonomic Impediment (SCBD 2010). The most basic requirement for people studying and working on biodiversity aspects is the availability of species identification guides. However, easy-to-use identification guides for non-taxonomists and the general public are scarce and available for relatively few taxonomic groups (SCBD 2010). Consequently, the other features of organisms (such as distribution, ecology, biology) remain poorly known (Costello et al. 2006, SCBD 2010).

Coccinellidae is a family of beetles popular and appreciated by naturalists and the general public. Because these animals have ecological and economic values as predators of pest insects (e.g. aphids, scale insects), their identification may be of importance for naturalists, amateurs and professionals (Hemptinne et al. 2005, Hodek and Honěk 2009, Ali et al. 2014). Several citizen science programs aim to describe the distribution patterns of this group: we can mention for instance the Harlequin Ladybird Survey (<http://www.harlequin-survey.org>) and the Ladybird Survey (<http://www.ladybird-survey.org>) in the UK, the Lost Ladybug Project (<http://www.lostladybug.org>) and the Buckeye Lady Beetle Blitz (<https://entomology.osu.edu/about-us/multi-media/buckeye-lady-beetle-blitz>) in the US, and the Coccinula Recording Scheme in Belgium (Bagnée et al. 2011). The data collected led to a significant number of scientific works that have been published (e.g., Brown et al. 2008, Comont et al. 2012, 2014, Gardiner et al. 2012, Purse et al. 2015).

Single-access identification keys consist of a series of identification steps that form a single and unique identification path for a given taxon. Although it is a very powerful tool for identifying species, the user cannot choose the character to be observed (the answer for every single step must be known), and the identification is impossible if some characters are missing (e.g., if the specimen is poorly preserved). Moreover, this type of keys cannot be modulated or adapted to various kinds of publics, environmental conditions, season, or geographical location.

Most North American or European ladybird identification keys are single-access and difficult to use for non-specialists (Dauguet 1949, Iablokoff-Khuzorian 1982, Gordon 1985, Chapin and Brou 1991). Others are mainly based on shape and colour, but most characters need specific vocabulary, which makes the key still too complicated for the general public in the perspective of citizen science programs (Belgium: Baugnée and Branquart 2000; West of France: Le Monnier and Livory 2003; British Isles: Roy et al. 2013; North of France: Declercq et al. 2014).

Modern tools developed along with digital technologies and data processing make identification easier for the user. In this perspective, several interactive identification keys (IIK) are available online (e.g., <http://www.ladybird-survey.org/bbc/spotter.php>, <http://www.discoverlife.org/20/q?guide=Ladybug>), but most of them are only digital versions of single-access keys and maintain the same difficulties for the user.

A multi-access interactive key is a computer-aided identification tool that makes it possible to find correct names of species where the user enters attributes (character-state values) of the specimen (Dallwitz et al. 2013). The advantages compared to conventional keys are as follows: characters can be used in any order, characters are ordered to start with the one that best separates the remaining taxa, keys can be completed with illustrations (pictures, drawings) and texts explaining the terminology used, correct identifications can be obtained despite errors made by the user (FloraBase – <https://florabase.dpaw.wa.gov.au/keys>; Dallwitz et al. 2013). The software also includes the possibility to print a single-access key for field identification if needed, and to weight characters according to the user skills and abilities (students, general public, naturalists...). Despite the advantages provided by multi-access interactive keys, none has been produced for ladybirds so far.

This study aims to i) release the first multi-access digital interactive identification key for French ladybirds based on colour that takes into account intraspecific variability; and ii) study and discuss the discriminating power of the characters: can we identify species by colour pattern only? What are the most discriminating characters?

Materials and methods

Taxonomic coverage

As the aim of the key is to provide an identification tool for the general public in the perspective of citizen science programs, we have restricted the taxonomic coverage to the tribes Chilocorini, Coccinellini and Epilachnini (Table 1). Members of these tribes are relatively large (3–9 mm) and display a great diversity of colours, making them easily detectable in their environment and identifiable by non-specialists. We also included the most common colour forms, trying to cover most of the intraspecific variability of these species.

The current taxonomy (Seago et al. 2011) and the species list follow Tronquet (2014) and include native, introduced, and acclimated species. Sixty-six taxa are in-

Table I. Taxonomic coverage of the study.

	Chilocorini Mulsant, 1846	<i>Chilocorus bipustulatus</i> Linnaeus, 1758	
		<i>Chilocorus renipustulatus</i> Scriba, 1791	
Coccinellinae Latreille, 1807		<i>Exochomus octosignatus</i> Gebler, 1830	
		<i>Exochomus quadripustulatus</i> Linnaeus, 1758	
		<i>Parexochomus nigromaculatus</i> Goeze, 1777	
		<i>Halyzia sedecimguttata</i> Linnaeus, 1758	
		<i>Psyllobora vigintiduopunctata</i> Linnaeus, 1758	
		<i>Vibidia duodecimguttata</i> Poda, 1761	
		<i>Anisosticta novemdecimpunctata</i> Linnaeus, 1758	
		<i>Coccinula quatuordecimpustulata</i> Linnaeus, 1758	
		<i>Coccinula sinuatoinarginata</i> Faldermann, 1837	
		<i>Tytthaspis sedecimpunctata</i> Linnaeus 1761 – f. <i>duodecimpunctata</i>	
		<i>Adalia bipunctata</i> Linnaeus, 1758 – f. <i>annulata</i>	
		<i>Adalia bipunctata</i> Linnaeus, 1758 – f. <i>pantherina</i>	
		<i>Adalia bipunctata</i> Linnaeus, 1758 – f. <i>quadrimaculata</i>	
		<i>Adalia bipunctata</i> Linnaeus, 1758 – f. <i>sempustulata</i>	
		<i>Adalia bipunctata</i> Linnaeus, 1758 – f. <i>typica</i>	
		<i>Adalia conglomerata</i> Linnaeus, 1758 – f. <i>decas</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>decempustulata</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>guttatopunctata</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>lutea</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>quadripunctata</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>scribai</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>subpunctata</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>terna</i>	
		<i>Adalia decempunctata</i> Linnaeus, 1758 – f. <i>typica</i>	
		<i>Anatis ocellata</i> Linnaeus, 1758	
		<i>Aphidecta oblitterata</i> Linnaeus, 1758 – f. <i>typica</i>	
	Coccinellini Latreille, 1807		<i>Calvia decemguttata</i> Linnaeus, 1767
			<i>Calvia quatuordecimguttata</i> Linnaeus, 1758
			<i>Calvia quindecimguttata</i> Fabricius, 1777
			<i>Ceratomegilla alpina</i> Villa A. & Villa G. B., 1835
			<i>Ceratomegilla notata</i> Laicharting, 1781
			<i>Ceratomegilla rufocincta</i> Mulsant, 1850
			<i>Ceratomegilla undecimnotata</i> Schneider D.H. 1792
			<i>Coccinella venusta</i> Weise, 1879
			<i>Coccinella hieroglyphica</i> Linnaeus, 1758
			<i>Coccinella magnifica</i> Redtenbacher, 1843
			<i>Coccinella quinquepunctata</i> Linnaeus, 1758
			<i>Coccinella septempunctata</i> Linnaeus, 1758
			<i>Coccinella undecimpunctata</i> Linnaeus, 1758
			<i>Harmonia axyridis</i> Pallas, 1773 – f. <i>conspicua</i>
			<i>Harmonia axyridis</i> Pallas, 1773 – f. <i>novemdecimsignata</i>
			<i>Harmonia axyridis</i> Pallas, 1773 – f. <i>spectabilis</i>
			<i>Harmonia conformis</i> Boisduval, 1835
			<i>Harmonia quadripunctata</i> Pontoppidan, 1763 – f. <i>sedecimpunctata</i>
			<i>Harmonia quadripunctata</i> Pontoppidan, 1763 – f. <i>typica</i>
			<i>Hippodamia septemmaculata</i> DeGeer, 1775 – f. <i>cestiva</i>
			<i>Hippodamia tredecimpunctata</i> Linnaeus, 1758
			<i>Hippodamia variegata</i> Goeze, 1777 – f. <i>quinquemaculata</i>
			<i>Hippodamia variegata</i> Goeze, 1777 – f. <i>undecimpunctata</i>
			<i>Hippodamia variegata</i> Goeze, 1777 – f. <i>carpini</i>
		<i>Hippodamia variegata</i> Goeze, 1777 – f. <i>constellata</i>	

Coccinellinae Latreille, 1807	Coccinellini Latreille, 1807	<i>Hippodamia variegata</i> Goeze, 1777 – f. <i>velox</i>
		<i>Myrrha octodecimguttata</i> Linnaeus, 1758
		<i>Myzia oblongoguttata</i> Linnaeus, 1758
		<i>Oenopia conglobata</i> Linnaeus, 1758
		<i>Oenopia doublieri</i> Mulsant, 1846
		<i>Oenopia impustulata</i> Linnaeus, 1767
		<i>Oenopia lyncea</i> Olivier, 1808
		<i>Propylea quatuordecimpunctata</i> Linnaeus, 1758 – f. <i>weisei</i>
		<i>Sospita vigintiguttata</i> Linnaeus, 1758
	Epilachnini Chevrolat in Dejean, 1837	<i>Henosepilachna argus</i> Geoffroy in Fourcroy, 1785
		<i>Henosepilachna elaterii</i> Rossi, 1794
		<i>Subcoccinella vigintiquatuorpunctata</i> Linnaeus, 1758 – f. <i>limbata</i>
		<i>Subcoccinella vigintiquatuorpunctata</i> Linnaeus, 1758 – f. <i>typica</i>

cluded in the key (47 species and 19 intraspecies colour forms). Several species were removed from this list: introduced and non-acclimated species (*Chilocorus kuwanae* Silvestri, 1909; *C. nigritus* Fabricius, 1798; *C. stigma* Say, 1835; *Hippodamia convergens* Guérin-Ménéville, 1842; *Olla v-nigrum* Mulsant, 1866), or doubtful records (*Anisosticta strigata* Thünberg, 1795; *Cynegetis impunctata* Linnaeus, 1767). *Platynaspis luteorubra* Goeze, 1777 (Chilocorinae, Platynaspini) was also removed due to its small size (2.5–3.5mm). Since only a few discriminating characters are known that are not reliable with colour patterns, *Henosepilachna angusticollis* Reiche, 1862 is not discriminated from its congener *H. argus* Geoffroy in Fourcroy, 1785 in the key.

Specimens were examined in the collection of the Muséum national d'Histoire naturelle, Paris, France (MNHN) and their data are available at <https://science.mnhn.fr/institution/mnhn/collection/ec/search>.

Characters used in the key

A list of 21 morphological characters based on colour and shape is defined, mainly from existing identification keys (Iablokoff-Khinzorian 1982, Bagnée and Branquart 2000, Roy et al. 2013, Declercq et al. 2014) (Table 2). Only characters that are visible to the naked eye or with a $\times 10$ hand lens are included. The character nomenclature follows Roy et al. (2013), except for characters #10, 11, 15, 16, 17, and 18. All characters were treated as discrete.

Interactive identification key construction and statistics

Digitalization of the 47 species was performed using Xper² v.2.3.2 (Ung et al. 2010) and transferred to Xper³ (Vignes-Lebbe et al. 2016). These softwares are dedicated to manage structured taxonomic descriptions, to analyse these descriptions and to produce keys (Kerner et al. 2011, Corvez and Grand 2014, Martin et al. 2015). A wiki and a documentation of Xper³ are available at <http://wiki.xper3.fr/index.php/UserManualXper3>.

Table 2. List of descriptors used in this study.

Pronotum
1. Pronotum colours
2. Pattern on pronotum
3. Number of pronotum patterns
4. Type of pronotum patterns
Elytra
5. Elytra main colour (background)
6. Elytra markings
7. Colour of elytra markings
8. Number of elytra markings
9. Type of elytra markings
10. Number of lateral lines in the elytra markings
11. Number of longitudinal lines in the elytra markings
12. A spot in the first third of the elytra
13. One of the spots reaches the rim of the elytra
14. Cream ring around dots
15. Dark sutural elytra band
16. Scutellar spot
17. Shape of the scutellar spot
18. White marks between the scutellar spot and the elytra basis
19. Distinct rim around the edge of the elytra
20. Elytra covered with short hairs
Underside
21. Small white triangular marks on the underside below both the middle and front legs

An Xper knowledge base is a set of items described using the same model and terminology, and documented by texts and images. In this key there are 66 items covering 47 species and 19 intraspecies colour forms. The descriptive model consists of a hierarchy of descriptors and a chosen terminology for expressing different possible values (states). The descriptors are the 21 morphological characters previously described. Some of them are consistent only if some conditions are true for another descriptor and these dependencies define a hierarchical structure of descriptors (Table 2). The complete terminology (descriptors and states) is documented by images and texts in order to avoid misinterpretation, a crucial point for relevant identifications with the key. Figure 1 presents the description of the species *Coccinella quinquepunctata* following these model and terms.

Xper³ was also used to compare species and genera. For each descriptor, the comparison tests are able to distinguish a pair of items. Three different measures are available (Burguière et al. 2013). The result is displayed as a table with different colours to separate three cases: (1) items have the same values for a given descriptor (= no discrimination), (2) one pair of items is completely distinct for a given descriptor (= total discrimination), (3) at least one pair of items has not equal values for the descriptor, but these values overlap (= partial discrimination). For a given descriptor the sum of the comparison for all pairs of items is a measure of its ability to distinct taxa (discriminatory power).

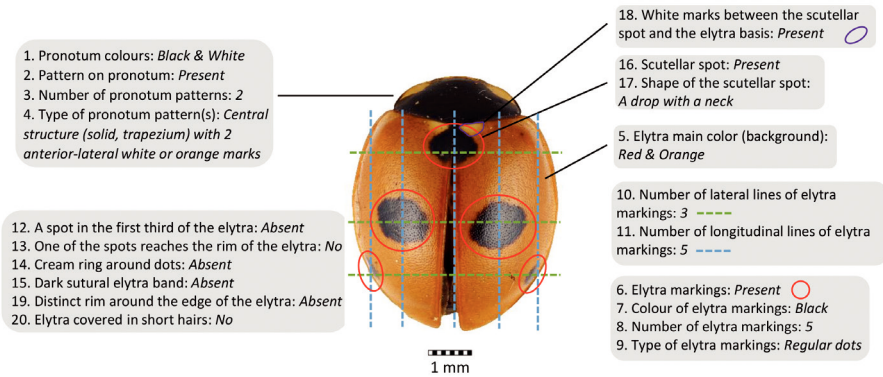


Figure 1. Description of *Coccinella quinquepunctata* following the list of terms used in this study.

The discriminatory power, which represents the quantitative assessments of the ability of a descriptor to distinguish taxa, is measured with the Xper original index (Lebbe 1991) implemented in the Xper² software. This index is based on the incompatibility between descriptions. Two taxa are incompatible (or dissimilar or discriminated) if for one given descriptor there are no states of descriptors in common. For each descriptor the index value ranges between 0 (null discriminatory power of the descriptor) and 1 (the descriptor can distinguish all taxa).

Comparisons within and between genera are made with the “compare groups” and “compare items” options of Xper³. For the comparison between genera, we estimated the number of discriminating characters, weighted or not by the number of colour forms. A subset of descriptors sufficient to discriminate the total of descriptors with the same efficiency was also calculated with the “minset” tool (Lebbe and Vignes 1992, Ziani et al. 1994).

Comparison with standard keys

Two types of keys are available: free-access keys and single-access keys (Hagedorn et al. 2010). A free-access key is a very flexible identification key allowing the user to choose the characters he or she wants to describe. Another web service (Ikey+) (Burguière et al. 2013) is a single-access key builder. A single-access key is a classical key in which descriptors are ordered steps. The topology of the key is a tree and it is possible to compute some indices on the tree: number of maximal steps, length of the paths, etc.

A single-access identification key was generated by IKey+ under Xper³ with the default option and the Xper score method. In this case we show four statistics by taxon: the number of steps, the length of the shortest and the longest paths, and the average length of paths. This key was then compared with five single-access keys for European ladybirds (Dauguet 1949, Bagnée and Branquart 2000, Le Monnier and Livory 2003, Roy et al. 2013, Declercq et al. 2014).

Results

Structure and analysis of the key

The consistency of the knowledge base has been tested with the “Checkbase” functionality of Xper³: no items share the same description and all items are described. The base is 100% complete. Twenty-one descriptors are used: five do not have any dependence (either father or son), four are parent descriptors (for which two are also child descriptors) and 14 are child descriptors (for which two are also parent descriptors). Ninety-eight states are described (minimal/maximal/average number of states: 2/12/4.67).

Discriminatory power of descriptors (Table 3)

The four most discriminating characters (XPER index >0.8) are the type of pronotum patterns (#5), the number of elytra markings (#8), and the number of lateral (#10) and longitudinal (#11) lines of elytra markings. These characters can separate taxa in 7 to 13 groups. For example, the two most discriminating characters (#5 and #8) split all the remaining taxa in 10–13 different groups including 2–13 taxa per group.

The characters #14 and #21 are the least discriminating as they both have an XPER index below 0.8. These characters are binary and split all taxa in two unequal groups (60 vs 2 for the character #14, 65 vs 1 for the character #21). Despite its weak discriminating power, the character #21 is the only one that can distinguish the two species *Coccinella septempunctata* and *C. magnifica*. Eleven descriptors are sufficient to separate all taxa (Table 3, in bold).

Comparison within and between genera

Comparison within a genus: *Coccinella* (Table 4)

Among the 21 characters, 12 are informative (in blue) whereas the other nine are constant and cannot discriminate within this genus (in red). The intersection column shows what is constant in *Coccinella*, therefore helping with the description of the genus: black and white pronotum with two patterns (Central structure - solid, trapezium with two anterior-lateral white or orange marks), elytra with different markings, but always devoid of rim around the edge, short down hairs, cream rings around dots, or dark sutural band.

Comparison between the 24 genera included in the study (Table 5)

Among the genera with at least two species studied, the most constant are *Coccinula* (5% of discriminating characters), *Henosepilachna* (14%) and *Chilocorus* (19%); the most variable are *Adalia* (76%) and *Harmonia* (67%). If weighted by the number of described colour forms per genera, the most constant are still *Coccinula* and *Henosepilachna*; whereas the most variable genera are *Ceratomegilla* and *Exochomus*.

Table 3. List of the discriminating power efficiency of descriptors. Those written in bold are sufficient for discriminating all taxa.

Category	Descriptor	XPER index	Number of descriptor states
Pronotum	4. Type of pronotum pattern(s)	0.85	10
Elytra	8. Number of elytra markings	0.85	12
Elytra	10. Number of lateral lines of elytra markings	0.83	8
Elytra	11. Number of longitudinal lines of elytra markings	0.82	7
Pronotum	3. Number of pronotum patterns	0.71	9
Elytra	17. Shape of the scutellar spot	0.68	4
Elytra	5. Elytra main colour (background)	0.53	7
Elytra	7. Colour of elytra markings	0.52	6
Elytra	13. One of the spots reaches the rim of the elytra	0.49	2
Elytra	18. White marks between the scutellary spot and the elytra basis	0.47	2
Elytra	16. Scutellar spot	0.46	2
Elytra	12. A spot in the first third of the elytra	0.4	2
Elytra	9. Type of elytra markings	0.32	6
Elytra	15. Dark sutural elytra band	0.18	2
Elytra	19. Distinct rim around the edge of the elytra	0.14	2
Elytra	6. Elytra markings	0.12	2
Elytra	20. Elytra covered in short hairs	0.12	2
Pronotum	2. Pattern on pronotum	0.11	2
Pronotum	1. Pronotum colours	0.09	7
Elytra	14. Cream ring around dots	0.06	2
Underside	21. Small white triangular marks on the underside below both the middle and front legs	0.03	2

Single-access identification key and comparison with standard keys

For each identification, the descriptive statistics of the generated key (Appendix 1) are: mean 4.2 steps (2–7), 1–4 paths leading to a taxon (mean 1.5). Unlike many other single-access keys, lots of steps for identifying a taxon do not follow the taxonomy. This is the case in the three tribes: for instance, the user can follow five different paths for identifying an Epilachnini species (in green). The same reasoning applies to *Coccinella* species (in red) with six different paths, and the colour forms of *A. decempunctata* (marked with a yellow star) with six different paths (Figure 2).

In comparison to other standard keys (Table 6), this newly generated key is more efficient for finding the taxon, despite its highest number of included species, except for *Coccinula quatuordecimpustulata* and the key of Dauguet (1949). For example, only five steps are required in the generated key for identifying *Coccinella septempunctata*, whereas 8–14 steps are needed in the other keys.

Table 4. Comparison within genus: *Coccinella* (six species), showing informative characters (in bold) and constant characters (in regular). The intersection column shows what is constant within the genus.

	<i>Coccinella venusta</i>	<i>Coccinella hieroglyphica</i>	<i>Coccinella magnifica</i>	<i>Coccinella quinquepunctata</i>	<i>Coccinella septempunctata</i>	<i>Coccinella undecimpunctata</i>	UNION	INTERSECTION
1. Pronotum colours	Black; White							
2. Elytra main colour (background)	Black; Orange	Black; Orange	Red; Orange	Red; Orange	Red; Orange	Red; Orange	Black; Orange; Red	Orange
3. Pattern on pronotum	Present							
4. Number of pronotum patterns	2							
5. Type of pronotum patterns	Central structure (solid, trapezium) with 2 anterior-lateral white or orange marks							
6. Elytra markings	Present							
7. Colour of elytra markings	Red; Orange; Black	Black	Black	Black	Black	Black	Red; Orange; Black	Black
8. Number of elytra markings	1; 2	7	7	5	7	10–14	1; 2; 5; 7; 10–14	
9. Type of elytra markings	Other	Ovoid shape spot; Other	Regular dots	Regular dots	Regular dots	Regular dots	Ovoid shape spot; Regular dots; Other	
10. Number of lateral lines of elytra markings	not applicable	3	3	3	3	3	3	
11. Number of longitudinal lines of elytra markings	not applicable	4	5	5	5	5	4; 5	
12. Scutellar spot	Absent	Absent	Present	Present	Present	Present	Absent; Present	
13. Distinct rim around the edge of the elytra	absent							
14. Elytra covered in short hairs	No							

	<i>Coccinella venusta</i>	<i>Coccinella hieroglyphica</i>	<i>Coccinella magnifica</i>	<i>Coccinella quinquepunctata</i>	<i>Coccinella septempunctata</i>	<i>Coccinella undecimpunctata</i>	UNION	INTERSECTION
15. Small white triangular marks on the underside below both the middle and front legs	Absent	Absent	Present	Absent	Absent	Absent	Absent; Present	
16. Cream ring around dots	Absent							
17. Dark sutural elytra band	Absent							
18. Shape of the scutellar spot	not applicable	not applicable	A drop (spot with a neck)	A drop (spot with a neck)	A drop (spot with a neck)	A drop (spot with a neck)	A drop (spot with a neck)	
19. White marks between the scutellar spot and the basis of elytra	not applicable	not applicable	Present	Present	Present	Present	Present	
20. One of the spots reaches the rim of the elytra	No	Yes	No	No	No	No	Yes; No	
21. A spot in the first third of the elytra	Absent	Present	Present	Absent	Present	Present	Absent; Present	

Table 5. Comparison between the 24 genera included in the study, showing the most constant or variable genus, weighted or not by the intraspecific variability taken into account in this study (number of colour forms).

	Number of species studied	Number of colour forms studied	Number of discriminating characters within genus	
			with all colours forms	weighted by the number of colour forms
<i>Adalia</i>	3	14	16 (76%)	1.1
<i>Calvia</i>	3	3	5 (23%)	1.7
<i>Ceratomegilla</i>	4	4	13 (62%)	3.3
<i>Chilocorus</i>	2	2	4 (19%)	2
<i>Coccinella</i>	6	6	11 (52%)	1.8
<i>Coccinula</i>	2	2	1 (5%)	0.5
<i>Exochomus</i>	2	2	7 (33%)	3.5
<i>Harmonia</i>	3	6	14 (67%)	2.3
<i>Henosepilachna</i>	2	2	3 (14%)	1.5
<i>Hippodamia</i>	3	7	10 (48 %)	1.4
<i>Oenopia</i>	4	4	10 (48%)	2.5
<i>Subcoccinella</i>	1	2	–	–
<i>Anatis</i>	1	1	–	–
<i>Anisosticta</i>	1	1	–	–
<i>Aphidecta</i>	1	1	–	–
<i>Halyzia</i>	1	1	–	–
<i>Myrrha</i>	1	1	–	–
<i>Myzia</i>	1	1	–	–
<i>Parexochomus</i>	1	1	–	–
<i>Propylea</i>	1	1	–	–
<i>Psyllobora</i>	1	1	–	–
<i>Sospita</i>	1	1	–	–
<i>Tytthaspis</i>	1	1	–	–
<i>Vibidia</i>	1	1	–	–

Discussion

The work presented in this study led to the release of the first multi-access interactive digital identification key for French ladybirds. The adaptability and great number of possibilities provided by this new generation tool are unparalleled for this group, and make the key very flexible and abundantly illustrated and described, thanks to images and texts. Since it is available online and open to experts for modification, the identification key can easily be improved. It will be possible to add ladybird taxa and to extend the geographic area (e.g., a key to all European ladybirds).

Most classical and single-access keys share characters that are quite difficult to observe for students, naturalists and the general public (e.g. for ladybirds in Dauguet 1949 or Roy et al. 2013: mandibles, tooth on tibia, tarsal claws, mesosternal epimera, abdominal post-coxal lines). Here, all taxa are distinguishable with only 11 characters focusing on markings (number and shape). All characters used in this new key are visible to the naked eye or with a x10 hand lens; therefore this tool is designed for non-

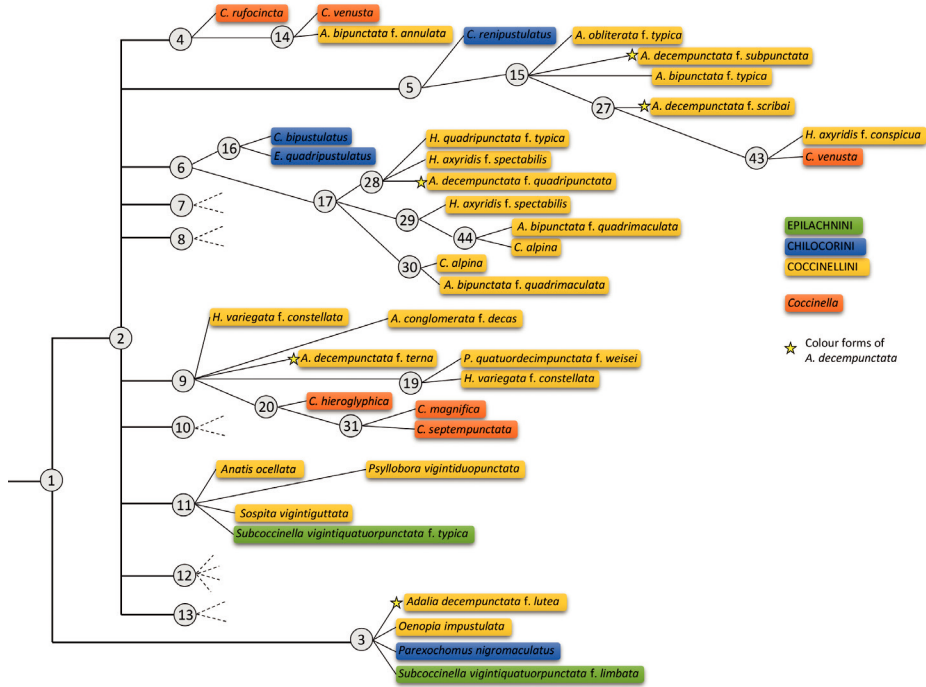


Figure 2. Representation of a part of the single-access identification key generated by IKey+ under Xper³ and the Xper score method (statistics detailed in Appendix 1). The taxonomy is highlighted (the three tribes included in this study, the genus *Coccinella* and the colour forms of *Adalia decempunctata*). Numbers in the circles represent the number of steps in the generated key.

specialists. Using this key, most species can be identified through pictures only, as it is already the case in the identification key for the photographic survey of flower visitors (Spipoll citizen science program, www.spipoll.org), also built with Xper³.

Identification in the field is traditionally realised with paper-printed keys, but recent developments of mobile devices make it possible to use portable version of digital keys. In this perspective, the project “KeyToNature” (www.keytonature.eu) aims to develop new, more convenient and paper-free identification tools, for use within schools and universities across Europe and available on a variety of mobile platforms (laptops, smartphones). However these keys are single access, e.g., MobileKey (Weber and Hagedorn 2010) or the iRecord Ladybirds mobile application, and do not support all the possibilities provided by the free-access keys. Recently, a set of software and applications for transferring the information present in a local Xper database to an Android application has been developed (Troudet 2012). For example, the Malaco-fr application provides an interactive way to identify French snails in the field and without internet connection (Gargominy and Ripken 2011). The free-access key of French ladybirds described in this study will soon be transferred to such a mobile application,

Table 6. Comparison between five single-access keys from bibliography and the generated key by IKey+ under Xper³ for 13 ladybird species: average number of steps (number of paths, length of the shortest and the longest paths).

	Generated Key - this study 47 species	Dauguet (1949) 35 species	Bagnée and Branquart (2000) 37 species	Le Monnier and Livory (2003) 35 species	Roy et al. (2013) 26 species	Derolez et al. (2014) 35 species
<i>Adalia bipunctata</i> f. <i>typica</i>	4 (1)	5 (1)	8 (1)	11 (1)	7 (1)	9 (1)
<i>Anatis ocellata</i>	3 (2)	6 (1)	4.6 (3, 3–7)	4 (1)	8 (3, 6–10)	8 (2, 7–9)
<i>Anisosticta novemdecimpunctata</i>	3 (1)	5 (1)	9 (1)	5 (1)	10 (1)	5.5 (2, 3–8)
<i>Calvia quatuordecimguttata</i>	5 (1)	11 (1)	12 (1)	10 (1)	8 (1)	11 (1)
<i>Chilocorus bipustulatus</i>	4 (1)	-	7 (1)	4 (1)	4 (1)	8 (1)
<i>Coccinella hieroglyphica</i>	4 (2)	10 (1)	6.5 (4, 4–9)	-	8 (1)	8 (2, 7–9)
<i>Coccinella quinquepunctata</i>	3 (1)	12 (1)	8 (1)	-	10 (1)	13 (1)
<i>Coccinella septempunctata</i>	5 (1)	13 (1)	8 (1)	12 (1)	11 (1)	14 (1)
<i>Coccinella undecimpunctata</i>	5 (1)	11 (1)	8.5 (2, 8–9)	11 (1)	10 (1)	10.5 (2, 10–11)
<i>Coccinula quatuordecimpustulata</i>	6 (2)	10 (1)	5 (1)	14 (1)	-	11 (1)
<i>Myzia oblongoguttata</i>	5.5 (2, 5–6)	6 (1)	6 (1)	6 (1)	7 (1)	7 (1)
<i>Oenopia conglobata</i>	4 (1)	10 (1)	11 (1)	9 (1)	-	6 (1)
<i>Psyllobora vigintiduopunctata</i>	3 (1)	8 (1)	11 (1)	12 (1)	12 (1)	7 (1)

which will make it possible to use this tool in educative programs such as “Vigie-nature école” (<https://www.vigienature-ecole.fr>).

Both experienced and inexperienced users are likely to succeed in identifying problematic species (Morse et al. 1996, Drinkwater 2009) if they use convenient, multi-access interactive digital keys. The general public can discriminate species and several colour forms only with colour/form characters and a ×10 lens. Moreover, Xper³ provides possibilities for collaborative work and editing through its website. All the online features make it possible to update the knowledge base easily (e.g. adding numerous colour forms or acclimated species, or new characters such as ecological data). This kind of key aims to increase the curiosity of the general public, and to collect more data on the biology and distribution of species.

Acknowledgements

We thank Clémentine Dantard for the illustrations, and Frédéric Chevillot, Fabrice Duvivier, Jean-François Godeau, David Gonzales, Donald Hobern, Guillaume Jacquemin, Julien Aït El Mekki, Andgelo Mombert, and Gilles San Martin for providing field pictures of the species. We also thank Sylvain Bouquin for his technical help with the Xper system, and Estelle Bourdon and Donald Davesne for proofreading.

References

- Ali A, Hussain A, Baloch MA, Ahmed S (2014) Population trend of canola aphid, *Lipaphis erysimi* (Kalt) (Homoptera: Aphididae) and its associated bio-control agent, *Coccinella septempunctata* (Linnaeus) (Coleoptera: Coccinellidae) in different brassica lines. *Journal of Biology, Agriculture and Healthcare* 4: 19–22.
- Baugnée JY, Branquart E (2000) Clef de terrain pour la reconnaissance des principales coccinelles de Wallonie (Chilocorinae, Coccinellinae & Epilachninae). *Jeunes & Nature Association and Faculté Universitaire des Sciences Agronomiques de Gembloux*, 43 pp.
- Baugnée JY, Branquart E, Maes D, Segers S (2011) Veldbepalingsatlas voor de lieveheersbeestjes van België en Nederland (Chilocorinae, Coccinellinae, Epilachninae & Coccidulinae): herziene druk met larventabel. *Jeugdbond voor Natuur en Milieu (Gent), Jeunes & Nature Association (Wavre) i.s.m. het Instituut voor Natuur- en Bosonderzoek, (Brussels)*, 79 pp.
- Brown PMJ, Roy HE, Rothery P, Roy DB, Ware RL, Majerus MEN (2008) *Harmonia axyridis* in Great Britain: analysis of the spread and distribution of a non-native coccinellid. *Bio-Control* 53: 55–67. <https://doi.org/10.1007/s10526-007-9124-y>
- Burguière T, Causse F, Ung V, Vignes-Lebbe R (2013) IKey+: A New Single-Access Key Generation Web Service. *Systematic Biology* 62: 157–161. <https://doi.org/10.1093/sysbio/sys069>
- Chapin JB, Brou VA (1991) *Harmonia axyridis* (Pallas), the third species of the genus to be found in the United States (Coleoptera: Coccinellidae). *Proceedings of the Entomological Society of Washington* 93: 630–635.
- Comont RF, Roy HE, Harrington R, Shortall CR, Purse BV (2014) Ecological correlates of local extinction and colonisation in the British ladybird beetles (Coleoptera: Coccinellidae). *Biological Invasions* 16: 1805–1817. <https://doi.org/10.1007/s10530-013-0628-3>
- Comont RF, Roy HE, Lewis OT, Harrington R, Shortall CR, Purse BV (2012) Using biological traits to explain ladybird distribution patterns. *Journal of Biogeography* 39: 1772–1781. <https://doi.org/10.1111/j.1365-2699.2012.02734.x>
- Corvez A, Grand A (2014) Enabling comparisons of characters using an Xper2 based knowledge-base of fern morphology. *Phytotaxa* 183: 145–158. <https://doi.org/10.11646/phytotaxa.183.3.2>
- Costello M, Bouchet P, Emblow C, Legakis A (2006) European marine biodiversity inventory and taxonomic resources: state of the art and gaps in knowledge. *Marine Ecology Progress Series* 316: 257–268. <https://doi.org/10.103354/meps316257>
- Dallwitz MJ, Paine TA, Zurcher EJ (2013) Principles of Interactive Keys. 20 pp. <http://delta-intkey.com/www/interactivekeys.pdf> [April 21, 2017]
- Dauguet P (1949) Les Coccinellini de France. *Éditions de l'entomologiste*, Paris, 46 pp.
- Derolez B, Orczyk N, Declercq S (2014) Clé de détermination d'identification des coccinelles du Nord-Pas-de-Calais, version 4.1, 84 pp. http://gon.fr/gon/wp-content/uploads/2015/03/cle_cox_NPdC_version4_1.pdf
- Drinkwater R (2009) Insights into the development of online plant identification keys based on literature review: an exemplar electronic key to Australian *Drosera*. *Bioscience Horizons* 2: 90–96. <https://doi.org/10.1093/biohorizons/hzp007>

- Fontaine B, van Achterberg K, Alonso-Zarazaga MA, Araujo R, Asche M, Aspöck H, Aspöck U, Audisio P, Aukema B, Bailly N (2012) New species in the Old World: Europe as a frontier in biodiversity exploration, a test bed for 21st century taxonomy. *PLoS ONE* 7: e36881. <https://doi.org/10.1371/journal.pone.0036881>
- Gardiner MM, Allee LL, Brown PM, Losey JE, Roy HE, Smyth RR (2012) Lessons from lady beetles: accuracy of monitoring data from US and UK citizen-science programs. *Frontiers in Ecology and the Environment* 10: 471–476. <https://doi.org/10.10189/110185>
- Gargominy O, Ripken TEJ (2011) Une collection de référence pour la malacofaune terrestre de France. *MalaCo Hors Série* 1: 1–108.
- Gordon RD (1985) The Coccinellidae (Coleoptera) of America north of Mexico. *Journal of the New York Entomological Society* 93: 1–912.
- Hagedorn G, Rambold G, Martellos S (2010) Types of identification keys. Tools for identifying biodiversity: progress and problems. *Nimis P.L. & Vignes-Lebbe R., Trieste, Italy*, 59–64.
- Hebert PDN, Cywinska A, Ball SL, deWaard JR (2003) Biological identifications through DNA barcodes. *Proceedings of the Royal Society B: Biological Sciences* 270: 313–321. <https://doi.org/10.1098/rspb.2002.2218>
- Hemptinne JL, Magro A, Majerus MEN (2005) *Les Coccinelles*. Delachaux et Niestlé, Paris, 190 pp.
- Hodek I, Honěk A (2009) Scale insects, mealybugs, whiteflies and psyllids (Hemiptera, Sternorrhyncha) as prey of ladybirds. *Biological Control* 51: 232–243. <https://doi.org/10.1016/j.biocontrol.2009.05.018>
- Iablokoff-Khnzorian SM (1982) *Les coccinelles, Coléoptères Coccinellidae. Tribu Coccinellini des régions Paléarctiques et Orientale*. Boubée, Paris, 568 pp.
- de Jong Y, Verbeek M, Michelsen V, Bjørn P de P, Los W, Steeman F, Bailly N, Basire C, Chylarecki P, Stloukal E, Hagedorn G, Wetzel F, Glöckler F, Kroupa A, Korb G, Hoffmann A, Häuser C, Kohlbecker A, Müller A, Güntsch A, Stoev P, Penev L (2014) *Fauna Europaea – all European animal species on the web*. *Biodiversity Data Journal* 2: e4034. <https://doi.org/10.103897/BDJ.2.e4034>
- Kerner A, Debrenne F, Vignes-Lebbe R (2011) Cambrian archaeocyathan metazoans: revision of morphological characters and standardization of genus descriptions to establish an online identification tool. *ZooKeys* 150: 381–395. <https://doi.org/10.103897/zookeys.150.1566>
- Le Monnier Y, Livory A (2003) *Atlas des coccinelles de la Manche: une enquête Manche-Nature 1998–2001*. *Manche-Nature*, Coutances, 208 pp.
- Lebbe J (1991) *Représentation des concepts en biologie et en médecine: introduction à l'analyse des connaissances et à l'identification assistée par ordinateur*. PhD thesis, Université Pierre et Marie Curie Paris, France.
- Lebbe J, Vignes R (1992) Sélection d'un sous ensemble de descripteurs maximale discriminant dans une base de connaissances, 3^{ème} journée Symbolic-Numérique, Paris, 219–232.
- Martin P, Schmelz RM, Dole-Olivier M-J (2015) Groundwater oligochaetes (Annelida, Clitellata) from the Mercantour National Park (France), with the descriptions of one new genus and two new stygobiont species. *Zoosystema* 37: 551–570. <https://doi.org/10.105252/z2015n4a2>

- Morse D, Tardival GM, Spicer J (1996) A Comparison of the Effectiveness of a Dichotomous Key and a Multi-access Key to Woodlice. <http://kar.kent.ac.uk/21343/1/WoodliceMorse.pdf> [Accessed November 01, 2017]
- Oliver J (1988) Crisis in biosystematics of arthropods. *Science* 240: 967–967. <https://doi.org/10.1016/science.3368789>
- Purse BV, Comont R, Butler A, Brown PMJ, Kessel C, Roy HE (2015) Landscape and climate determine patterns of spread for all colour morphs of the alien ladybird *Harmonia axyridis*. *Journal of Biogeography* 42: 575–588. <https://doi.org/10.1111/jbi.12423>
- Roy HE, Brown PMJ, Comont RF, Poland RL, Sloggett JJ, Majerus M, Kearns PWE (2013) *Naturalists' Handbook 10: Ladybirds*. Pelagic Publishing, Exeter, 142 pp.
- SCBD (2010) What is the Problem? The Taxonomic Impediment. <https://www.cbd.int/gti/problem.shtml> [Accessed 16 April 2018]
- Seago AE, Giorgi JA, Li J, Ślipiński A (2011) Phylogeny, classification and evolution of ladybird beetles (Coleoptera: Coccinellidae) based on simultaneous analysis of molecular and morphological data. *Molecular Phylogenetics and Evolution* 60: 137–151. <https://doi.org/10.1016/j.ympev.2011.03.015>
- Silvertown J (2009) A new dawn for citizen science. *Trends in Ecology & Evolution* 24: 467–471. <https://doi.org/10.1016/j.tree.2009.03.017>
- Smith MA, Rodriguez JJ, Whitfield JB, Deans AR, Janzen DH, Hallwachs W, Hebert PDN (2008) Extreme diversity of tropical parasitoid wasps exposed by iterative integration of natural history, DNA barcoding, morphology, and collections. *Proceedings of the National Academy of Sciences* 105: 12359–12364. <https://doi.org/10.1073/pnas.0805319105>
- Tronquet M (2014) *Catalogue des Coléoptères de France*. Association Roussillonnaise d'Entomologie, 1052 pp.
- Troudet J (2012) Des bases de connaissances en biodiversité aux outils de diffusion sur smartphones. Stage de Master 2, Systématique et Évolution, Muséum national d'Histoire naturelle, 30 pp.
- Ung V, Dubus G, Zaragüeta-Bagils R, Vignes-Lebbe R (2010) Xper2: introducing e-taxonomy. *Bioinformatics* 26: 703–704. <https://doi.org/10.1093/bioinformatics/btp715>
- Vander Zanden MJ, Hansen GJA, Higgins SN, Kornis MS (2010) A pound of prevention, plus a pound of cure: Early detection and eradication of invasive species in the Laurentian Great Lakes. *Journal of Great Lakes Research* 36: 199–205. <https://doi.org/10.1016/j.jglr.2009.11.002>
- Vignes-Lebbe R, Chesselet P, Diep Thi M-H (2016) Xper3: new tools for collaborating, training and transmitting knowledge on botanical phenotypes. In: *Botanists of the twenty-first century: roles, challenges and opportunities*. Noëline R. Rakotoarisoa, Stephen Blackmore and Bernard Riera, Paris, 228–239.
- Weber G, Hagedorn G (2010) Wiki keys on mobile devices. In: *Tools for Identifying Biodiversity: Progress and Problems*. Nimis P. L., Vignes Lebbe R., Trieste, Italy, 89–93.
- Ziani D, Khalil Z, Vignes R (1994) Finding a minimum subset of variables from symbolic objects. *Proceedings IPMU, 5th international conference*, Springer-Verlag Berlin Heidelberg, Paris, France, 794–799.

Appendix I

STATISTICS				
Taxon	Number of paths leading to taxon	Length of the shortest path leading to taxon	Average length of paths leading to taxon	Length of the longest path leading to taxon
<i>Adalia bipunctata</i> f. <i>annulata</i>	1	4	4	4
<i>Adalia bipunctata</i> f. <i>pantherina</i>	1	4	4	4
<i>Adalia bipunctata</i> f. <i>quadrimaculata</i>	2	5	5.5	6
<i>Adalia bipunctata</i> f. <i>sexpustulata</i>	2	3	3.5	4
<i>Adalia bipunctata</i> f. <i>typica</i>	1	4	4	4
<i>Adalia conglomerata</i> f. <i>decas</i>	1	3	3	3
<i>Adalia decempunctata</i> f. <i>decempustulata</i>	2	7	7	7
<i>Adalia decempunctata</i> f. <i>guttatopunctata</i>	4	7	7	7
<i>Adalia decempunctata</i> f. <i>lutea</i>	1	2	2	2
<i>Adalia decempunctata</i> f. <i>quadripunctata</i>	1	5	5	5
<i>Adalia decempunctata</i> f. <i>scribai</i>	1	5	5	5
<i>Adalia decempunctata</i> f. <i>subpunctata</i>	1	4	4	4
<i>Adalia decempunctata</i> f. <i>terna</i>	1	3	3	3
<i>Adalia decempunctata</i> f. <i>typica</i>	2	6	6	6
<i>Anatis ocellata</i>	2	3	3	3
<i>Anisosticta novemdecimpunctata</i>	1	3	3	3
<i>Aphidecta oblitterata</i> f. <i>typica</i>	2	2	3	4
<i>Calvia decemguttata</i>	2	6	6	6
<i>Calvia quatuordecimguttata</i>	1	5	5	5
<i>Calvia quindecimguttata</i>	2	6	6	6
<i>Ceratomegilla alpina</i>	2	5	5.5	6
<i>Ceratomegilla notata</i>	2	6	7	8
<i>Ceratomegilla rufocincta</i>	1	3	3	3
<i>Ceratomegilla undecimnotata</i>	1	8	8	8
<i>Chilocorus bipustulatus</i>	1	4	4	4
<i>Chilocorus renipustulatus</i>	1	3	3	3
<i>Coccinella hieroglyphica</i>	2	4	4	4
<i>Coccinella magnifica</i>	1	5	5	5
<i>Coccinella quinquepunctata</i>	1	3	3	3
<i>Coccinella septempunctata</i>	1	5	5	5
<i>Coccinella undecimpunctata</i>	1	5	5	5
<i>Coccinella venusta</i>	2	4	5	6
<i>Coccinula quatuordecimpustulata</i>	2	6	6	6
<i>Coccinula sinuatomarginata</i>	1	5	5	5
<i>Exochomus octosignatus</i>	1	2	2	2
<i>Exochomus quadripustulatus</i>	1	4	4	4

STATISTICS				
Taxon	Number of paths leading to taxon	Length of the shortest path leading to taxon	Average length of paths leading to taxon	Length of the longest path leading to taxon
<i>Halyzia sedecimguttata</i>	1	5	5	5
<i>Harmonia axyridis</i> f. <i>conspicua</i>	2	6	6	6
<i>Harmonia axyridis</i> f. <i>novendecimsignata</i>	2	4	4.5	5
<i>Harmonia axyridis</i> f. <i>spectabilis</i>	2	5	5	5
<i>Harmonia conformis</i>	3	3	4	5
<i>Harmonia quadripunctata</i> f. <i>sedecimpunctata</i>	2	3	3	3
<i>Harmonia quadripunctata</i> f. <i>typica</i>	1	5	5	5
<i>Henosepilachna argus</i>	2	4	5	6
<i>Henosepilachna elaterii</i>	1	4	4	4
<i>Hippodamia septemmaculata</i> f. <i>cestiva</i>	2	3	3.5	4
<i>Hippodamia tredecimpunctata</i>	1	4	4	4
<i>Hippodamia variegata</i> f. <i>undecimpunctata</i>	2	5	6	7
<i>Hippodamia variegata</i> f. <i>quinquemaculata</i>	1	3	3	3
<i>Hippodamia variegata</i> f. <i>carpini</i>	1	4	4	4
<i>Hippodamia variegata</i> f. <i>constellata</i>	2	3	3.5	4
<i>Hippodamia variegata</i> f. <i>velox</i>	2	5	5.5	6
<i>Myrrha octodecimguttata</i>	1	5	5	5
<i>Myzia oblongoguttata</i>	4	5	5.5	6
<i>Oenopia conglobata</i>	1	4	4	4
<i>Oenopia dublieri</i>	1	4	4	4
<i>Oenopia impustulata</i>	1	2	2	2
<i>Oenopia lyncea</i>	4	5	5.5	6
<i>Parexochomus nigromaculatus</i>	1	2	2	2
<i>Propylea quatuordecimpunctata</i> f. <i>weisei</i>	2	4	4	4
<i>Psyllobora vigintiduopunctata</i>	1	3	3	3
<i>Sospita vigintiguttata</i>	1	3	3	3
<i>Subcoccinella vigintiquatuor punctata</i> f. <i>limbata</i>	1	2	2	2
<i>Subcoccinella vigintiquatuor punctata</i> f. <i>typica</i>	1	3	3	3
<i>Tytthaspis sedecimpunctata</i> f. <i>duodecimpunctata</i>	2	5	5.5	6
<i>Vibidia duodecimguttata</i>	1	5	5	5
AVERAGE	1.53	4.197	4.364	4.53

Annexe 2 : Descripteurs dans la clé des larves d'odonates (T1)

54 descripteurs dont 12 descripteurs-fils

Taille générale

1. Longueur du corps (en mm)

Abdomen

2. Extrémité de l'abdomen

3. Longueur des lamelles caudales

4. Extrémité des lamelles caudales

5. Longueur de la ligne foncée sur le bord des lamelles caudales

6. Trachée médiane des lamelles caudales

7. Longueur relative des lignes foncées sur le bord des lamelles

8. Échancrure horizontale au milieu des lamelles caudales

9. Ratio longueur/largeur des lamelles caudales

10. Lamelle caudale centrale

11. Longueur relative des cerques et des paraproctes

12. Sur le dessus de la pyramide anale, une expansion terminée en pointe

13. Longueur relative de l'expansion de l'épiprocte et des cerques

14. Forme de l'abdomen

15. Largeur de l'abdomen

16. Longueur de l'ovipositeur

17. L'ovipositeur atteint ou dépasse la marge postérieure du sternite 9

18. Soies sur les tergites 7, 8 et 9 de l'abdomen

19. Soies sur la bordure arrière des tergites 8 et 9 de l'abdomen

20. Rangée d'épines sur le second sternite abdominal

21. Point noir médian sur les sternites abdominaux

22. Épine dorsale sur les tergites abdominaux 4 et/ou 5

23. Épine dorsale sur le tergite abdominal 7

24. Épine dorsale sur le tergite abdominal 9

25. Épines latérales du tergite abdominal 5

26. Épines latérales du tergite abdominal 6

27. Épines latérales du tergite abdominal 7

- 28. Longueur relative des épines latérales du tergite abdominal 8
- 29. Longueur des épines latérales des tergites abdominaux 8 et 9
- 30. Taille des épines latérales du tergite abdominal 9 par rapport aux cerques
- 31. Taille des épines latérales du tergite abdominal 9 par rapport au tergite 10
- 32. Dimensions du sternite 10
- 33. Peigne de soies sur le bord postérieur des sternites abdominaux 7 et/ou 8

Tête

- 34. Aspect des antennes
- 35. Longueur du premier article antennaire
- 36. Diamètre de l'œil
- 37. Forme de l'œil
- 38. Côtés latéraux de la tête
- 39. Deux taches claires à l'arrière de la tête

Thorax et pattes

- 40. Épines sur le poststernite du thorax
- 41. Paires de pattes 1 et 2

Masque

- 42. Type de masque
 - 43. Dents des palpes labiaux
- 44. Sillon à la base du masque
- 45. Fente médiane à l'extrémité du masque
- 46. Forme du prémentum du masque
- 47. Entaille profonde sur les palpes labiaux
- 48. Nombre de soies sur le crochet mobile des palpes labiaux
- 49. Nombre de longues soies à l'intérieur du masque
- 50. Dernière dent du peigne du palpe labial
- 51. Partie proéminente du palpe labial
- 52. Ratio longueur/largeur du masque

Tête

- 53. Protubérance(s) entre les deux yeux
- 54. Expansions latérales à l'arrière de la tête

Annexe 3 : Descripteurs dans la clé des larves d'éphémères (T2)

52 descripteurs dont 5 descripteurs-fils

Aspect général du corps

1. Carapace dorsale
2. Aplatissement dorso-ventral
3. Longueur du corps

Thorax et abdomen

4. Position des branchies
 5. Marge antérieure du mésosternum
 6. Ligne claire au milieu du pronotum
7. Branchies visibles dorsalement
8. Nombre de paires de branchies
9. Marge des branchies sclérifiée
10. Aspect des branchies
 11. Branchies 2 à 6 : simples ou doubles
 12. Extrémité des branchies
13. Forme des branchies
14. Épines sur le bord droit des branchies
15. Angles postérieurs des segments abdominaux
16. Forme du pronotum
17. Marges latérales du pronotum
18. Protubérances sur le pronotum
19. Patrons de coloration sur les tergites abdominaux 7 à 9
20. Deux grandes taches foncées diffuses sur le tergite 2
21. Ligne centrale antéro-postérieure foncée sur les tergites abdominaux
22. Bord postérieur du sternite abdominal 9
23. Frange de soies sur la marge postérieure des segments abdominaux 7 et 8
24. Apparence tachetée des tergites abdominaux

Pattes

25. Griffes du tarse sur la première paire de pattes

26. Longueur de la griffe du tarse
27. Dents sur la griffe du tarse
28. Projection dorso-apicale sur le tibia des pattes antérieures
29. Aspect du tibia
30. Tache(s) foncée(s) sur le fémur
31. Bande(s) foncée(s) sur le tarse
32. Longueur des soies sur le bord postérieur du fémur

Extrémité de l'abdomen

33. Localisation des soies sur les cerques
34. Longueur des cerques
35. Taille relative du paracerque
36. Bande sombre unique clairement délimitée vers le milieu des cerques
37. Nombre d'anneaux sombres entre le dernier segment de l'abdomen et la bande sombre sur les cerques
38. Coloration des cerques

Tête

39. Position des yeux
40. Longueur des antennes
41. Expansion sur le premier segment de l'antenne
42. Trois protubérances sur la tête
43. Marques claires irrégulières sur le sommet de la tête

Pièces buccales

44. Processus des mandibules
45. Apex des maxilles
46. Nombre d'articles du palpe labial
47. Segment terminal du palpe labial
48. Taille des soies sur les palpes maxillaires et labiaux
49. Largeur des glosses et paraglosses
50. Forme des glosses
51. Dents sur les mandibules
52. Forme du labre

Annexe 4 : Descripteurs dans la clé des éphémères imagos (T3)

25 descripteurs dont 9 descripteurs-fils

Ailes

1. Ailes postérieures
 2. Processus costal
 3. Nombre de veines longitudinales (postérieures vestigiales)
4. Nervures transverses de l'aile antérieure
5. Points noirs sur l'aile antérieure
6. Nervures intercalaires de l'aile antérieure
7. Taille de l'aile antérieure

Tête

8. Position et taille des yeux composés
9. Coloration des yeux
10. Aspect particulier des yeux

Parties génitales

11. Sexe
 12. Segmentation des pinces génitales
 13. Lobes péniens
 14. Forme de la couverture pénienne
 15. Titillateurs
 16. Sillon médian des pinces génitales

Thorax

17. Prosternum
18. Segments du tarse antérieur le plus grand
19. Coloration du thorax
20. Base des pattes avant

Abdomen

21. Paracerque

22. Processus postéromédian (Tg2)

23. Coloration de l'abdomen

24. Type d'ornementation

25. Coloration (cas des taches)

Annexe 5 : Descripteurs dans la clé des larves de plécoptères (T4)

12 descripteurs, aucun descripteur-fils

Taille générale

1. Longueur du corps

Tête

2. Longueur des glosses et paraglosses

3. Segment terminal des palpes maxillaires

4. Protubérances chitineuses sur les segments basaux des antennes

5. Position des yeux

Thorax

6. Longueur des articles du tarse

7. Longueur des pattes arrière

8. Orientation des fourreaux alaires

9. Articles 1 et 3 du tarse

Abdomen

10. Protubérances dorsales sur l'abdomen

11. Extrémité des paraproctes

12. Nombre de segments abdominaux présentant deux parties distinctes (tergite et sternite)

Annexe 6 : Descripteurs dans la clé des plécoptères imagos (T5)

13 descripteurs dont 1 descripteur-fils

Taille générale du corps

1. Taille du corps

Tête

2. Glosses / Paraglosses

3. Forme de la tête

Thorax

4. Longueur des articles du tarse

5. Position des ailes au repos

6. Taille des ailes

7. Pièce en X

Abdomen

8. Nombre d'article des cerques

9. Plaque sous/post génitale

10. Gonopore

11. Forme des cerques

12. Détail de la plaque sous génitale

13. Taille des cerques

Annexe 7 : Descripteurs dans la clé des coléoptères aquatiques (T6)

22 descripteurs dont 4 descripteurs-fils

Aspect général

1. Taille de l'organisme
2. Forme générale du corps
3. Ornementation
 4. Sillons
5. Reflets métalliques

Tête

6. Division des yeux
7. Aspect des antennes
8. Longueur des antennes
9. Rapport des palpes et des antennes

Thorax

10. Nombre de segments aux tarses postérieurs
11. Taille des griffes postérieures
12. Motifs du pronotum
 13. Couleur secondaire du pronotum
14. Couleur principale du pronotum
15. Coxa postérieure
 16. Plaque de la coxa postérieure

Élytres

17. Soies des élytres
18. Écusson
19. Motifs des élytres
 20. Couleur secondaire des élytres
21. Couleur principale des élytres
22. Ponctuation

Annexe 8 : Descripteurs dans la clé des coccinelles de France (T7)

21 descripteurs dont 15 descripteurs-fils

Pronotum

1. Pronotum colours
2. Pattern on pronotum
 3. Number of pronotum patterns
 4. Type of pronotum patterns

Elytra

5. Elytra main colour (background)
6. Elytra markings
 7. Colour of elytra markings
 8. Number of elytra markings
 9. Type of elytra markings
 10. Number of lateral lines in the elytra markings
 11. Number of longitudinal lines in the elytra markings
 12. A spot in the first third of the elytra
 13. One of the spots reaches the rim of the elytra
 14. Cream ring around dots
 15. Dark sutural elytra band
 16. Scutellar spot
 17. Shape of the scutellar spot
 18. White marks between the scutellar spot and the elytra basis
 19. Distinct rim around the edge of the elytra
20. Elytra covered with short hairs

Underside

21. Small white triangular marks on the underside below both the middle and front legs

Annexe 9 : Article (à soumettre dans *PLoS ONE* ou *PeerJ*) sur les clés des larves aquatiques et le test de la clé des larves d'odonates

An innovative digital tool to improve the identification of aquatic insects in an urban area (Insecta: Odonata, Ephemeroptera and Plecoptera)

Mathilde Delaunay¹, Régine Vignes-Lebbe¹, Aurélie Dubois¹, Sylvain Bouquin¹, Michel Brulin², Romain Nattier¹

¹ Institut de Systématique, Evolution, Biodiversité, ISYEB – UMR 7205 CNRS MNHN UPMC EPHE, Muséum national d'Histoire naturelle, Sorbonne Universités, 57 rue Cuvier, CP50, 75005 Paris, France

² Office pour les insectes et leur environnement (OPIE) – Chemin rural n°7, La Minière, 78280 Guyancourt, France

Corresponding author: mdelaunay@mnhn.fr

Abstract

[...]

Keywords

Damselflies, Dragonflies, France, Ile-de-France, Interactive identification key, Mayflies, Stoneflies, Xper

1. Introduction

Aquatic habitats – including wetlands and watercourses – provide some crucial ecosystem services like carbon absorption or water filtration (Barbier 2011), and support a high diversity of species (Dudgeon *et al.* 2006). In Europe, more than 10% of insect species strictly depend on these environments during the totality or part of their life cycle (Nilsson 1996a). However those habitats and the fauna and flora they host are greatly threatened by human population activities such as habitat degradation, water pollution and over-exploitation (Dudgeon *et al.* 2006). The largest urban areas are the most affected by this phenomenon, as it has been documented in several European big cities. In the Région Ile-de-France, the highly urbanized area of Paris and its suburbs, 90% of ponds and pools have disappeared since 1900 (SNPN 2012).

Since the beginning of the XXIth century there has been growing awareness of these habitats decline and thus, actions to protect them have been implemented. The conservation of aquatic environments is juridically framed by some texts, e.g. the Water Framework Directive (Directive 2000/60/EC of the European Parliament and of the Council of 23 October 2000 establishing a framework for Community action in the field of water policy) [1] at the European level, and a law about water and aquatic environments (LEMA: Loi n° 2006-1772 du 30 décembre 2006 sur l'eau et les milieux aquatiques) [2] in France.

Efficient protection measures for aquatic environments require advanced knowledge about the species living in these habitats, particularly because biodiversity inventories implemented for several texts of law are based on the presence or absence of certain taxa in a given location – e.g. IBG-DCE, the “Global biological normalized index” for the European Water Framework Directive, which focuses on “aquatic invertebrates” living in running waters, for their value as indicators for water quality. The larva stages are of particular importance because, unlike adults that have wings in most insect orders, larva cannot disperse far from their main place of life and are thus reliable to assess the precise position of populations.

In this context, an accurate identification of the species is essential, as it is for most of ecological studies (species inventories, ecosystem management, establishment and improvement of environmental public policies, taxonomic reviews...) (Vander Zanden *et al.* 2010, Jackson *et al.* 2014, Leal *et al.* 2016, Troudet *et al.* 2017). It is thus of importance for different types of public – naturalists, amateurs and professionals (Farnsworth *et al.* 2013).

However aquatic insect taxa are concerned, first of all, by what we know as the “Taxonomic impediment” (SCBD 2010): due to their small size and the difficulties encountered for their morphological identification – some are even considered cryptic species (Williams *et al.* 2006) – these organisms require expert skills for accurate identification, and there is a generalized lack of experts for aquatic insects. Moreover, we observe a taxonomic bias in biodiversity and ecological data for these taxa. In fact generally, a fairly small number of species attract most of the public and scientific attention, whereas most taxa – including insects – are poorly studied and known, compared to their number of species (Troudet *et al.* 2017). Thus the databases for aquatic insects are still very incomplete, whether for ecological parameters or for fine morpho-anatomy identification, with variations depending on orders (Vieira *et al.* 2006, Jackson *et al.* 2014).

The most basic requirement for people working on biodiversity aspects is the availability of identification guides. However, simple-to-use identification guides for non-taxonomists and the general public are rare and available for relatively few taxonomic groups (SCBD 2010). Consequently, other organisms distribution, biology, ecology and genetics remain poorly known (Costello *et al.* 2006, SCBD 2010). As for other groups, most new species of aquatic insects are described by non-professional taxonomists (Fontaine *et al.* 2012) and the distribution of a majority of organisms remains poorly known. For mayflies, stoneflies, dragonflies and damselflies, citizen science programs (participation of amateurs and general public to the inventory and the description of life) have been created that aim to fill that gap, for example in the USA and the UK (e.g. Inletkeeper program [3], the Dragonfly Pond Watch Project [4] and Anglers’ Riverfly Monitoring Initiative [5]). In France, so far the citizen science programs only focus on dragonflies and damselflies (e.g. the Steli, *Suivi temporel des libellules* that is part of the Vigie-Nature platform (<http://vigienature.mnhn.fr/>) and concerns the entire metropolitan France; *L’observatoire des odonates* managed by the Groupe ornithologique et naturaliste in North of France).

In this perspective and for the programs that have to come, easy-to-use visual and interactive tools for the identification of species offer enormous potential for the general public (Riccamboni *et al.* 2010).

In addition to the identification guides, another type of tools is available as a help for the determination of taxa: the identification keys. The classic and very used single-access identification keys consist of a series of identification steps that form a single and unique identification path for a given taxon. Although they are powerful tools and very useful to identify species, the user cannot choose the character to be observed (the

answer for every single step must be known), and the identification is impossible if some characters are missing – e.g. if the specimen is of poor quality) (Hagedorn *et al.* 2010). Moreover, this type of keys cannot be modulated nor adapted to various kinds of publics, environmental conditions, season, nor geographical location.

Most of European aquatic insects identification keys are difficult to use for the non-specialist, because of the terminology specific to entomology and the small dimension of most of these taxa – e.g. Tachet *et al.* 2010 (for all aquatic invertebrates), Nilsson 1996b, Dijkstra and Lewington 2006 and Grand *et al.* 2014 (for dragonflies and damselflies), Hynes 1977, Nilsson 1996a and Pryce *et al.* 2007 (for stoneflies), Nilsson 1996a, Macadam and Bennett 2010, and Bauernfeind and Soldán 2012 (for mayflies).

Modern tools developed along with digital technologies can make the identification easier for the user. In this perspective, several interactive identification keys are available online for aquatic invertebrates, e.g. that of the *Groupe d'éducation et d'écovigilance de l'eau* for Québec (http://www.g3e-ewag.ca/documents/jeux/cle_macro/cle.html) or the *Perla database* for France (<http://www.perla.developpement-durable.gouv.fr/index.php/arbre/>). Those keys can be abundantly illustrated with pictures and drawings and can include additional resources like hyperlinks and sound files. However, most of them are only digital versions of single-access keys and maintain the same difficulties for the user, in particular having to follow a series of steps in an imposed order.

Conversely, a multi-access interactive key is a computer-aided identification tool that assists with finding correct names of species where the user enters attributes (character-state values) of the specimen. Several benefits over conventional keys have been shown (Edwards and Morse 1995, Dallwitz *et al.* 2013): characters can be used in any order, characters are ordered to start with the one that best separates the remaining taxa, several character states can be selected for a same character – allowing doubts), the keys can be completed with illustrations and texts explaining the terminology used, correct identifications can be obtained despite small errors made by the user. Some softwares also include the possibility to print a single-access key for field identification if needed, as well to weight characters according to the target audience competences and abilities (students, children, general public, naturalists...) (Ung *et al.* 2010).

So far there is no multi-access interactive identification key for aquatic insects of the Région Ile-de-France (Paris area), despite a need for this kind of tools that can be used directly in the field for biodiversity inventories and surveys, which are crucial in highly degraded areas.

In this perspective, the main aim of this study is to produce multi-access digital interactive identification keys at the species level for the orders Ephemeroptera, Odonata and Plecoptera of the Région Ile-de-France, for mature larvae (nymphs). Those orders are used for surveys (ex: IBG-DCE) and enough information is available for larva stages morphology at the species level to complete a morphological database adequate to build a reliable identification key.

In this perspective, the main aims of this study are:

i) to produce multi-access digital interactive identification keys at the species level for the orders Ephemeroptera, Odonata and Plecoptera of the Région Ile-de-France, for mature larvae (nymphs). Those orders are used for surveys (ex: IBG-DCE) and enough information is available for larva stages morphology

at the species level to complete a morphological database adequate to build a reliable identification key. We study and discuss the discrimination power of the characters: what is the minimal number of characters needed to identify each species? What are the most discriminating characters?

ii) to describe a new system that was linked to one of the identification keys to record in details the identification paths followed by the users of the key. A series of tests with different types of users has been realized. We present the analysis of these data and the hypothesis we can make about the use of the key. The results allow to improve and complete the key in order to provide the most reliable and easy-to-use tool possible.

We study and discuss the discrimination power of the characters: what is the minimal number of characters needed to identify each species? What are the most discriminating characters?

In this perspective, we describe a new system that was linked to one of the identification keys to record in details the identification paths followed by the users of the key. A series of tests with different types of users has been realized. We present the analysis of these data and the hypothesis we can make about the use of the key. The results allow to improve and complete the key in order to provide the most reliable and easy-to-use tool possible.

2. Materials and Methods

2.1 Taxonomic coverage for the three keys

We included all the species known to be living at larval stages in the Région Ile-de-France for the three orders Ephemeroptera, Odonata and Plecoptera (Table 1). The list of species mainly comes from recent data (since 1950) with verified occurrence in the field at the larval stages (Office Pour les Insectes et leur Environnement – OPIE, Inventory of French mayflies and Inventory of French stoneflies programs, Société Française d’Odonatologie – SFO, Cettia database, Inventaire cartographique des libellules de France – Invod, Cilif program for Odonata, Inventaire National du Patrimoine Naturel – INPN...). Some species were also added to the list from assumptions of occurrence made by experts of the orders, when the species can be found in areas adjoining the Région Ile-de-France (Michel Brulin, Jacques Le Doaré, pers. comm.).

As we wanted the keys to be the most exhaustive and precise possible, they allow the identification at the species level for mature larvae (last larval stage before emergence) and exuviae.

To complete the descriptions made with data coming from descriptive books and existing keys, some specimens were examined from collections of the OPIE for the order Ephemeroptera.

2.2 Characters used in the keys

For each of the three keys, we defined a list of morphological characters (“descriptors”), mainly from descriptive books and existing identification keys (Hynes 1977, Dijkstra and Lewington 2006, Pryce *et al.* 2007, Doucet 2010, Macadam and Bennett 2010, Tachet *et al.* 2010, Bauernfeind and Soldán 2012, Grand *et al.* 2014) (Table 2). We only included morphological characters that are visible with a binocular magnifier

and do not require dissections; thereby are excluded the observation of genitalia and the tiny mouthparts. All descriptors were treated as “categorical” (qualitative) – (ie. they have at least two descriptor states for the user to choose between), except for four descriptors that were treated as “numerical” (ie. the user has to insert a number to characterize the descriptor) (Table 2).

2.3 IIK construction and statistics

Digitalization of the 108 species was performed using Xper³ (Vignes-Lebbe *et al.* 2016). This software is dedicated to manage structured taxonomic descriptions, to analyze these descriptions and to produce keys (Kerner *et al.* 2011, Corvez & Grand 2014, Martin *et al.* 2015). A wiki and a documentation of Xper³ are available at wiki.xper3.fr/index.php/UserManualXper3.

An Xper knowledge base is a set of items (in our case, species), all described following the same model and terminology, and documented by texts and images. The descriptive model is compounded of a hierarchy of descriptors (discrete and continue characters) and a chosen terminology to express the different possible values (states). The complete terminology (descriptors and states) is documented by images and texts to avoid misinterpretation, a crucial point for relevant identification with the key.

The “discriminant power”, which represents the quantitative assessments of a descriptor ability to distinguish between items, is measured with the Xper original index (Lebbe 1991) implemented in the Xper² software – a previous offline version of the software (Ung *et al.* 2010). This index is based on incompatibility between descriptions. Two taxa are incompatible (or dissimilar or discriminated) if for one descriptor there is no common states of descriptors. For each descriptor, and for all the taxa, the index value is between 0 (no discriminatory power of the descriptor) and 1 (the descriptor can discriminate all taxa). Comparisons within and between genera are made with the “compare groups” and “compare items” options in Xper³.

2.4 The Xperience system for recording identification paths

A new system called Xperience was implemented in the research team. It is possible to link any Xper³ database to this system, then Xperience records every identification step that is validated by the user during the specimen identification. Thereby, a new database is obtained including all the validated choices of the users. The chosen descriptors and descriptor states are recorded – in the right order, and so is the final identification (species proposed by the user as “correct”). Every item, descriptor and descriptor state has a specific number for Xperience (Table 2), allowing the analysis of the determination paths.

2.5 Validation of the use of the key

In order to test the efficiency of the Xperience system, we asked different people to use one of the keys. We chose the key to Odonata larvae, because their identification is the easiest between the three orders and it is relatively easy to find specimens of different species in the field all year long. The 24 testers belonged to five categories of people susceptible to use this kind of identification tool for their personal or professional

activities: 7 professional entomologists, 7 amateur naturalists (members of a naturalistic association), 4 students in a Masters' degree in natural sciences, 4 middle and high school teachers, and 2 experts in aquatic biodiversity surveys.

We collected specimens of three ordinary Odonata species of the Région Ile-de-France, in a pond located in the University Paris-Sud campus – Orsay, France (GPS coordinates: 48.702480, 2.171445). The first two species are *Libellula quadrimaculata* (Anisoptera, Libellulidae) and *Pyrrhosoma nymphula* (Zygoptera, Coenagrionidae); the specimens are both mature larvae (last larval stage before emergence). The third taxon cannot be identified at the species level because the specimen (female) is not mature enough (antepenultimate larval stage); the species is either *Anax imperator*, *Anax parthenope* or *Hemianax ephippiger* (Anisoptera, Aeshnidae). This however, follows the reality of the field: often, the larvae are not fully developed and it is not possible to identify at the species level.

Before using the identification key, a two-minute video was shown to the testers that explains how to navigate in the key and validate each step of the identification. For every identification, the same magnifier was used, an Olympus binocular magnifier with a magnification capacity of x7.5 to x64. The specimens were enlightened thanks to an optical fiber. 70% ethanol and two pairs of tweezers (hard and soft) were provided to the testers. The specimens were presented in a plate and stored in a 90% ethanol solution between tests.

The three specimens were presented to each tester, but the order of presentation was chosen at random, so the first, second and third specimen to be identified were not the same each time. We did not impose any time limitation for the identifications. No help were given to the testers, except for potential technical problems or difficulties with entomological or English terms (the testers did not speak English as their native language).

2.6 Analysis of test data

Every identification step (validation of the choice of a descriptor state) was recorded in the Xperience database. The entire identification path has the form of a series of numbers, each number corresponding to a descriptor or a descriptor state (Table 2); for example, *36:77; 46:100:101*; means that the user first selected the descriptor state #77 for the descriptor #36, then the descriptor states #100 and #101 for the descriptor #46. The database also records, at the end of each identification path, one or several numbers corresponding to the final species identification(s) made by the user.

3. Results

3.1 Structure and analysis of the identification keys

The three identification keys were checked and validated for their content by experts of the different orders.

The identification keys are available at those links:

Odonata: odonata-larvae-test.identificationkey.org

Ephemeroptera: ephemeroptera-larvae-paris.identificationkey.org

Plecoptera: plecoptera-larvae-paris.identificationkey.org

For the three keys, we have a total of 108 species: 39 Ephemeroptera, 59 Odonata, and 10 Plecoptera (Table 1, Table 2). There are a total of 117 descriptors and 253 descriptor states (Table 2). Each species, descriptor and descriptor state is illustrated by at least one image (picture or drawing).

The table 3 shows the numbers of items, descriptors and descriptor states for each key. When counting the number of descriptor states, we ignored the quantitative characters as they do not have a precise number of states.

The consistency of the knowledge base has been tested with the “Checkbase” option in Xper³: no items share the same description and all items are described.

The user interface shows the descriptors in order of discrimination power, the descriptor at the top being the most discriminant for the list of remaining species. After each descriptor/descriptor state validation, the list of possible species changes and so the order of the remaining descriptors (again, beginning with the most discriminant for the remaining species).

Table 4 shows the discrimination power of descriptors for the three identification keys.

3.2 Validation of the use of the key

The 72 identification paths recorded in the Xperience database are appropriate data to better understand how the Odonata larvae identification key is used by different types of publics. The analysis of the paths allows to know – among others things – the total number of steps taken in each identification path, the order in which the tester selected the different descriptors or descriptor states, or the steps where the tester “doubted” (i.e. selected several descriptor states for the same descriptor) or made mistakes (i.e. did not choose the descriptor state corresponding to the specimen being identified, did not selected the right species for the final identification).

The number of testers in each of the five groups being too small (between 2 and 7 people), we did not compare the results between groups, but we analyzed the data for the entire set of the 24 testers. There were a total of 465 steps (choices of a descriptor/descriptor state(s) couples) for 72 identification paths (3 paths per tester).

We looked at the number and percentage of correct and wrong final identifications; as well as the “correct included” final identifications, which are when the user validate the correct species but also some other species in their final identification. For the *Anax/Hemianax* complex, the identification was considered correct if only the three species were proposed by the user as the “correct identification”.

We looked at the numbers of doubts in the final identifications (2 out of 72), the doubts in the identification steps (14 out of 465) and the mistakes in the identification steps (45 out of 465). We calculated the corresponding percentages. The results are shown in Table 5.

For each taxon, the three most chosen descriptors at the first identification step were also the first three appearing in the list of descriptors in the interface. The detailed results are shown in Table 6.

Ten descriptors led to doubts; the details are shown in Table 7.

The Table 8 presents the descriptors that led to mistakes.

The graph in the Figure 1 allows to compare the percentage of mistakes made for the descriptors, and the rate of choice of each descriptor. We can see that some descriptors are often chosen by the users and at the same time, lead to a significant rate of mistakes (e.g. descriptor #36, sixth most chosen descriptor and more that 30% mistakes made).

3.3 Second version of the key

We used the results of the tests to implement a new version of the identification key in order to improve its quality and reliability.

[...]

4. Discussion

The classical and single-access keys are not interactive, cannot be adapted to different types of public, and present classical identification problems e.g. the impossibility to determine the taxon if some essential morphological characters are not visible. Most of the specimens in the taxa studied here, especially the larvae, are particularly prone to be damaged before the identification session because of their small size, their delicacy and/or the fact that they live in aquatic habitats and are not well preserved in dry environments. Moreover, collection techniques in the field for biodiversity inventories and surveys often involve strenuous contacts between the specimens and some elements of their habitats, e.g. stones, sand or wood, which often lead to the loss of the most delicate parts of the body (legs, cerci, antennas...).

The interactive digital identification keys developed here bypass these barriers and allow an accurate identification of all the specimens even if all the morphological structures are not apparent. For the majority of the taxa (items) of our six keys, identification at the species level is possible even if some data are not available or if the user is not sure of the character states. As all the characters used in this new key are visible to the naked eye or with a x20 lens, these keys can be used directly in the field, or indoors but without damaging the specimens by dissecting them.

Identification on the field is traditionally made with paper-printed keys, but the recent development of mobile devices allows the use of a portative version of digital keys. In this perspective, the project “Key of Nature” (www.keytonature.eu) aim to develop new, much easier and paper-free identification tools, for use within schools and universities across Europe and available on a variety of mobile platforms (laptops, smartphones). However, these keys are single access, e.g. MobileKey (Weber and Hagedorn 2010) or the iRecord Ladybirds mobile application, and do not support all the possibility of free access keys. Recently, a set of softwares and applications to transfer the information present in a local Xper database to an Android

application has been developed (Troudet 2012). For example, the Malaco-fr application provide an interactive way to identify the snails of France (Gargominy and Ripken 2011) on the field and without internet connection. The free access keys of mayflies, stoneflies and dragonflies & damselflies of the Région Ile-de-France described in this study will be soon transferred to such a mobile application. This system will make easier the work of naturalists and managers in the field. We also plan to integrate some of the keys in citizen science programs, for example the Steli (Suivi Temporel des Libellules), which is part of the Vigie-Nature platform (www.vigienature.mnhn.fr). The Steli is designed for non-beginner participants, i.e. people knowing some basics of entomology and natural sciences, and involves species identification in the field. Thus the digital identification key described in this study, based on morphological characters visible to the naked eye or with a small lens, is fully adapted. The Steli currently only focuses on Odonata adults; our goal is to extend this program to larval stages for this order.

In this study we tried a novel approach for the digital identification keys, by testing the use of this type of identification tool. We showed that it is possible to analyze some key parameters, for example the rates of mistakes and doubts in the final identifications and in all the identification steps during the use of the identification key. With a larger sample of testers in each category of publics, it would be possible to compare the results between the different categories. For example, do we observe a better rate of correct identifications when the user is a professional entomologist? Which descriptors lead to the biggest number of mistakes for the “beginners” categories? The analysis of this type of results allows to improve the identification keys, for example, add details, written complements or additional pictures for the problematic or difficult descriptors or descriptor states. [...]

5. Conclusion

[...]

Acknowledgements

We would like to warmly thank Guillaume Doucet (Odonata), Michel Brulin (Ephemeroptera), Alexandre Ruffoni and Jacques Le Doaré (Plecoptera) for their taxonomic expertise and the scientific validation of the identification keys. We also thank Michel Brulin, Christophe Brochard, the Perla database managers, and all the photographers from Flickr, for providing great illustrations. We also are very grateful to all the people who gave their time to test the key to Odonata: Camille Anceau, Lucie Bauret, Nicolas Bekkouche, Paola Campos, Benjamin Carbuccia, Laetitia Carrive, Donald Davesne, Claire David, Claudie Doums, Emilie Dutouillon, Felipe Espinosa, Romain Garrouste, Frédéric Legendre, Nicolas Mokuenko, Nathacha Ouvrié, Adrien Perrard, Patrick Perret, Lucile Rabeau, Camille Ramambason, Tony Robillard, Christine Rollard, Rodolphe Rougerie, Karen Salazar, Adeline Soulier-Perkins, Jean-Luc Suberchicot, Sébastien Turpin, Amir Yassin, and Julie Zalko. Finally, we thank the Région Ile-de-France for funding this project.

Références

- Barbier EB (2011) Wetlands as natural assets. *Hydrological Sciences Journal* 56: 1360–1373. doi: 10.1080/02626667.2011.629787
- Bauernfeind E, Soldán T (2012) The Mayflies of Europe (Ephemeroptera). *Apollo Boo*. 781 pp.
- Burguiere T, Causse F, Ung V, Vignes-Lebbe R (2013) IKey+: A new single-access key generation web service. *Systematic Biology* 62: 157–161. doi: 10.1093/sysbio/sys069
- Corvez A, Grand A (2014) Enabling comparisons of characters using an Xper2 based knowledge-base of fern morphology. *Phytotaxa* 183: 145–158. doi: 10.11646/phytotaxa.183.3.2
- Costello MJ, Bouchet P, Emblow CS, Legakis A (2006) European marine biodiversity inventory and taxonomic resources: State of the art and gaps in knowledge. *Marine Ecology Progress Series* 316: 257–268. doi: 10.3354/meps316257
- Dallwitz MJ, Paine TA, Zurcher EJ (2013) Principles of interactive keys. : 20. Available from: <http://delta-intkey.com/www/interactivekeys.pdf>.
- Doucet G (2010) Clé de détermination des exuvies des odonates de France. *Société Française d'Odonatologie*, 64 pp.
- Dudgeon D, Arthington AH, Gessner MO, Kawabata Z-I, Knowler DJ, Lévêque C, Naiman RJ, Prieur-Richard A-H, Soto D, Stiassny MLJ, Sullivan C a (2006) Freshwater biodiversity: importance, threats, status and conservation challenges. *Biological reviews of the Cambridge Philosophical Society* 81: 163–182. doi: 10.1017/S1464793105006950
- Edwards M, Morse DR (1995) The potential for computer-aided identification in biodiversity research. *Trends in Ecology & Evolution Reviews* 10: 153–158. doi: 10.1016/S0169-5347(00)89026-6
- Fontaine B, van Achterberg K, Alonso-Zarazaga MA, Araujo R, Asche M, Aspöck H, Aspöck U, Audisio P, Aukema B, Bailly N (2012) New species in the Old World: Europe as a frontier in biodiversity exploration, a test bed for 21st century taxonomy. *PLoS ONE* 7: e36881.
- Gargominy O, Ripken TEJ (2011) Une collection de référence pour la malacofaune terrestre de France. *MalaCo Hors Série*: 1–108.
- Grand D, Boudot J-P, Doucet G (2014) Cahier d'identification des Libellules de France, Belgique, Luxembourg et Suisse. *Biotope Ed. France*, 136 pp.
- Hagedorn G, Rambold G, Martellos S (2010) Types of identification keys. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems.* , 59–64.
- Hynes HBN (1977) A key to the adults and nymphs of the British stoneflies (Plecoptera). *Scientific Association FB (Ed)*. 90 pp.
- Jackson JK, Battle JM, White BP, Pilgrim EM, Stein ED, Miller PE, Sweeney BW (2014) Cryptic biodiversity in streams: a comparison of macroinvertebrate communities based on morphological and DNA barcode identifications. *Freshwater Science* 33: 312–324. doi: 10.1086/675225

- Kerner A, Debrenne F, Vignes-Lebbe R (2011) Cambrian archaeocyathan metazoans: Revision of morphological characters and standardization of genus descriptions to establish an online identification tool. *ZooKeys* 150: 381–395. doi: 10.3897/zookeys.150.1566
- Leal MC, Hilário A, Munro MHG, Blunt JW, Calado R (2016) Natural products discovery needs improved taxonomic and geographic information. *Nat. Prod. Rep.*: 747–750. doi: 10.1039/C5NP00130G
- Lebbe J (1991) *Représentation des concepts en biologie et en médecine: introduction à l'analyse des connaissances et à l'identification assistée par ordinateur*. Université Pierre et Marie Curie, Paris, France
- Macadam C, Bennett C (2010) *A pictorial guide to British Ephemeroptera*. AIDGAP. Field Studies Council Publications, 128 pp.
- Martin P, Schmelz RM, Dole-Olivier MJ (2015) Groundwater oligochaetes (Annelida, Clitellata) from the Mercantour National Park (France), with the descriptions of one new genus and two new stygobiont species. *Zoosystema* 37: 551–569. doi: 10.5252/z2015n4a2
- Nilsson A (1996a) *Aquatic Insects of North Europe: A Taxonomic Handbook. Volume 1: Ephemeroptera-Plecoptera-Heteroptera-Neuroptera-Megaloptera-Coleoptera-Trichoptera-Lepidoptera*. Nilsson A (Ed). Apollo Books, 274 pp.
- Nilsson A (1996b) *Aquatic Insects of North Europe: A Taxonomic Handbook. Volume 2: Odonata - Diptera*. Nilsson A (Ed). Books, Apollo, 440 pp.
- Pryce D, Macadam C, Brooks S (2007) *Guide to the British stonefly (Plecoptera) families: adults and larvae*. AIDGAP. Field Studies Council Publications.
- Riccamboni R, Mereu A, Boscarol C (2010) Keys to Nature: a test on the iPhone market. In: Nimis PL, Vignes-Lebbe R (Eds), *Tools for identifying biodiversity: progress and problems*. Trieste, Italy, 445–450.
- SCBD (2010) *What is the Problem? The Taxonomic Impediment*.
- SNPN (2012) *Bilan de l'inventaire des mares 2011*.
- Tachet H, Richoux P, Bournaud M, Usseglio-Polatera P (2010) *Invertébrés d'eau douce: systématique, biologie, écologie*. CNRS Editions, Paris, 608 pp.
- Troudet J (2012) *Des bases de connaissances en biodiversité aux outils de diffusion sur smartphones*.
- Troudet J, Grandcolas P, Blin A, Vignes-Lebbe R, Legendre F (2017) Taxonomic bias in biodiversity data and societal preferences. *Scientific Reports* 7: 9132. doi: 10.1038/s41598-017-09084-6
- Ung V, Dubus G, Zaragueta-Bagils R, Vignes-Lebbe R (2010) Xper2: introducing e-taxonomy. *Bioinformatics* 26: 703–704. doi: 10.1093/bioinformatics/btp715
- Vieira BNKM, Poff NL, Carlisle DM, Ii SRM, Koski ML, Kondratieff BC, Ii SRM (2006) *A Database of Lotic Invertebrate Traits for North America*. Director: 19. Available from: <http://pubs.water.usgs.gov/ds187>.
- Vignes-Lebbe R, Chesselet P, Diep Thi M-H (2016) Xper3: new tools for collaborating, training and transmitting knowledge on botanical phenotypes. In: Rakotoarisoa N, Blackmore S, Riera B (Eds), *Botanists of the twenty-first century: roles, challenges and opportunities*. Paris, 228–239.

Weber G, Hagedorn G (2010) Wiki keys on mobile devices. In: Nimis PL, Vignes-Lebbe R (Eds), Tools for identifying biodiversity: progress and problems. Trieste, Italy, 89–93.

Williams HC, Ormerod SJ, Bruford MW (2006) Molecular systematics and phylogeography of the cryptic species complex *Baetis rhodani* (Ephemeroptera, Baetidae). *Molecular Phylogenetics and Evolution* 40: 370–382. doi: 10.1016/j.ympev.2006.03.004

Vander Zanden MJ, Hansen GJA, Higgins SN, Kornis MS (2010) A pound of prevention, plus a pound of cure: Early detection and eradication of invasive species in the Laurentian Great Lakes. *Journal of Great Lakes Research* 36: 199–205. doi: 10.1016/j.jglr.2009.11.002

Tables and figures

Table 1: Taxonomic coverage of the study

Table 2: List of descriptors used in the study

Table 3: Number of items, descriptors and descriptor states for the three databases

Table 4: List of the discriminating power efficiency of descriptors

Table 5: Numbers and percentage of doubts in the final identifications and in the identification steps and mistakes in the identification steps

Table 6: Numbers and names of the different descriptors chosen at the first identification step

Table 7: List of descriptors leading to doubts in the identification steps

Table 8: List of descriptors leading to mistakes in the identification steps

Figure 1: General morphology of an aquatic insect larva

Figure 2: Comparison between the number of times each descriptor was chosen, and the percentage of mistakes for each descriptor

Table 1: Taxonomic coverage of the study

EPHEMEROPTERA Hyatt & Arms, 1891

Ameletidae McCafferty, 1991		<i>Metreletus balcanicus</i> Ulmer, 1920
Baetidae Leach, 1815	Baetinae Leach, 1815	<i>Baetis buceratus</i> Eaton, 1870
		<i>Baetis fuscatus</i> Linnaeus, 1761
		<i>Baetis liebenauae</i> Keffermüller, 1974
		<i>Baetis lutheri</i> Müller-Liebenau, 1967

		<i>Baetis rhodani</i> F.J. Pictet, 1843
		<i>Baetis scambus</i> Eaton, 1870
		<i>Baetis vernus</i> Curtis, 1834
		<i>Labiobaetis atrebatinus</i> Eaton, 1870
	Cloeoninae Newman, 1853	<i>Baetopus tenellus</i> Albarda, 1878
		<i>Centroptilum luteolum</i> O.F. Müller, 1776
		<i>Cloeon dipterum</i> Linnaeus, 1761
		<i>Cloeon simile</i> Eaton, 1870
		<i>Procloeon bifidum</i> Bengtsson, 1912
		<i>Procloeon pennulatum</i> Eaton, 1870
Caenidae Newman, 1853	Brachycercinae Lestage, 1924	<i>Brachycercus harrisellus</i> Curtis, 1834
		<i>Cercobrachys minutus</i> Tshernova, 1952
	Caeninae Newman, 1853	<i>Caenis horaria</i> Linnaeus, 1758
		<i>Caenis lactea</i> Burmeister, 1839
		<i>Caenis luctuosa</i> Burmeister, 1839
		<i>Caenis macrura</i> Stephens, 1836
		<i>Caenis pusilla</i> Navás, 1913
		<i>Caenis robusta</i> Eaton, 1884
Ephemerellidae Klapálek, 1909	Ephemerellinae Klapálek, 1909	<i>Serratella ignita</i> Poda, 1761
Ephemeridae Latreille, 1810		<i>Ephemera danica</i> Müller, 1764
		<i>Ephemera glaucops</i> F.J. Pictet, 1843
		<i>Ephemera lineata</i> Eaton, 1870
		<i>Ephemera vulgata</i> Linnaeus, 1758
Heptageniidae Needham, 1901	Ecdyonurinae Ulmer, 1920 (1905)	<i>Ecdyonurus dispar</i> Curtis, 1834
		<i>Ecdyonurus venosus</i> Fabricius, 1775
		<i>Electrogena ujhelyii</i> Sowa, 1981
	Heptageniinae Needham, 1901	<i>Heptagenia flava</i> Rostock, 1878
		<i>Heptagenia sulphurea</i> O.F. Müller, 1776
Leptophlebiidae Banks, 1900	Habrophlebiinae Kluge, 1994	<i>Habrophlebia lauta</i> Eaton, 1884
	Leptophlebiinae Banks, 1900	<i>Leptophlebia marginata</i> Linné, 1767
		<i>Paraleptophlebia submarginata</i> Stephens, 1836
Polymitarcyidae Banks, 1900		<i>Ephoron virgo</i> Olivier, 1791
Potamanthidae Albarda, 1888		<i>Potamanthus luteus</i> Linné, 1767
Prosopistomatidae Laméere, 1917		<i>Prosopistoma pennigerum</i> O.F. Müller, 1785

**ODONATA Fabricius,
1793**

Anisoptera Sélys, 1854	Aeshnidae Rambur, 1842	<i>Aeshna affinis</i> Vander Linden, 1823
		<i>Aeshna cyanea</i> Müller, 1764
		<i>Aeshna grandis</i> Linnaeus, 1758
		<i>Aeshna isoceles</i> Müller, 1767
		<i>Aeshna mixta</i> Latreille, 1805
		<i>Anax imperator</i> Leach, 1815
		<i>Anax parthenope</i> Sélys, 1839
		<i>Boyeria irene</i> Fonscolombe, 1838
		<i>Brachytron pratense</i> Müller, 1764
		<i>Hemianax ephippiger</i> Burmeister, 1839
	Cordulegastridae Hagen, 1875	<i>Cordulegaster boltonii</i> Donovan, 1807
	Corduliidae Sélys, 1871	<i>Cordulia aenea</i> Linnaeus, 1758
		<i>Epithea bimaculata</i> Charpentier, 1825
		<i>Oxygastra curtisii</i> Dale, 1834
		<i>Somatochlora flavomaculata</i> Vander Linden, 1825
		<i>Somatochlora metallica</i> Vander Linden, 1825
	Gomphidae Rambur, 1842	<i>Gomphus pulchellus</i> Sélys, 1840
		<i>Gomphus simillimus</i> Sélys, 1840
		<i>Gomphus vulgatissimus</i> Linnaeus, 1758
		<i>Onychogomphus forcipatus</i> Linnaeus, 1758
	Libellulidae Rambur, 1842	<i>Crocothemis erythraea</i> Brullé, 1832
		<i>Leucorrhinia caudalis</i> Charpentier, 1840
		<i>Leucorrhinia pectoralis</i> Charpentier, 1825
		<i>Libellula depressa</i> Linnaeus, 1758
		<i>Libellula fulva</i> Müller, 1764
		<i>Libellula quadrimaculata</i> Linnaeus, 1758
		<i>Orthetrum albistylum</i> Sélys, 1848
		<i>Orthetrum brunneum</i> Fonscolombe, 1837
		<i>Orthetrum cancellatum</i> Linnaeus, 1758
		<i>Orthetrum coerulescens</i> Fabricius, 1798
		<i>Sympetrum danae</i> Sulzer, 1776
		<i>Sympetrum flaveolum</i> Linnaeus, 1758
		<i>Sympetrum fonscolombii</i> Sélys, 1840
		<i>Sympetrum meridionale</i> Sélys, 1841
		<i>Sympetrum sanguineum</i> Brullé, 1832
		<i>Sympetrum striolatum</i> Charpentier, 1840
		<i>Sympetrum vulgatum</i> Linnaeus, 1758

Zygoptera Sélys, 1854	Calopterygidae Sélys, 1850	<i>Calopteryx splendens</i> Harris, 1782
		<i>Calopteryx virgo</i> Linnaeus, 1758
		<i>Chalcolestes viridis</i> Vander Linden, 1825
		<i>Lestes barbarus</i> Fabricius, 1798
		<i>Lestes dryas</i> Kirby, 1890
		<i>Lestes sponsa</i> Hanseman, 1823
		<i>Lestes virens</i> Charpentier, 1825
		<i>Sympecma fusca</i> Vander Linden, 1820
	Coenagrionidae Kirby, 1890	<i>Ceriagrion tenellum</i> de Villers, 1789
		<i>Coenagrion mercuriale</i> Charpentier, 1840
		<i>Coenagrion puella</i> Linnaeus, 1758
		<i>Coenagrion pulchellum</i> Vander Linden, 1825
		<i>Coenagrion scitulum</i> Rambur, 1842
		<i>Enallagma cyathigerum</i> Charpentier, 1840
		<i>Erythromma lindenii</i> Sélys, 1871
		<i>Erythromma najas</i> Hanseman, 1823
		<i>Erythromma viridulum</i> Charpentier, 1840
		<i>Ischnura elegans</i> Vander Linden, 1820
		<i>Ischnura pumilio</i> Charpentier, 1825
		<i>Pyrrhosoma nymphula</i> Charpentier, 1840
	Platycnemididae Jacobson & Bianchi, 1905	<i>Platycnemis acutipennis</i> Sélys, 1849
		<i>Platycnemis pennipes</i> Pallas, 1771

PLECOPTERA
Burmeister, 1839

Capniidae Klapálek, 1905	<i>Zwicknia bifrons</i> Newman, 1833	
Leuctridae Klapálek, 1905	<i>Leuctra geniculata</i> Stephens, 1836	
	<i>Leuctra nigra</i> Olivier, 1811	
Nemouridae Newman, 1853	<i>Nemoura cinerea</i> <i>cinerea</i> Retzius, 1783	
	<i>Nemurella</i> <i>pictetii</i> Klapálek, 1900	
Perlodidae Klapálek, 1912	<i>Isogenus</i> <i>nubecula</i> Newman, 1833	
	<i>Isoperla</i> <i>grammatica</i> Poda, 1761	

Taeniopterygidae Klapálek, 1905	<i>Brachyptera risi</i> Morton, 1896	
	<i>Taeniopteryx</i> <i>nebulosa</i> Linnaeus, 1758	
	<i>Taeniopteryx</i> <i>schoenemundi</i> Mertens, 1923	

Table 2: List of descriptors used in the study

EPHEMEROPTERA

Number	Name	Type
1	Dorsal shell	Categorical
2	Gills visible in dorsal view	Categorical
3	Number of gill pairs	Categorical
4	Dorso-ventral flattening	Categorical
5	Position of eyes	Categorical
6	Position of gills	Categorical
7	Anterior margin of the mesosternum	Categorical
8	A pale line in the middle of the pronotum	Categorical
9	Location of setae on the cerci	Categorical
10	Body length	Numerical
11	Antenna length	Categorical
12	Cerci length	Categorical
13	Sclerotized gill margin	Categorical
14	Mandible process	Categorical
15	Aspect of gills	Categorical
16	Gills 2 to 6: simple or double	Categorical
17	Gill end	Categorical
18	Gill shape	Categorical
19	Spines on the straight edge of gills	Categorical
20	Posterior angles of abdominal segments	Categorical
21	Relative size of the paracercus	Categorical
22	Apex of the maxilles	Categorical
23	Number of segments in the labial palp	Categorical
24	Last segment of the labial palp	Categorical
25	Length of setae on labial and maxillary palps	Categorical
26	Tarsal claw on the first pair of legs	Categorical
27	Length of the tarsal claw	Categorical
28	Teeth on the tarsal claw	Categorical
29	Expansion on the first antennal segment	Categorical

30	Width of glossae and paraglossae	Categorical
31	Shape of glossae	Categorical
32	One clearly defined dark strip around the middle of the cerci	Categorical
33	Number of dark rings between the last abdominal segment and the dark strip on the cerci	Categorical
34	Pronotum shape	Categorical
35	Lateral margins of the pronotum	Categorical
36	Protuberances on the pronotum	Categorical
37	Three protuberances on the head	Categorical
38	Dorso-apical extension in the tibia of the first pair of legs	Categorical
39	Aspect of the tibia	Categorical
40	Coloration patterns on abdominal tergites 7 to 9	Categorical
41	Two diffuse dark large spots on tergite 2	Categorical
42	Dark spot(s) on femur	Categorical
43	Teeth on mandibles	Categorical
44	Dark stripe(s) on the tarsus	Categorical
45	Length of bristles on the posterior margin of the femur	Categorical
46	Labrum shape	Categorical
47	Cerci coloration	Categorical
48	Dark antero-posterior line in the middle of the abdominal tergites	Categorical
49	Posterior edge of abdominal sternite 9	Categorical
50	Fringe of setae on the posterior margin of abdominal segments 7 and 8	Categorical
51	Spots on the abdominal tergites	Categorical
52	Irregular light marks on top of the head	Categorical

ODONATA

Number	Identifier for Xperience	Name	Type
1	0	Body length	Numerical
2	1	Length of the caudal gills	Numerical
3	2	End of the abdomen	Categorical
4	3	End of the caudal gills	Categorical
5	4	Length of the dark line bordering the caudal gills	Categorical
6	5	Medial trachea of the caudal gills	Categorical
7	6	Relative length of the dark lines bordering the caudal gills	Categorical
8	7	Horizontal node in the middle of the caudal gills	Categorical
9	8	Ratio length/width of the caudal gills	Categorical
10	9	Central caudal gill	Categorical
11	10	Relative length of the cerci and paraprocts	Categorical

12	11	On top of the anal pyramid, an expansion with a pointed end	Categorical
13	12	Relative length of the epiproct expansion and of the cerci	Categorical
14	13	Abdomen shape	Categorical
15	14	Antenna shape	Categorical
16	15	Length of the first antennal segment	Categorical
17	16	Type of mask	Categorical
18	17	Teeth on the labial palps	Categorical
19	18	Furrow at the base of the mask	Categorical
20	19	Pairs of legs 1 and 2	Categorical
21	20	Median slot at the end of the mask	Categorical
22	21	Shape of the prementum	Categorical
23	22	Protuberance(s) between the eyes	Categorical
24	23	Lateral expansions at the back of the head	Categorical
25	24	Deep cut in the labial palps	Categorical
26	25	Number of setae in the movable hook of the labial palps	Categorical
27	26	Number of long setae inside of the mask	Categorical
28	27	Ovipositor length	Categorical
29	28	The ovipositor reaches or exceeds the posterior margin of sternite 9	Categorical
30	29	Last tooth of the comb of the labial palp	Categorical
31	30	Setae on abdominal tergites 7, 8 and 9	Categorical
32	31	Setae on the posterior margin of abdominal tergites 8 and 9	Categorical
33	32	Row of spines on abdominal tergite 2	Categorical
34	33	Spines on thorax poststernite	Categorical
35	34	Black spot on abdominal sternites	Categorical
36	35	Dorsal spine on abdominal tergites 4 and/or 5	Categorical
37	36	Dorsal spine on abdominal tergite 7	Categorical
38	37	Dorsal spine on abdominal tergite 9	Categorical
39	38	Lateral spines on abdominal tergite 5	Categorical
40	39	Lateral spines on abdominal tergite 6	Categorical
41	40	Lateral spines on abdominal tergite 7	Categorical
42	41	Relative length of the lateral spines on abdominal tergite 8	Categorical
43	42	Length of the lateral spines on abdominal tergites 8 and 9	Categorical
44	43	Length of the lateral spines on abdominal tergite 9 in relation to the cerci	Categorical
45	44	Length of the lateral spines on abdominal tergite 9 in relation to tergite 10	Categorical
46	45	Dimensions of sternite 10	Categorical
47	46	Diameter of the eye	Categorical

48	47	Eye shape	Categorical
49	48	Lateral sides of the head	Categorical
50	49	Row of setae on the posterior margin of abdominal sternites 7 and/or 8	Categorical
51	50	Prominent part of the labial palp	Categorical
52	51	Ratio length/width of the mask	Categorical
53	52	Two light spots on the back of the head	Categorical

PLECOPTERA

Number	Name	Type
1	Length of glossae and paraglossae	Categorical
2	Distal segment of the maxillary palp	Categorical
3	Length of the tarsal segments	Categorical
4	Dorsal protrusions on the abdomen	Categorical
5	Length of the hind leg	Categorical
6	Direction of the wing pads	Categorical
7	End of the paraprocts	Categorical
8	Number of abdominal segments divided into upper and lower halves	Categorical
9	Chitinous processes on the basal segments of the antennae	Categorical
10	Eye position	Categorical
11	Body length	Numerical
12	Tarsal segments 1 and 3	Categorical

Table 3: Number of items, descriptors and descriptor states for the three databases

Database	Number of items (taxa)	Number of descriptors	Number of descriptor states	Min/Max/Average of number states per descriptor
Odonata	59	53	114	2/4/2.24
Ephemeroptera	39	52	115	2/4/2.27
Plecoptera	10	12	24	2/3/2.18

Table 4: List of the discriminating power efficiency of descriptors

Name of descriptor	Xper index
3. End of the abdomen	0,48
14. Abdomen shape	0,48
17. Type of mask	0,48
21. Median slot at the end of the mask	0,45
36. Dorsal spine on abdominal tergites 4 and/or 5	0,42
37. Dorsal spine on abdominal tergite 7	0,4
4. End of the caudal gills	0,34
15. Antenna shape	0,3
27. Number of long setae inside of the mask	0,28
24. Lateral expansions at the back of the head	0,24
6. Medial trachea of the caudal gills	0,21
8. Horizontal node in the middle of the caudal gills	0,21
19. Furrow at the base of the mask	0,2
38. Dorsal spine on abdominal tergite 9	0,18
5. Length of the dark line bordering the caudal gills	0,17
40. Lateral spines on abdominal tergite 6	0,17
41. Lateral spines on abdominal tergite 7	0,17
47. Diameter of the eye	0,17
10. Central caudal gill	0,16
22. Shape of the prementum	0,13
26. Number of setae in the movable hook of the labial palps	0,11
23. Protuberance(s) between the eyes	0,1
18. Teeth on the labial palps	0,09
9. Ratio length/width of the caudal gills	0,08
16. Length of the first antennal segment	0,07
20. Pairs of legs 1 and 2	0,06
25. Deep cut in the labial palps	0,06
49. Lateral sides of the head	0,04
11. Relative length of the cerci and paraprocts	0,02
12. On top of the anal pyramid, an expansion with a pointed end	0,02
28. Ovipositor length	0,02
33. Row of spines on abdominal stergite 2	0,02
39. Lateral spines on abdominal tergite 5	0,02
30. Last tooth of the comb of the labial palp	0,01
35. Black spot on abdominal sternites	0,01
48. Eye shape	0,01
1. Body length	0
2. Length of the caudal gills	0
7. Relative length of the dark lines bordering the caudal gills	0

13. Relative length of the epiproct expansion and of the cerci	0
29. The ovipositor reaches or exceeds the posterior margin of sternite 9	0
31. Setae on abdominal tergites 7, 8 and 9	0
32. Setae on the posterior margin of abdominal tergites 8 and 9	0
34. Spines on thorax poststernite	0
42. Relative length of the lateral spines on abdominal tergite 8	0
43. Length of the lateral spines on abdominal tergites 8 and 9	0
44. Length of the lateral spines on abdominal tergite 9 in relation to the cerci	0
45. Length of the lateral spines on abdominal tergite 9 in relation to tergite 10	0
46. Dimensions of sternite 10	0
50. Row of setae on the posterior margin of abdominal sternites 7 and/or 8	0
51. Prominent part of the labial palp	0
52. Ratio length/width of the mask	0
53. Two light spots on the back of the head	0

Table 5: Numbers and percentage of doubts in the final identifications and in the identification steps and mistakes in the identification steps

	Number	%
Final identification (correct/correct included/wrong)	32/2/38	44.44/2.78/52.78
Doubts in the final identification	2	2.78
Doubts in the identification steps	14	3.01
Mistakes in the identification steps	45	9.68

Table 6: Numbers and names of the different descriptors chosen at the first identification step

Most chosen descriptors	Times chosen	Position
Taxon A : <i>Libellula quadrimaculata</i>		
Type of mask	10	1
Abdomen shape	9	2
End of the abdomen	3	3
Dorsal spine on abdominal tergites 4 and/or 5	1	5
Diameter of the eye	1	12
Taxon B : <i>Pyrrhosoma nymphula</i>		
Abdomen shape	9	2
End of the abdomen	7	3
Type of mask	7	1
Median slot at the end of the mask	1	4
Taxon C : “<i>complexe Anax/Hemianax</i>”		

Type of mask	8	1
End of the abdomen	7	3
Abdomen shape	4	2
Median slot at the end of the mask	2	4
Number of long setae inside of the mask	1	8
Dorsal spine on abdominal tergites 4 and/or 5	1	5
Diameter of the eye	1	12

Table 7: List of descriptors leading to doubts in the identification steps

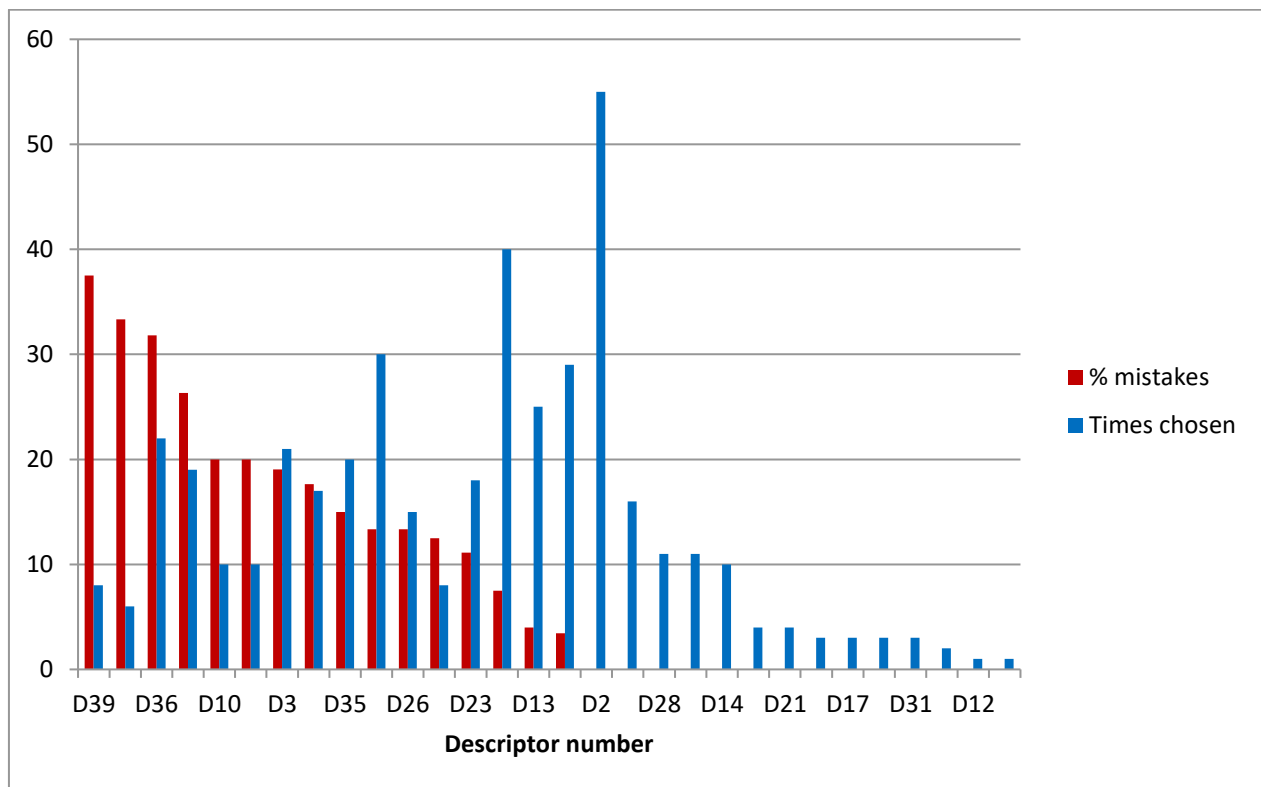
Number and name of the descriptor	Number of doubts
#37: Dorsal spine on abdominal tergite 9	3
#3: End of the caudal gills	2
#46: Diameter of the eye	2
#14: Antenna shape	1
#18: Furrow at the base of the mask	1
#19: Pairs of legs 1 and 2	1
#22: Protuberance(s) between the eyes	1
#23: Lateral expansions at the back of the head	1
#26: Number of long setae inside of the mask	1
#31: Setae on the posterior margin of abdominal tergites 8 and 9	1

Table 8: List of descriptors leading to mistakes in the identification steps

		% of errors	Number of times the descriptor was chosen
D39	Lateral spines on abdominal tergite 6	37,50	8
D48	Lateral sides of the head	33,33	6
D36	Dorsal spine on abdominal tergite 7	31,82	22
D18	Furrow at the base of the mask	26,32	19
D10	Relative length of the cerci and paraprocts	20,00	10
D19	Pairs of legs 1 and 2	20,00	10
D3	End of the caudal gills	19,05	21
D38	Lateral spines on abdominal tergite 5	17,65	17
D35	Dorsal spine on abdominal tergites 4 and/or 5	15,00	20
D16	Type of mask	13,33	30
D26	Number of long setae inside of the mask	13,33	15
D32	Row of spines on abdominal stergite 2	12,50	8
D23	Lateral expansions at the back of the head	11,11	18
D46	Diameter of the eye	7,50	40

D13	Abdomen shape	4,00	25
D20	Median slot at the end of the mask	3,45	29
D2	End of the abdomen	0	55
D37	Dorsal spine on abdominal tergite 9	0,00	16
D28	The ovipositor reaches or exceeds the posterior margin of sternite 9	0,00	11
D40	Lateral spines on abdominal tergite 7	0,00	11
D14	Antenna shape	0,00	10
D7	Horizontal node in the middle of the caudal gills	0,00	4
D21	Shape of the prementum	0,00	4
D11	On top of the anal pyramid, an expansion with a pointed end	0,00	3
D17	Teeth on the labial palps	0,00	3
D22	Protuberance(s) between the eyes	0,00	3
D31	Setae on the posterior margin of abdominal tergites 8 and 9	0,00	3
D9	Central caudal gill	0,00	2
D12	Relative length of the epiproct expansion and of the cerci	0,00	1
D25	Number of setae in the movable hook of the labial palps	0,00	1

Figure 2: Comparison between the number of times each descriptor was chosen, and the percentage of mistakes for each descriptor



Annexe 10 : Détails des 85 descripteurs de la base de données Xper³ du Spipoll

Numéro	Nom du descripteur	Partie du corps	Type	Nbre d'états	Nbre de taxons concernés	Nb gens qui l'ont choisi	Nbre de doutes	% de doutes
0	Période d'observation	N/A	Temporel	12	630	81	5	6.17
1	Quelle est l'allure générale de votre spécimen à identifier ?	Corps entier	Forme générale	6	630	1979	33	1.67
2	Forme des antennes de votre papillon	Antennes	Forme précise	2	231	60	0	0.00
3	Face du papillon observée	Ailes	Présence/absence	3	133	52	0	0.00
4	Type d'élytres	Ailes	Forme générale	4	166	216	6	2.78
5	Ecartement des élytres à leur extrémité	Ailes	Forme précise	2	139	19	2	10.53
6	Présence ou absence d'un "nez"	Tête	Présence/absence	2	139	56	2	3.57
7	Taille (Coléoptère, Hémiptère)	Corps entier	Taille	2	139	40	1	2.50
8	Forme des antennes	Antennes	Forme précise	2	166	159	0	0.00
9	Coloration des antennes	Antennes	Couleur	2	135	23	1	4.35
10	Présence ou absence de motifs sur les élytres ou demi-élytres	Ailes	Motifs	2	166	283	2	0.71
11	Couleurs principales des élytres ou demi-élytres	Ailes	Couleur	8	105	96	11	11.46
12	Motifs principaux sur les élytres et demi-élytres	Ailes	Motifs	3	105	43	1	2.33
13	Présence ou absence de motifs sur le thorax (Coléoptère, Hémiptère)	Thorax	Motifs	2	166	161	1	0.62
14	Couleur du thorax (Coléoptère, Hémiptère)	Thorax	Couleur	8	136	78	12	15.38
15	Motifs sur le thorax (Coléoptère, Hémiptère)	Thorax	Motifs	4	60	18	2	11.11
16	Type de coloration de l'abdomen	Abdomen	Couleur	2	196	686	21	3.06
17	Motifs et taches sur l'abdomen	Abdomen	Motifs	5	166	293	31	10.58
18	Couleur de l'extrémité de l'abdomen	Abdomen	Couleur	4	20	12	0	0.00
19	Couleur de la bande sur l'abdomen	Abdomen	Couleur	3	6	0	0	N/A
20	Couleur de l'abdomen uni	Abdomen	Couleur	5	81	103	11	10.68
21	Couleurs principales de l'abdomen	Abdomen	Couleur	6	166	232	38	16.38
22	Type de coloration du thorax	Thorax	Couleur	2	196	571	9	1.58
23	Couleur unie du thorax	Thorax	Couleur	6	144	216	19	8.80
24	Couleurs principales du thorax	Thorax	Couleur	6	134	92	7	7.61
25	Motifs sur le thorax	Thorax	Motifs	5	134	56	3	5.36
26	Taches claires sur le corps, les pattes, la face	Corps entier	Motifs	5	196	16	3	18.75
27	Type de coloration des ailes	Ailes	Couleur	4	196	384	28	7.29
28	Elargissement et pilosité de la patte postérieure	Pattes	Forme précise	4	196	345	40	11.59

29	Forme des yeux	Yeux	Forme précise	2	196	553	3	0.54
30	Couleurs et motifs des yeux	Yeux	Motifs	4	196	242	29	11.98
31	Forme de la cellule marginale	Ailes	Détail	2	196	28	2	7.14
32	Taille de votre spécimen	Corps entier	Taille	2	196	16	0	0.00
33	Longueur de la jonction thorax-abdomen	Thorax	Taille	2	196	151	0	0.00
34	Forme du bout de l'abdomen	Abdomen	Forme précise	5	196	40	2	5.00
35	Bout des antennes blanc	Antennes	Couleur	2	196	14	0	0.00
36	Longueur des antennes	Antennes	Taille	3	378	895	106	11.84
37	Forme du corps	Corps entier	Forme générale	2	201	440	10	2.27
38	Forme du corps (Coléoptère, Hémiptère)	Corps entier	Forme générale	2	166	194	1	0.52
39	Pilosité	Corps entier	Présence/absence	2	399	419	1	0.24
40	Présence ou absence de reflets métalliques	Corps entier	Couleur	2	593	115	3	2.61
41	Type de coloration du corps	Corps entier	Couleur	2	24	12	0	0.00
42	Couleur principale du corps	Corps entier	Couleur	8	24	43	5	11.63
43	Couleur principale des élytres ou demi-élytres	Ailes	Couleur	8	90	120	20	16.67
44	Présence ou absence d'une trompe, d'un rostre ou d'un "bec"	Tête	Présence/absence	3	217	102	5	4.90
45	Silhouette de votre papillon "de jour"	Corps entier	Forme générale	2	133	23	1	4.35
46	Forme des ailes	Ailes	Forme générale	4	133	33	3	9.09
47	Couleur principale du dessus des ailes	Ailes	Couleur	7	133	28	5	17.86
48	Motifs du dessus des ailes antérieures	Ailes	Motifs	5	40	0	0	N/A
49	Motifs sur le dessus des ailes orange	Ailes	Motifs	7	34	7	4	57.14
50	Motifs particuliers du dessus des ailes	Ailes	Motifs	3	15	2	0	0.00
51	Motifs sur le dessus des ailes marron à fauve	Ailes	Motifs	6	75	7	3	42.86
52	Couleur principale du dessous des ailes	Ailes	Couleur	6	133	32	5	15.63
53	Motifs du dessous des ailes	Ailes	Motifs	6	29	2	0	0.00
54	Couleurs et motifs sur le dessous des ailes	Ailes	Motifs	5	99	14	0	0.00
55	Couleur du dessous des ailes antérieures	Ailes	Couleur	2	32	6	0	0.00
56	Taches orange sur le dessous des ailes antérieures	Ailes	Couleur	2	26	5	0	0.00
57	Trait blanc sur le dessous des ailes postérieures	Ailes	Couleur	2	26	3	0	0.00
58	Bande blanche sur le dessous des ailes postérieures	Ailes	Couleur	2	26	4	0	0.00
59	Motifs sur le dessous des ailes postérieures	Ailes	Motifs	5	22	0	0	N/A
60	Taille de votre papillon "de nuit"	Corps entier	Taille	2	99	2	0	0.00

61	Silhouette de votre papillon "de nuit"	Corps entier	Forme générale	8	99	4	1	25.00
62	Aspect de la bordure des ailes	Ailes	Forme précise	2	96	0	0	N/A
63	Couleur principale du dessus des ailes antérieures	Ailes	Couleur	8	77	3	0	0.00
64	Motifs sur le dessus des ailes antérieures	Ailes	Motifs	3	77	1	0	0.00
65	Couleur des principaux motifs sur le dessus des ailes antérieures	Ailes	Couleur	6	72	1	0	0.00
66	Le dessus des ailes antérieures et postérieures sont-ils de couleurs semblables ?	Ailes	Couleur	2	99	2	0	0.00
67	Couleur principale du dessus des ailes postérieures	Ailes	Couleur	5	46	1	0	0.00
68	Positionnement géographique	N/A	Spatial	3	629	60	9	15.00
69	Massif montagneux	N/A	Spatial	3	574	3	0	0.00
70	Localisation Méditerranée	N/A	Spatial	2	601	3	0	0.00
71	Nombre de parties visibles du corps	Corps entier	Forme générale	2	20	92	0	0.00
72	Type de coloration du céphalothorax	Tête	Couleur	2	19	27	0	0.00
73	Motifs sur le céphalothorax	Tête	Motifs	3	18	9	1	11.11
74	Couleurs de l'abdomen	Abdomen	Couleur	2	19	43	1	2.33
75	Présence de motifs bien distincts sur l'abdomen	Abdomen	Motifs	2	19	49	1	2.04
76	Type de motifs sur l'abdomen	Abdomen	Motifs	5	18	14	3	21.43
77	Forme de l'abdomen	Abdomen	Forme générale	3	19	67	5	7.46
78	Type de coloration des pattes	Pattes	Couleur	2	19	4	0	0.00
79	Présence de très grandes épines sur les pattes	Pattes	Détail	2	19	14	0	0.00
80	Port et forme des pattes	Pattes	Détail	4	19	63	4	6.35
81	Présence ou absence d'ailes	Ailes	Présence/absence	2	24	50	0	0.00
82	Ailes translucides aux bordures colorées	Ailes	Couleur	2	231	55	0	0.00
83	Longueur des antennes de votre papillon	Antennes	Taille	2	231	5	0	0.00
84	Motifs sur le dessus des ailes blanches	Ailes	Motifs	5	15	2	0	0.00

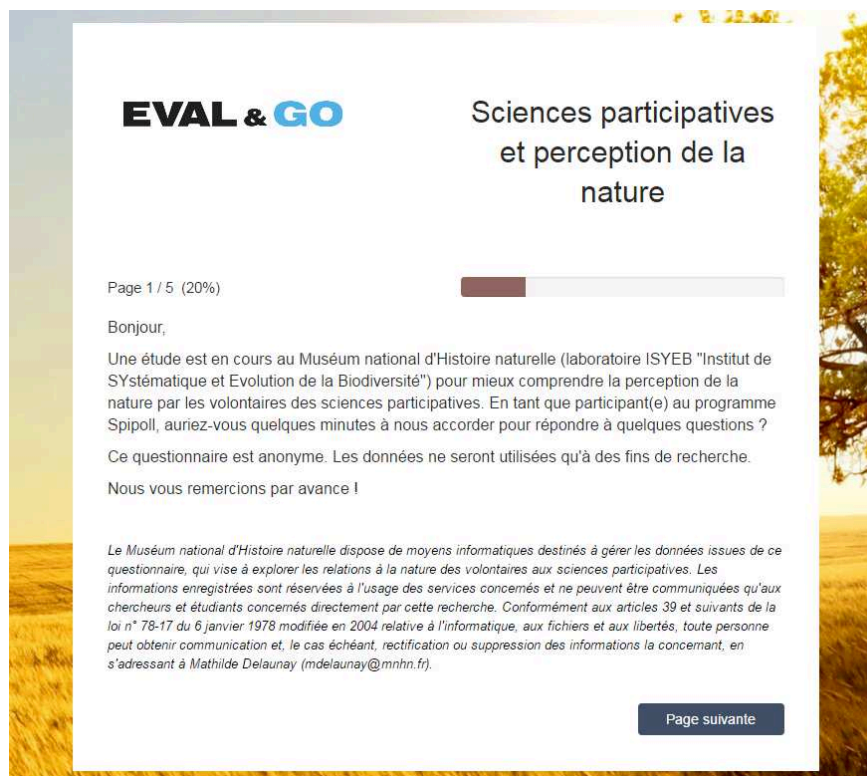
Annexe 11 : Liste des descripteurs de la base classés selon leur pouvoir discriminant d'après l'indice Xper (le plus discriminant en premier)

Descripteur	Indice XPER	Rang
Quelle est l'allure générale de votre spécimen à identifier ?	0,698	1
Couleurs et motifs sur le dessous des ailes	0,696	2
Motifs sur le dessus des ailes orange	0,686	3
Couleur principale du dessus des ailes postérieures	0,612	4
Couleur principale du dessus des ailes antérieures	0,600	5
Motifs du dessous des ailes	0,599	6
Motifs sur le dessous des ailes postérieures	0,589	7
Motifs sur le dessus des ailes blanches	0,581	8
Couleur principale du dessus des ailes	0,578	9
Motifs particuliers du dessus des ailes	0,543	10
Couleurs principales des élytres ou demi-élytres	0,514	11
Longueur des antennes	0,513	12
Motifs sur le dessus des ailes marron à fauve	0,496	13
Forme des yeux	0,492	14
Forme des antennes de votre papillon	0,487	15
Taches orange sur le dessous des ailes antérieures	0,474	16
Silhouette de votre papillon de nuit""	0,469	17
Couleur principale des élytres ou demi-élytres	0,454	18
Couleur principale du dessous des ailes	0,448	19
Couleur de l'extrémité de l'abdomen	0,437	20
Couleurs principales du thorax	0,436	21
Couleur du thorax (Coléoptère, Hémiptère)	0,405	22
Couleurs principales de l'abdomen	0,394	23
Elargissement et pilosité de la patte postérieure	0,388	24
Port et forme des pattes	0,380	25
Couleur des principaux motifs sur le dessus des ailes antérieures	0,363	26
Le dessus des ailes antérieures et postérieures sont-ils de couleurs semblables ?	0,350	27
Présence ou absence de motifs sur les élytres ou demi-élytres	0,339	28
Forme du corps (Coléoptère, Hémiptère)	0,333	29
Couleur de l'abdomen uni	0,328	30
Motifs principaux sur les élytres et demi-élytres	0,316	31
Forme de l'abdomen	0,310	32
Forme du corps	0,304	33
Bande blanche sur le dessous des ailes postérieures	0,302	34
Forme des antennes	0,294	35

Type de coloration des ailes	0,289	36
Type d'élytres	0,282	37
Présence ou absence d'ailes	0,279	38
Couleur du dessous des ailes antérieures	0,278	39
Motifs sur le thorax (Coléoptère, Hémiptère)	0,273	40
Motifs sur le dessus des ailes antérieures	0,259	41
Couleur unie du thorax	0,243	42
Motifs et taches sur l'abdomen	0,238	43
Présence ou absence de motifs sur le thorax (Coléoptère, Hémiptère)	0,232	44
Motifs du dessus des ailes antérieures	0,232	45
Forme des ailes	0,227	46
Type de motifs sur l'abdomen	0,216	47
Pilosité	0,213	48
Trait blanc sur le dessous des ailes postérieures	0,203	49
Motifs sur le thorax	0,202	50
Type de coloration de l'abdomen	0,181	51
Aspect de la bordure des ailes	0,175	52
Type de coloration du thorax	0,169	53
Longueur de la jonction thorax-abdomen	0,132	54
Silhouette de votre papillon de jour""	0,113	55
Couleur principale du corps	0,112	56
Coloration des antennes	0,107	57
Présence ou absence d'une trompe, d'un rostre ou d'un bec""	0,105	58
Couleurs de l'abdomen	0,105	59
Présence de très grandes épines sur les pattes	0,105	60
Nombre de parties visibles du corps	0,100	61
Taille de votre papillon de nuit""	0,095	62
Taille (Coléoptère, Hémiptère)	0,094	63
Ecartement des élytres à leur extrémité	0,092	64
Présence ou absence d'un nez""	0,083	65
Couleurs et motifs des yeux	0,081	66
Motifs sur le céphalothorax	0,072	67
Présence ou absence de reflets métalliques	0,065	68
Type de coloration du céphalothorax	0,035	69
Taches claires sur le corps, les pattes, la face	0,035	70
Ailes translucides aux bordures colorées	0,026	71
Type de coloration des pattes	0,023	72
Forme du bout de l'abdomen	0,018	73
Type de coloration du corps	0,018	74
Présence de motifs bien distincts sur l'abdomen	0,018	75
Longueur des antennes de votre papillon	0,017	76

Période d'observation	0,015	77
Forme de la cellule marginale	0,010	78
Bout des antennes blanc	0,010	79
Taille de votre spécimen	0,010	80
Positionnement géographique	0,008	81
Localisation Méditerranée	0,004	82
Massif montagneux	0,001	83
Face du papillon observée	0,000	84
Couleur de la bande sur l'abdomen	0,000	85

Annexe 13 : Questionnaire



EVAL & GO Sciences participatives
et perception de la
nature

Page 1 / 5 (20%)

Bonjour,

Une étude est en cours au Muséum national d'Histoire naturelle (laboratoire ISYEB "Institut de SYstématique et Evolution de la Biodiversité") pour mieux comprendre la perception de la nature par les volontaires des sciences participatives. En tant que participant(e) au programme Spipoll, auriez-vous quelques minutes à nous accorder pour répondre à quelques questions ?

Ce questionnaire est anonyme. Les données ne seront utilisées qu'à des fins de recherche.

Nous vous remercions par avance !

Le Muséum national d'Histoire naturelle dispose de moyens informatiques destinés à gérer les données issues de ce questionnaire, qui vise à explorer les relations à la nature des volontaires aux sciences participatives. Les informations enregistrées sont réservées à l'usage des services concernés et ne peuvent être communiquées qu'aux chercheurs et étudiants concernés directement par cette recherche. Conformément aux articles 39 et suivants de la loi n° 78-17 du 6 janvier 1978 modifiée en 2004 relative à l'informatique, aux fichiers et aux libertés, toute personne peut obtenir communication et, le cas échéant, rectification ou suppression des informations la concernant, en s'adressant à Mathilde Delaunay (mdelaunay@mnhn.fr).

Page suivante

Page 2 / 5 (40%)

VOTRE RAPPORT A LA NATURE ET VOTRE EXPERIENCE DES SCIENCES PARTICIPATIVES :

1. Où avez-vous passé l'essentiel de votre enfance ? *

- Hameau (moins de 200 habitants)
- Village (entre 200 et 2 000 habitants)
- Petite ville (entre 2 000 et 20 000 habitants)
- Ville moyenne (entre 20 000 et 200 000 habitants)
- Grande ville (plus de 200 000 habitants)

2. Actuellement, où vivez-vous ? *

- Hameau (moins de 200 habitants)
- Village (entre 200 et 2 000 habitants)
- Petite ville (entre 2 000 et 20 000 habitants)
- Ville moyenne (entre 20 000 et 200 000 habitants)
- Grande ville (plus de 200 000 habitants)

3. A quel point vous sentez-vous en phase avec la nature ? *

Je ne me sens pas du tout proche de la nature Je ne fais qu'un(e) avec la nature

1 2 3 4 5

4. Quels sont les programmes de sciences participatives auxquels vous participez ou avez participé ? *

- Spipoll
- Opération Papillons
- Opération Escargots
- Observatoire des bourdons
- Sauvages de ma rue
- Oiseaux des jardins et/ou BirdLab
- BioLit
- Autre(s)

5. Quelles motivations vous ont poussé(e) à débiter votre implication dans un (des) programme(s) de sciences participatives ? *

- Motivation 1 : Agir pour des valeurs qui vous paraissent importantes (humanisme, protection de l'environnement...)
- Motivation 2 : Prendre part à un projet scientifique (suivi d'un protocole, acquisition de données scientifiques, découverte des résultats...)
- Motivation 3 : Apprendre des choses (scientifiques, naturalistes, organisationnelles...) et/ou exercer des compétences (observer scientifiquement la nature, suivre un protocole précis...)
- Motivation 4 : Passer du temps avec vos proches (amis, famille...) en partageant une activité ensemble
- Motivation 5 : Passer du temps à l'extérieur, au contact de la nature
- Motivation 6 : Vous développer psychologiquement, vous sentir mieux dans votre vie
- Motivation 7 : Gagner des compétences et/ou de l'expérience utiles pour votre carrière (pour trouver un travail, être plus performant au travail...)
- Motivation 8 : Passer le temps et/ou vous changer les idées (vous éloigner de vos propres problèmes)
- Autre

6. **Classez les motivations que vous avez cochées à la question précédente par ordre d'importance (en commençant par la motivation qui a été la plus forte dans votre volonté de faire des sciences participatives) en indiquant leurs numéros** *

Exemple de réponse : "4, 7, 2, 1"

Page précédente

Page suivante

Page 3 / 5 (60%)



QUELQUES QUESTIONS SUR LE PROGRAMME SPIPOLL :

7. **Quelle(s) année(s) avez-vous participé au Spipoll (au moins une collection déposée sur le site) ?** *

2010

2011

2012

2013

2014

2015

2016

8. **Combien de collections avez-vous déposées sur le site web du Spipoll au total ?** *

1 collection correspond à 1 session d'observation (au moins 20 min) de Spipoll

9. Parmi les facteurs suivants, quels sont ceux qui... ?



	Vous ont poussé(e) à commencer à participer au Spipoll	Vous motivent pour continuer à participer au Spipoll	Ne sont pas une motivation pour le Spipoll
Prendre des photos	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Observer des insectes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Observer des plantes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Développer vos connaissances en entomologie	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Développer vos connaissances en botanique	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Faire partie d'une communauté de volontaires aux sciences participatives	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Utiliser une clé d'identification pour nommer les insectes	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

10. Quels ont été, pour vous, les bénéfices de votre participation au Spipoll ?



	Pas du tout			Beaucoup	
	1	2	3	4	5
J'ai agi pour des valeurs qui me paraissent importantes (humanisme, protection de l'environnement...)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai pris part à un projet scientifique	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai appris des choses et/ou exercé des compétences	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai passé du temps avec mes proches en partageant cette activité avec eux	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je me suis fait de nouveaux amis au sein de la communauté des Spipolliens	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai passé du temps dehors, au contact de la nature	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Je me suis développé(e) psychologiquement, je me sens mieux dans ma vie	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai gagné des compétences et/ou de l'expérience utiles pour ma carrière	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
J'ai passé le temps et/ou je me suis changé les idées	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

11. Les affirmations suivantes vous correspondent-elles, avant votre implication dans les sciences participatives, et après votre implication ? *

	Avant mon implication dans un programme de sciences participatives	Suite à mon implication dans un programme de sciences participatives	Jamais
Je suis attentif(ve) à la nature et à la biodiversité qui m'entourent	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je me documente sur les problématiques environnementales	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je m'implique dans des associations naturalistes ou de protection de la nature	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
J'assiste à des conférences dans le domaine de la biodiversité	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je m'implique politiquement dans des actions de protection de la nature (manifestations, groupes de parole politiques, sitting dans les ZAD...)	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Je cherche à passer plus de temps au contact de la nature	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

12. Quel(s) facteur(s) pourrai(en)t vous faire stopper votre participation au Spipoll ou réduire la fréquence de votre participation ? *

Le manque de temps

La difficulté à trouver des fleurs pour faire des sessions d'observation

La faible quantité d'insectes pollinisateurs présents lors de vos précédentes sessions d'observation

La difficulté à identifier les insectes

Des difficultés techniques sur le site web du Spipoll (site lent, trop souvent en maintenance...)

La difficulté à comprendre le protocole et/ou les informations demandées lors de l'enregistrement des sessions d'observation

Le fait de ne pas avoir de "bon" appareil photo

Le manque de retour auprès des Spipolliens sur les résultats scientifiques obtenus à partir des données récoltées

La difficulté d'obtenir de l'aide de la part des responsables du Spipoll

Autre (si vous le souhaitez, vous pourrez préciser votre réponse dans la dernière question du sondage)

Rien de tout cela

13. Connaissez-vous la clé d'identification "Insectes" associée au programme Spipoll ? *

Oui Non

14. A quelle fréquence utilisez-vous cette clé d'identification pour nommer les insectes présents sur vos photos du Spipoll ? *

15. Vous arrive-t-il d'utiliser la clé d'identification "Insectes" du Spipoll pour un usage personnel en dehors du Spipoll ? *

Oui Non

16. Etes-vous satisfait(e) de la clé d'identification "Insectes" du Spipoll ? *

	Très insatisfait(e)			Très satisfait(e)	
	1	2	3	4	5
Design, mise en page	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Facilité de prise en main	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Rapidité du site qui héberge la clé	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Clarté des textes et explications	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Qualité des illustrations (schémas, figures, photos)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
Utilité de l'outil pour le programme (par rapport à s'il n'existait pas)	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

17. N'hésitez pas à nous faire part des remarques que vous auriez sur cette clé d'identification, ainsi que des points à améliorer dans l'outil.

18. Des remarques, demandes ou critiques sur le programme Spipoll de manière générale ? Ici, n'hésitez pas à décrire votre expérience Spipoll, à détailler les problèmes que vous rencontrez sur le site web ou dans votre participation au programme.

Page précédente

Page suivante

Annexe 14 : Poster sur l'expérience des volontaires aux sciences participatives




WANNA HELP SCIENCE?

Motivations of citizen science volunteers

Mathilde DELAUNAY, Régine VIGNES-LEBBE, Romain NATTIER

Institut de Systématique, Evolution, Biodiversité – UMR 7205 CNRS-MNHN-UPMC-EPHE, Muséum national d'Histoire naturelle, Sorbonne Universités, 57 rue Cuvier, 75005 Paris, France



Projet soutenu par la région – Dispositif PICRI

INTRODUCTION

What is a **citizen scientist**? It is "a volunteer who collects and/or processes data as part of a scientific enquiry" [1], within the framework of standardized protocols [2]. In various scientific fields, since 10 years, the number of citizen science programs has greatly increased, especially in Europe and the US.

A better understanding of conservation volunteers' behavior can provide insight into the relationship between humans and nature. Since they are not paid for this work, what are their **motivations** to participate, their **expectations** towards the projects? Which **benefits** can they gain as citizen science volunteers?

In 1999 Gil Clary & Mark Snyder [3] proposed the "**Matching hypothesis**" that says people could become volunteers if the work matches personal needs; they could pursue their commitment if the benefits from their participation are in harmony with their initial motivations. Here we propose to test this hypothesis.

MATERIAL & METHODS

For this study we focused on the biggest **French citizen science program about pollination**. In the Spipoll program, volunteers take pictures of pollinator insects in the field, then transfer the data to a dedicated website [4].

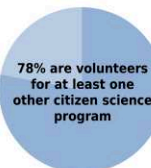
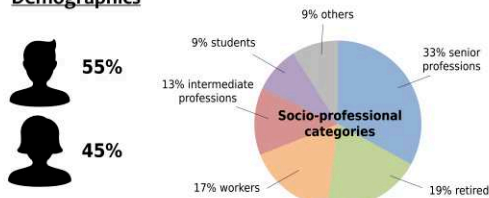


We created an **online survey** [5] with 23 questions for the Spipoll volunteers, about their profile (age, living environment, socio-professional category...) and their citizen science **prior experience**. We asked about their motivation, satisfaction and benefits gained as citizen science volunteers.

Question design was inspired by recent similar studies (e.g. [6]) and by Clary & Snyder's work. The survey was sent to the 1283 people engaged in the Spipoll from the beginning of the program in 2010. 271 of them responded, for a **21% response rate**.

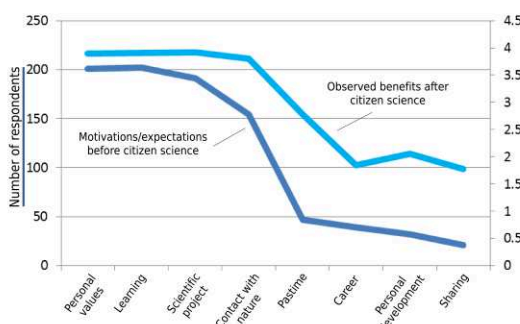
RESULTS & DISCUSSION

Demographics

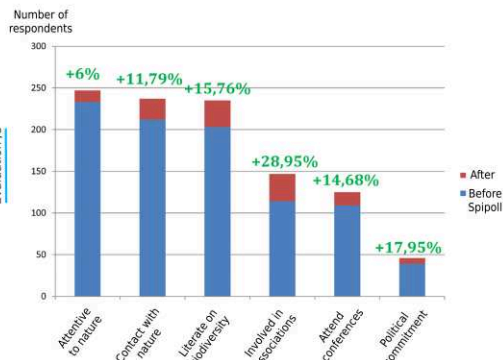


Respondents are predominantly **highly educated**. 1 out of 5 is retired, the mean age is **46 years old**. These demographics are similar to those in other environmental citizen science projects (e.g. [7]). Spipoll volunteers are **greatly aware of biodiversity issues**: most of them are involved in at least one other citizen science program in conservation.

Motivations and benefits



Our findings agree with Clary & Snyder's "Matching hypothesis": **benefits gained thanks to involvement in citizen science match the volunteers' initial motivations and expectations. The Spipoll program seems to bring satisfaction** to its participants, thus they are more likely to pursue their commitment. Moreover, volunteers gained **unexpected benefits**, for example for their career or their personal development.



Furthermore, participating in citizen science programs had interesting **aftereffects** in the life spheres of numerous volunteers, as observed in most of the other studies (e.g. [8]). After participating, **volunteers are more aware of nature**, ask more about biodiversity issues and attend more conferences on the subject. The biggest increase is for **involvement in associations** for biodiversity conservation.

CONCLUSION

Biodiversity-related citizen science programs contribute to raise awareness of nature in every type of public. Volunteers are more likely to be part of major actions for environment conservation. A good understanding of citizen science volunteers can provide insight into the relationship between humans and nature, but is also really helpful to **improve those programs** and win the loyalty of volunteers to **pursue citizen science development**.

REFERENCES

- [1] Silvertown J. (2009) A new dawn for citizen science. *Trends in Ecology and Evolution* 24(9): 467–471.
- [2] Irwin A. (1995) *Citizen science: a study of people, expertise and sustainable development*. Psychology Press.
- [3] Clary E. G. & Snyder M. (1999) The motivations to volunteer: theoretical and practical considerations. *Current Directions in Psychological Science* 8(5): 156–159.
- [4] <http://spipoll.org>
- [5] The survey was created with the Eval&Go platform: <http://evalandgo.com/>
- [6] Domrose M. C. & Johnson E. A. (2016) Why watch bees? Motivations of citizen science volunteers in the Great Pollinator Project. *Biological Conservation* – <http://dx.doi.org/10.1016/j.biocon.2016.08.020>
- [7] Brossard D., Lewenstein B., Bonney R. (2005) Scientific knowledge and attitude change: the impact of a citizen science project. *International Journal of Science Education* 27(9): 1099–1121.
- [8] Lewandowski E. J. & Oberbauer K. S. (2016) Buttery citizen scientists in the United States increase their engagement in conservation. *Biological Conservation* – <http://dx.doi.org/10.1016/j.biocon.2015.07.029>

RÉSUMÉ

Comment faciliter l'identification de l'entomofaune ? Construction, évaluation et amélioration de clés d'identification numériques

Les clés d'identification numériques représentent des outils efficaces pour identifier les êtres vivants, et sont également utilisées dans le cadre de programmes de sciences participatives. Accessible en ligne, la plateforme Xper³ permet de créer des bases de données de biodiversité et de générer des clés d'identification interactives.

Ce travail de thèse a testé diverses approches dans l'objectif de faciliter l'identification de l'entomofaune de France, en particulier de groupes d'insectes vivant dans des milieux menacés, ou ayant une importance écologique et patrimoniale notable : les coccinelles et trois ordres d'insectes aquatiques (odonates, éphémères et plécoptères), aux stades larvaires et adultes.

Nous avons créé pour ces groupes sept clés d'identification numériques fiables, intuitives et accessibles à tous les publics, disponibles à l'adresse web suivante : infosyslab.fr/?q=fr/node/113. Grâce au système Xperience permettant d'enregistrer le détail des identifications réalisées par les utilisateurs des clés Xper³, nous avons testé l'utilisation de notre clé des larves d'odonates et celle de la clé "Insectes" du programme de sciences participatives Spipoll sur les insectes pollinisateurs. Nous avons ainsi pu mieux comprendre comment ces clés numériques sont utilisées et quels problèmes sont rencontrés par leurs utilisateurs. Des pistes d'amélioration ont été proposées pour rendre les clés d'identification plus efficaces et faciles à prendre en main.

Le travail réalisé dans cette thèse a également montré que la précision des clés d'identification augmente avec l'intégration de critères écologiques en sus des caractères morpho-anatomiques. Cet ajout est particulièrement pertinent pour les jeunes larves d'insectes aquatiques, qui peuvent ne pas être identifiables à l'espèce avec les données morpho-anatomiques seules.

ABSTRACT

How to facilitate the identification of the entomofauna? Construction, evaluation and improvement of digital identification keys

Digital identification keys are effective tools for identifying plants and animals, and are also used in citizen science programs. Accessible online, the Xper³ platform makes it possible to create biodiversity databases and to generate interactive identification keys.

We investigated various approaches in order to facilitate the identification of French entomofauna, especially some insect taxa living in endangered environments, or having ecological and heritage significance: ladybirds and three orders of aquatic insects (dragonflies and damselflies, mayflies and stoneflies), at the larval and adult stages.

For all these groups, we created seven digital identification keys that are reliable, intuitive and accessible to everybody, available online at the following address: infosyslab.fr/?q=fr/node/113. Thanks to the Xperience system that can record details of the identifications made by the Xper³ key users, we surveyed the use of our key for odonate larvae and the "Insects" key of the citizen science program on pollination Spipoll. We were able to better understand how these digital keys are used and what problems are encountered by their users. Areas for improvement were proposed to make identification keys more effective and easier to use.

Our work also showed that the precision of digital identification keys increases when ecological data are integrated alongside morpho-anatomical characters. This addition is particularly relevant for young aquatic insect larvae, for which the identification at the species level can be impossible with morpho-anatomical data alone.