

Les référentiels en botanique

Notions de base permettant de comprendre comment sont constitués les référentiels et comment sont organisées ceux de Tela Botanica.

Daniel Mathieu, Joël Mathez, Valéry Malécot, David Mercier, Benoit Bock

Résumé : cet article constitue la synthèse et l'aboutissement des discussions qui ont eu lieu en 2017 au sein de Tela Botanica, notamment entre des membres de son conseil scientifique, pour définir la stratégie de constitution et d'organisation des index synonymiques et nomenclaturaux du site Internet de Tela Botanica utilisés pour agréger les connaissances sur la flore. Une part importante des discussions a porté sur la définition des termes utilisés en matière de taxinomie, de nomenclature et de référentiels de façon générale, pour conclure sur les principes pratiques qui seront mis en oeuvre dans les bases de données et les référentiels gérés par le réseau Tela Botanica.

Principes de taxinomie botanique

Les taxons

Les objets biologiques que nous avons à nommer sont bien particuliers, ils correspondent à des unités des classifications biologiques que l'on appelle des "**taxons**", c'est-à-dire à des ensembles discrets morphologiquement bien caractérisés, parmi lesquels les espèces satisfont à des critères biologiques classiques (ressemblance, interfécondité, sauf dans le cas des hybrides souvent stériles qui sont des taxons un peu particuliers définis par leur ascendance...). Nous verrons plus loin comment se définit plus précisément un taxon.

Les classifications biologiques

Les espèces (vivantes ou fossiles) sont regroupées en ensembles organisés en **système hiérarchisé** : selon différents critères, plusieurs espèces sont regroupées en genres, plusieurs genres en familles, etc. Dans ce système ont été définis un certain nombre de niveaux hiérarchiques, ou **rangs** dans lesquels la notion d'**espèce est une catégorie fondamentale**. En se limitant aux principaux, ces rangs sont :

espèce < genre < famille < ordre < classe < embranchement

En sens opposé, on reconnaît au sein de certaines espèces des taxons de rang inférieur à celui d'espèce, dont les principaux sont :

espèce > sous-espèce > variété > forme

Chacun de ces ensembles d'organismes, quel que soit son rang, constitue donc une unité de la classification ou **taxon**.

La construction d'une classification biologique revient à établir une liste hiérarchisée de taxons en choisissant pour chacun d'eux selon des critères variables :

- sa **délimitation** : sur quels critères décide-t-on qu'un individu appartient ou non à ce taxon ? ;
- son **rang** dans la hiérarchie biologique ;

- sa **position** dans la classification, c'est-à-dire le taxon de rang immédiatement supérieur dans lequel on décide de l'inclure.

Exemples de taxons : Fagaceae (niveau famille) > *Quercus* (niveau genre), > *Quercus ilex* (niveau espèce) > *Quercus ilex* subsp. *ballota* (sous-espèce).

Il importe de comprendre (et d'accepter !) que, selon les critères utilisés, les regroupements peuvent se faire de nombreuses façons différentes. Il n'existe donc pas une classification unique, ou unanimement admise par tous les botanistes... Nous verrons que ceci est lourd de conséquences sur la nomenclature et ses règles. Une forte tendance actuelle est de construire les classifications biologiques sur ce que l'on sait (ou suppose) de la phylogénie des êtres vivants : dans la mesure où il n'existe qu'un seul scénario évolutif, une seule phylogénie réelle, on pourrait espérer lui faire coïncider une classification biologique unique et donc universelle. À cet effet, on s'efforce de ne construire que des taxons "monophylétiques", c'est-à-dire rassemblant l'ensemble des individus issus d'un même ancêtre. Toutefois, il existe un très grand nombre de solutions possibles au découpage d'un groupe en taxons monophylétiques, parmi lesquels il faut choisir un nombre limité de rangs. Ce choix ne repose sur aucun critère objectif et reste donc largement arbitraire.

La classification APG (Angiosperm Phylogeny Group)

La **classification APG** (1998), ou classification phylogénétique, est une classification botanique des angiospermes établie selon les travaux de l'Angiosperm Phylogeny Group. Elle traduit, dans le domaine des angiospermes, les efforts faits en systématique pour que le système de classification des êtres vivants reflète au plus près la « réalité historique » des liens de parenté, ou phylogénie, entre les espèces, qu'elles soient actuelles ou éteintes.

Reflet d'un consensus sur les connaissances acquises lors de sa publication en 1998 (APG), cette classification a été révisée en 2003 (APG II) puis en 2009 (APG III) et la dernière en 2016 (APG IV) (les dernières modifications sont disponibles sur le [Angiosperm Phylogeny Website](#)). Cette classification phylogénétique, basée à l'origine sur les analyses de deux gènes chloroplastiques et d'un gène nucléaire, reflète aujourd'hui le consensus général sur les relations de parenté entre groupes d'angiospermes révélées puis confirmées par des centaines d'études dont les plus récentes sont basées sur des génomes entiers. Elle est aujourd'hui adoptée par la plupart des botanistes professionnels et reprise dans tous les grands herbiers du monde ainsi que de nombreuses bases de données de référence, telles que [The Plant List](#). Elle introduit des changements notables avec les classifications antérieures, fondées sur les ressemblances morphologiques. Par exemple, l'ancienne famille des Liliaceae est maintenant éclatée en une dizaine de familles.

Taxinomie, nomenclature et systématique

La discipline scientifique concernée par ces opérations est la **taxinomie**. Celle qui concerne la dénomination des taxons relève de la **nomenclature**. Elle n'intervient en effet qu'en aval des opérations taxinomiques : l'attribution d'un nom à un taxon doit obligatoirement obéir aux règles de nomenclature régies par le *Code international de nomenclature pour les algues, les champignons et les plantes*, mais ces règles ne peuvent être appliquées que si au préalable on a choisi sa délimitation, son rang et sa position. Or ces choix sont d'essence scientifique, c'est-à-dire appuyés sur une argumentation par essence contestable. Toutefois,

taxinomie et **nomenclature** entretiennent des relations évidentes ; on peut considérer que la **systématique** est la discipline qui les associe.

Résumé des règles de nomenclature botanique

Le Code international de nomenclature pour les algues, les champignons et les plantes (CIN / Melbourne juillet 2011)

Il s'agit d'un ensemble de règles et recommandations qui définissent le choix ou l'élaboration des noms scientifiques attribués aux plantes, et autres organismes « traditionnellement considérés comme plantes » bien que n'en étant pas tous selon les critères actuels, comme les champignons et les cyanobactéries. Il était appelé *Code International de Nomenclature Botanique* (CINB) jusqu'au 31 décembre 2011, son nom ayant été modifié par le XVIII^e Congrès international de botanique à Melbourne en juillet 2011 pour devenir le Code international de nomenclature pour les algues, les champignons et les plantes (CIN / Melbourne juillet 2011). Le dernier congrès s'est déroulé en juillet 2017 à Shenzhen en Chine.

Nom binomial, trinomial et combinaison nomenclaturale

Un **nom binomial** est la combinaison de deux mots (un nom de genre suivi d'une épithète spécifique) servant à désigner un **taxon** au rang d'espèce. Formalisée par Linné au cours du XVIII^e siècle, la nomenclature binominale constitue le « système linnéen ». Exemple : *Quercus ilex* pour désigner le Chêne vert.

Au rang de genre et au dessus (famille, embranchement, ...) le nom d'un taxon ne comprend qu'un seul mot : exemple *Quercus* pour le genre Chêne, ou *Fagaceae* pour la famille des Fagacées.

Entre le rang de genre et celui d'espèce (sous-genre, section, sous-section, série, ...) les noms sont des combinaisons dites **infragénériques**. Elles comportent deux ou trois mots (nom de genre suivi d'une épithète précédée de l'indication de rang). Exemple : *Quercus* subgen. *Quercus*, ou *Quercus* sect. *Lobatae*.

En dessous du rang d'espèce (sous-espèce, variété, forme, etc.) les noms sont des combinaisons dites **infraspécifiques** et comportent deux épithètes plus le nom du genre. Elles sont dites **trinominales**. Exemple : *Quercus ilex* subsp. *ballota*, ou *Quercus ilex* var. *angustifolia*.

Le Code ne permettant pas d'attribuer plus de deux épithètes au nom d'un taxon infraspécifique, ce nom ne reflète pas complètement la position du taxon dans la hiérarchie : ainsi, il ne donne aucune indication sur la sous-espèce à laquelle est rattachée une variété ou une forme. Cette information ne peut être exprimée qu'en désignant explicitement le nom du taxon de rang supérieur.

D'un point de vue typographique, les noms latins des taxons de rang inférieur ou égal au genre sont écrits en *italique* ; les noms d'auteur ainsi que l'abréviation désignant un rang infraspécifique ou infragénérique (var., subsp., f. ; subgen., sect.) sont écrits en caractères droits (et à l'inverse dans un texte écrit en italique). Dans un nom d'espèce, seul le nom du genre débute toujours par une majuscule.

Exemple : *Rubus serpens* subsp. *puripulvis* Sudre

Autorités

Chaque fois que la rigueur devient nécessaire, on doit faire suivre le nom d'une indication bibliographique sommaire, souvent restreinte à l'autorité, c'est-à-dire à la **citation de l'auteur** de la description originale ou de la combinaison, sous forme abrégée.

Exemple : *Quercus ilex* L. Le nom d'auteur Linné abrégé en " L." est noté en caractères droits.

Lorsqu'il s'agit d'un nom d'un nouveau taxon, la citation est simple : le nom est suivi du nom de l'auteur qui le publie (ou des noms des auteurs s'il sont plusieurs). On utilise très souvent une abréviation plus ou moins consacrée par l'usage (Cf. *Authors of plant names* de R.K. Brummitt et C.E. Powell <http://www.ipni.org/ipni/authorsearchpage.do>).

De façon facultative, il est possible d'indiquer, si c'est le cas, le nom du créateur du nom s'il n'est pas celui qui l'a validement¹ publié : le nom d'auteur est alors précédé du nom du créateur suivi de "ex".

Exemple : *Carex bipartita* Bellardi ex All. : l'auteur est Allioni, abrégé en All., et le créateur est Bellardi.

De façon facultative également, il est possible d'indiquer, si la publication a lieu sous l'égide d'un ou plusieurs coordinateur(s) principal(-aux) qui n'est (ne sont) pas l'auteur de la partie concernant le taxon, le nom du coordinateur de la publication : le nom d'auteur est alors suivi du nom du coordinateur précédé de "in".

Exemple : *Rubus* subgen. *Idaeobatus* Focke in Aschers. & Graebn.

Lorsqu'un systématicien estime que le genre ou le rang choisi n'est pas le meilleur pour un taxon déjà publié, il peut décider de transférer ce taxon dans un autre genre (par exemple suite à la création d'un nouveau genre) ou à un autre rang. Dans ce cas, le nom de l'auteur de la combinaison princeps (basionyme**) demeure, mais il est placé entre parenthèses.**

Exemple : l'espèce *Anagallis arvensis* L. a été transférée dans le genre *Lysimachia* par Ulrika Manns & Arne A. Anderberg et doit se nommer *Lysimachia arvensis* (L.) U.Manns & Anderb. Dans ce cas, le basionyme de la combinaison nouvelle est *Anagallis arvensis* L.

Homonymes : l'autorité permet également de distinguer deux noms identiques homonymiques se référant à deux taxons distincts, comme *Daucus carota* L. et *Daucus carota* Mill. par exemple.

Daucus carota L. [1753] est un nom affecté par Linné en 1753 à la Carotte sauvage de nos contrées, alors que *Daucus carota* Mill. 1768 est le nom (synonyme) affecté en 1758 par Philip Miller à la Carotte que nous consommons. Ne pas confondre les deux !

¹ Voir plus loin la définition de cette expression en nomenclature.

Date de publication

Le nom d'auteur est parfois suivi de la **date de publication effective**² du livre ou de la revue dans lequel le taxon a été nommé pour la première fois sous ce nom. Cette date est indispensable pour appliquer le principe de priorité lorsqu'on estime que le même taxon a été désigné successivement par des noms différents ou lorsqu'un auteur a publié le même nom pour désigner différents taxons à différentes dates.

Exemple : *Anagallis arvensis* L. [1753] a été dénommé par Linné en 1753.

Source bibliographique

Afin de faciliter la recherche d'informations et faire des vérifications concernant un nom, on peut faire suivre celui-ci de la **référence bibliographique** plus précise de la publication. On y trouve en effet la description du taxon ou la liste des caractères le distinguant des autres taxons proches (diagnose).

Exemple : *Anagallis arvensis* L. [1753, Sp. Pl., 1 : 148] a été publié par Linné dans la publication *Species Plantarum*, volume 1, page 148. Voir :

<http://biodiversitylibrary.org/page/358167>

La notion de type nomenclatural

Le Code de nomenclature impose de désigner un échantillon de collection (ou parfois une illustration), comme "étalon" du nouveau nom d'un taxon : c'est cet échantillon que le CIN désigne sous le nom de "**type nomenclatural**", c'est-à-dire de type du nom donné au taxon. Le Code précise qu'un type nomenclatural est l'élément auquel le nom du taxon est attaché de façon permanente, qu'il s'agisse du nom correct du taxon ou d'un synonyme du nom correct (voir plus loin).

Il peut s'agir du spécimen unique ayant servi à la dénomination du taxon ou désigné explicitement parmi d'autres comme type par l'auteur du nom (**holotype**), ou d'un spécimen désigné après la description parmi les récoltes (appelés **syntypes** avant désignation d'un type parmi eux) ayant servi à sa description (**lectotype**). Plus rarement, si le ou les spécimens ayant servi à la description ont disparu, ou rarement, si le ou les spécimens possibles sont en contradiction avec un usage unanimement accepté, un autre spécimen représentant le taxon tel qu'il est interprété est désigné comme type (**néotype**). Encore plus rarement, dans le cas d'un spécimen type (holotype, lectotype ou néotype) incomplet et d'appartenance taxinomique impossible à établir, on peut désigner, en plus de celui-ci, un spécimen présentant les organes nécessaires à une identification certaine (**épitype**).

Le type nomenclatural n'est pas nécessairement l'élément le plus typique ou le plus représentatif d'un taxon. Le recours à l'examen du type s'avère cependant indispensable en bien des circonstances au cours des travaux de systématique : par exemple lorsque l'on est amené à préciser ou modifier la délimitation du taxon à l'aide de critères nouveaux (palynologiques, moléculaires...) et de façon générale lorsqu'on s'interroge sur ce que recouvre réellement le nom attribué à un taxon.

² Voir plus loin la définition de ce terme.

Désignation, nom validement publié

Pour qu'une **désignation** soit reconnue comme nom au sens du CIN , elle doit être "**validement publiée**", c'est-à-dire satisfaire à des conditions formelles:

1. La publication doit être "**effective**" : il s'agit notamment de garantir que les informations soient réellement accessibles aux scientifiques.
2. La publication doit par ailleurs être "**valide**", c'est-à-dire conforme à une série de **règles** du CIN

Par exemple, lorsqu'il s'agit d'un nom d'une nouvelle espèce, la publication doit notamment satisfaire les critères suivants :

- le nom doit répondre à des **critères d'acceptabilité** (forme latine, binominale pour les espèces, trinominale pour les taxons de rang inférieur, épithètes tels que "genuinus", "typicus" non acceptées, épithète de l'espèce devant être différente du nom de genre...).
- le taxon à nommer doit être explicitement défini par sa **délimitation**, son **rang** et sa **position** - que nous abrégeons par "**DRP**" - dans la classification, trois critères découlant des choix taxinomiques préalables,
- délimitation : présence d'une **description** ou d'une **diagnose** différentielle et choix des descripteurs acceptables ;
- le nouveau nom doit être **formellement exprimé**, et pas seulement de façon allusive, provisoire ou conditionnelle ;
- un **type doit être explicitement désigné** dans les publications distribuées après le 1er janvier 1958 ;
- un nom de rang inférieur au genre (série, espèce, variété, etc.) n'est validement publié que si le nom du genre auquel on l'attribue est également validement publié ;

La date d'un nom est celle de sa publication valide et donc pas nécessairement celle indiquée sur la revue ou l'ouvrage.

Nom légitime, nom illégitime

Un nom doit être légitime pour pouvoir être éventuellement utilisable comme nom correct, c'est à dire satisfaire à certains critères définis par le Code de nomenclature.

Et pour qu'un nom soit légitime, il faut qu'il réponde aux caractères suivants :

1. **Le nom doit être dépourvu d'homonyme antérieurement publié.** Ex. *Poa alpina* Pall. ex Roem. et Schult. publié en 1817 est un nom illégitime, car il existe *Poa alpina* L. publié en 1753. Il est à noter que l'indication de rang n'intervient pas dans la notion d'homonymie (par ex. : *Rubus hedycarpus* subsp. *pubescens* Weihe ex Focke et *Rubus hedycarpus* var. *pubescens* Sudre sont homonymes). Des noms à orthographes très proches (ex. *Rubus sciaphilus* Lange et *Rubus sciophilus* Lefèvre et P.J. Müll.) peuvent être également considérés comme homonymes.
2. Le taxon auquel il se rapporte, et tel que défini par l'auteur du nom au moment de la publication valide de ce dernier, **ne doit pas inclure le type d'un taxon déjà pourvu d'un nom légitime de même rang pourvu d'une épithète finale différente** (rang inférieur au genre) ou pourvu d'un nom différent (rang supérieur ou égal au genre : ex. 1), à moins que cela ne permette d'éviter une homonymie ou une publication invalide (ex. 2), ou bien qu'il s'agisse d'une recombinaison (ex. 3). - Exemple 1 : *Pinus excelsa* Lam. publié en 1779 inclut le type de *Pinus abies* L. (publié en 1753), et est donc illégitime. La recombinaison *Picea excelsa* (Lam.) Link (publiée en 1842), est également illégitime, puisque l'espèce est déjà pourvue d'un nom légitime de même rang avec une épithète différente et disponible sous ce genre *Picea* : l'épithète

abies de *Pinus abies* L. Le nom légitime au rang d'espèce sous le genre *Picea* est donc *Picea abies* (L.) H. Karst, publié en 1881. Cet exemple montre que ce n'est donc pas toujours le nom le plus ancien qui est légitime ! - Exemple 2 : *Persicaria mitis* Delarbre est basé sur le même type que *Polygonum persicaria* L., mais est un nom légitime, car les tautonymes (ici, *Persicaria persicaria*) ne sont pas autorisés par le Code. Il s'agit en fait de la publication d'un nomen novum (voir ci-dessous). - Exemple 3 : *Chloris radiata* (L.) Sw. publié en 1788 est une combinaison nouvelle fondée sur le basionyme *Agrostis radiata* L. (1759), et la mention de *Andropogon fasciculatus* L. (publié en 1753) dans le protologue n'engendre pas l'illégitimité du nom.

3. Dans de rares cas, malgré une illégitimité par l'application des deux règles précédentes, un statut spécial de conservation du nom peut être obtenu, dans un but de stabilité de la nomenclature (*nomen conservandum*). Ou bien, à l'inverse, un nom en principe légitime peut être rejeté pour diverses raisons (*nomen rejiciendum*).
4. Dans le cas de **genres hybrides**, le nom doit être constitué de la fusion d'une partie ou totalité du nom des genres parentaux. Ex. : x *Sasinaria* Demoly (1995) est le nom légitime de *Arundinaria* x *Sasa*, et remplace *Sasaella* Makino (1929) qui devient illégitime suite à la découverte qu'il s'agit d'un genre hybride.

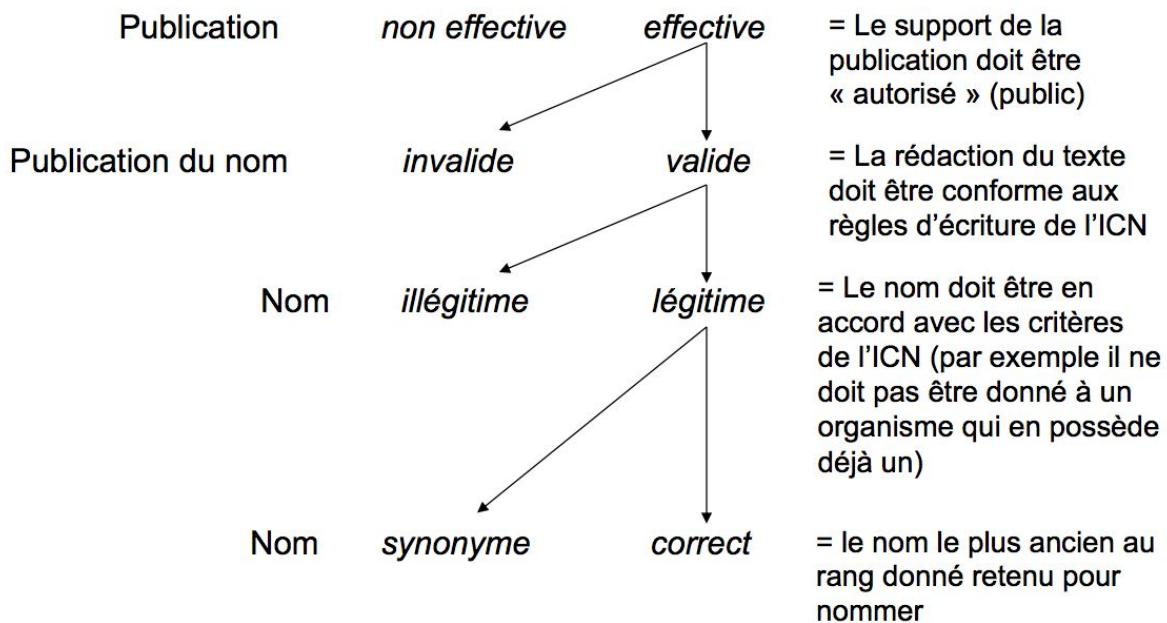
Nom correct, synonyme

Chaque taxon tel que compris par un ou plusieurs auteur(s) comme défini par une délimitation, un rang et une position systématique (DRP), reçoit **un nom et un seul, appelé nom correct**. Lorsqu'un nom correct est défini, les autres noms dont le type appartient au taxon sont relégués au statut de **synonyme**.

Ce nom doit être légitime et répondre aux exigences suivantes :

- le type du nom correct (qui est un spécimen, rarement une illustration) doit appartenir au taxon tel que délimité ;
- le rang du nom correct doit être évidemment en accord avec celui du taxon (par exemple un nom au rang d'espèce pour une espèce) ;
- si le taxon est de rang inférieur au genre, le nom correct est une combinaison qui doit être évidemment en accord avec la systématique du taxon (par exemple, si le taxon appartient au genre *Veronica*, la combinaison doit être sous le genre *Veronica*, et pas un autre).

Parmi les noms légitimes disponibles (ex.) ou potentiels (note), le nom correct (nom de rang égal ou supérieur au genre) ou l'épithète finale du nom correct (combinaison de rang inférieur au genre) doit être le plus anciennement publié au rang du taxon concerné. - Exemple : de nombreux noms au rang d'espèce possèdent un spécimen type appartenant à *Picea abies* (L.) H.Karst. Parmi ceux-ci, *Pinus abies* L. est le plus ancien nom légitime, et l'épithète *abies* a donc priorité. Bien que le nom *Picea abies* (L.) H.Karst ne date que de 1882, il a notamment priorité sur le nom légitime *Picea rubra* A.Dietr. publié en 1824. - Note : il peut arriver que parmi les noms disponibles, aucun ne convienne, et qu'il soit nécessaire de créer un nouveau nom. Ce nouveau nom doit alors également respecter les règles de priorité telles que définies ici.



Relations entre les “objets” nomenclaturaux (Valéry Malécot)

Basionyme

Nom légitime à partir duquel un nom nouveau est donné, et dont l'épithète finale est identique. Par définition, ces noms se basent sur le même spécimen type. Il peut s'agir d'une combinaison nouvelle (ex. 1) ou seulement d'un changement de rang (ex. 2).

Exemple 1 : *Pinus abies* L. est le basionyme de *Picea abies* (L.) H.Karst. -
 Exemple 2 : *Rubus serpens* subsp. *puripulvis* Sudre est le basionyme de *Rubus serpens* var. *puripulvis* (Sudre) Sudre.

Ou plus rarement, l'épithète d'un nom de rang intermédiaire entre genre et espèce est élevée au rang de nom de genre, ou inversement. Ex. *Cerasus* Mill. est le basionyme de *Prunus* subgen. *Cerasus* (Mill.) Focke.

Synonyme remplacé

Nom légitime ou illégitime pour lequel un nom nouveau pourvu d'une épithète finale nouvelle est donné (ex. 1) ou bien nom illégitime pour lequel un nom nouveau pourvu de la même épithète finale est donné (ex. 2). Par définition, ces noms se basent sur le même spécimen type.

Exemple 1 : *Polygonum persicaria* L. est le synonyme remplacé de *Persicaria mitis* Delarbre.
 Exemple 2 : Le nom illégitime *Rubus robustus* P.J. Müll., est le synonyme remplacé de *Rubus pubescens* Weihe subvar. *robustus* P.J. Müll. ex Nyman.

Diagnose

Une diagnose, en biologie, est une description scientifique, concise, permettant d'isoler ou de différencier un taxon (espèce, genre, famille, etc.), des taxons qui lui sont proches.

En botanique, depuis 1935 et jusqu'en 2011, une diagnose devait être publiée en latin pour être valide. Le *Code international de nomenclature botanique*, amendé lors du XVIII^e

Congrès international de botanique qui s'est tenu à Melbourne en juillet 2011, autorise aussi, depuis le 1^{er} janvier 2012, la publication d'une diagnose en anglais.

Protologue

Tous les éléments associés à la publication valide du nom, tels que la description, diagnose, illustration, références, synonymie, données géographiques, mention de spécimens, discussion et commentaires.

Locus typicus

Localité de récolte du spécimen type.

Synonymie

Les synonymes des noms corrects sont de 2 types différents :

1. **les synonymes nomenclaturaux** : ce sont tous les noms qui ont été attribués au même type nomenclatural. Ils ne peuvent faire l'objet de discussion car ils se réfèrent au même objet (il sont **homotypiques**). C'est le cas par exemple lorsqu'on affecte le type d'une espèce à un genre différent ou que l'on change le rang du taxon qu'il désigne. Le premier nom attribué au type nomenclatural est le **basionyme**. Ainsi, **tous les synonymes nomenclaturaux se réfèrent au même basionyme** et à la même publication initiale de la description du taxon (protologue). Sauf cas particuliers, les différents synonymes nomenclaturaux d'une même espèce partagent la même épithète spécifique, celle du basionyme.
2. **les synonymes taxinomiques** expriment les choix et opinions taxinomiques des auteurs. C'est le cas lorsque deux noms distincts désignent des taxons qui semblent avoir la **même délimitation**, ou une délimitation tellement semblable que l'on pense avoir affaire à un même taxon. L'examen et la comparaison des types nomenclaturaux de ces noms (planches d'herbiers) permet en principe d'en avoir le cœur net : le botaniste convaincu que deux types appartiennent effectivement à la même espèce doit en tirer les conséquences et appliquer le principe de priorité : c'est le nom validement publié à la date la plus ancienne qui est **légitime** pour ce taxon. Les autres sont des **synonymes taxinomiques** ou "**hétérotypiques**" (car ils se réfèrent à des types différents).

Noms vernaculaires

Un **nom vernaculaire** ou **nom commun** est un nom usuel, en langue locale, donné à une ou plusieurs espèces végétales dans son pays ou sa région d'origine. Un tel nom peut donc désigner des taxons très différents selon le contexte. Les noms vernaculaires sont souvent repris pour former la base des **noms normalisés**, équivalents des noms scientifiques dans une langue donnée. Malgré leurs imprécisions, ils ont souvent la faveur dans le langage courant et font partie de la culture populaire. Par contre, il y a souvent plusieurs noms vernaculaires pour le même taxon et il est illusoire de vouloir en privilégier un seul.

Nom français normalisé

Un guide a été élaboré pour attribuer à chaque taxon de la **flore vasculaire de France** un “**nom français normalisé**” (NFN) unique (MATHIEU, D., DURECU, M., MERCIER, D., MATHEZ, J., CHAUVET, M., 2015. *Guide de nomenclature des noms normalisés en français pour les plantes trachéophytes de France métropolitaine*. Code NFN Version 2.4 - novembre 2014. *J. Bot. Soc. Bot. France*, 70). Ce nom est de caractère binomial. Il est écrit en caractères droits avec une majuscule au nom du genre. Exemple *Quercus ilex* L. => Chêne vert. L’attribution de ces NFN s’effectue dans le cadre d’un programme de travail collaboratif animé par Tela Botanica en deux étapes : d’abord l’attribution des genres en français puis des espèces et des rangs inférieurs.

Stratégie envisagée pour les référentiels botaniques de Tela Botanica

Référentiels taxinomiques

Ils concernent les **inventaires de taxons régionaux** (flores de différentes entités géographiques) ou les **inventaires thématiques** (révision d'une famille par exemple) et, dans ce cadre restreint, supposent de la part de leurs auteurs toute une série de choix qui expriment des opinions personnelles (ou consensuelles). Ce sont des outils indispensables puisqu'ils mettent à disposition un vocabulaire permettant d'échanger des informations sur les taxons.

Le nom de chaque taxon est **le nom retenu par les auteurs**, nom correct correspondant aux choix taxinomiques dans le cadre du référentiel. Le nom retenu est accompagné par tout ou partie de ses **synonymes** (nomenclaturaux et taxinomiques). Chaque version successive d'APG est un référentiel mondial pour les familles et rangs supérieurs d'Angiospermes.

Causes des difficultés de fusion des référentiels taxinomiques

Lorsqu'il s'agit d'un **référentiel territorial**, les informations données pour chaque taxon sous le nom retenu ne sont **valables qu'à l'intérieur de ses limites territoriales** : les indications morphologiques (description, illustrations...), biologiques, écologiques, chorologiques ne peuvent être appliquées sans une vérification rigoureuse aux organismes désignés ailleurs, dans d'autres référentiels et dans d'autres contrées, par le même nom. La simple agrégation d'informations recueillies à partir des noms et synonymes de référentiels taxinomiques de territoires différents ne peut produire qu'une “**synthèse hypothétique**” à soumettre à la critique scientifique.

C'est le travail qui est réalisé par le MNHN pour la réalisation du référentiel TaxRef, en agrégeant les référentiels taxinomiques des différents territoires nationaux, c'est également ce travail qui est en cours de réalisation dans le cadre du projet “The Plant List” depuis de nombreuses années, et qui à ce jour a permis de regrouper 1 064 035 noms dont 350 699 noms retenus et 470 624 synonymes de différentes sources, mais avec encore 242 712

noms qui ne sont pas résolus <http://www.theplantlist.org/>. Le même travail a été réalisé par Cyril Chatelain pour l'Afrique en regroupant les référentiels de l'Afrique Tropicale, l'Afrique Australe et l'Afrique du Nord dans son référentiel African Plant Database (APD) <http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php>

Les référentiels taxinomiques à Tela Botanica

Compte tenu des difficultés énoncées ci-dessus, **Tela Botanica ne fusionne pas les référentiels taxinomiques provenant de sources différentes.** De ce fait, Tela Botanica utilise plusieurs référentiels dans ses applications (actuelles et à venir), notamment :

- **BDNFF** (Base de données nomenclaturales de la Flore de France qui est en fait une base taxinomique), en cours d'extension à l'échelle de l'Europe par Benoît Bock
http://www.tela-botanica.org/page:eflore_bdtfx?langue=fr
- **EuroSL 1.0** établi à partir de Euro+med en complétant par Flora Europaea pour les pays du Nord permettant d'avoir une liste des plantes d'Europe (en restant au niveau spécifique la plupart du temps), afin de disposer rapidement d'un référentiel de travail harmonisé pour développer EIVE 1.0, la base européenne des valences écologiques des plantes.
- **TaxRef** (MNHN) pour les territoires et départements d'OutreMer
<https://inpn.mnhn.fr/programme/referentiel-taxonomique-taxref>
- **APD** (African Plant Database) du Conservatoire et Jardin botaniques de la Ville de Genève pour l'Afrique (hors Madagascar)
<http://www.ville-ge.ch/musinfo/bd/cjb/africa/recherche.php>
- **VASCAN** (Base de données des plantes vasculaires du Canada) pour le Canada, notamment le Québec, <http://data.canadensys.net/vascan/>
- **D'autres référentiels** comme Lebanon Flora pour le Liban
<http://www.lebanon-flora.org/>, Madagascar avec le Missouri Botanical Garden, etc.

Définition d'un référentiel nomenclatural

Un référentiel nomenclatural est la compilation de tous les noms attribués aux taxons dans la littérature et en particulier dans les différents référentiels taxinomiques. Il comporte tous les noms valides de ces référentiels (noms corrects et synonymes) mais peut être étendu aux autres désignations (avec mention des arguments justifiant le caractère non valide de leur publication).

Un tel référentiel devrait facilement faire l'objet d'un consensus universel dans la mesure où il repose sur des informations objectives vérifiables : combinaisons nomenclaturales publiées, auteur, date, références bibliographiques...

Cependant, il existe des variantes de présentation formelle des noms dans les différents référentiels taxinomiques, ce qui complique la compilation et nécessite une vérification soignée de ces informations :

1. **variantes orthographiques** = désignations : par ex : *Rubus stereacanthus* / *stereacanthos* / *steracanthos* / *steracanthus* ou *sylvatica* et *silvatica*.
2. **variantes d'autorité** = il y en a beaucoup dans la Flore de Bonnier et Layens, par exemple ; mais on en trouve partout. Notamment, *Flora gallica* a éliminé les "ex", et par exemple *Abies concolor* (Gordon & Glend.) Lindl. ex Hildebr. devient *Abies concolor* (Gordon & Glend.) Hildebr.

3. **variantes de conventions d'écriture** = par ex : *Solanum nigrum* subsp. *schultesii* (Opiz) Wessely / *Solanum nigrum* L. subsp. *schultesii* (Opiz) Wessely) ; "et" au lieu de "&" (entre les auteurs), les abréviations d'auteur (Matzke-Hajek est devenu récemment Matzk.), les noms d'auteurs parfois non abrégés... Une base des données de correspondance sur les noms d'auteurs serait donc nécessaire.
4. **variantes sur le statut d'hybride ou non** = par exemple *Populus* (x) *canescens*.

Utilité d'un référentiel nomenclatural

Pour chaque nom, la mise en ligne des protogues, l'identification précise, la localisation et la mise en ligne des images des types nomenclaturaux des noms valides sont des indications extrêmement précieuses pour faciliter le travail des systématiciens et l'élaboration de référentiels taxinomiques pertinents.

Comme par ailleurs, ces informations sont de nature nomenclaturale, elles sont partagées par l'ensemble des référentiels taxinomiques qui mentionnent ces noms, d'où l'intérêt de tous les regrouper dans un seul référentiel nomenclatural.

Position par rapport à l'IPNI

L'IPNI (*International Plant Names Index*) est un référentiel nomenclatural élaboré en collaboration entre [The Royal Botanic Gardens \(Kew\)](#), [The Harvard University Herbaria](#), et [the Australian National Herbarium](#). Il regroupe l'ensemble des noms reconnus par ces organismes, leur associe leur bibliographie de base (avec les liens vers les publications quand elles sont disponibles) et leur affecte un identifiant unique. L'IPNI tient également à jour un index des noms d'auteurs avec leurs abréviations standard et une bibliothèque de publications. C'est donc une source d'informations extrêmement précieuses et incontournable sur le plan nomenclatural.

Cependant, nous ne pouvons nous en satisfaire pour les raisons suivantes :

- **incomplétude nomenclaturale** : la couverture de l'IPNI est mondiale, mais ne retient que les noms connus de ses fondateurs. De ce fait, elle ignore nombre de noms qui sont contenus dans NOS référentiels
- **incomplétude bibliographique** : l'IPNI ne mentionne que peu de choses concernant chaque nom : source bibliographique et basionyme.

Il convient donc de réaliser notre propre référentiel nomenclatural contenant :

- **tous les noms cités dans tous nos référentiels** ;
- **toutes les données dont nous disposons sur ces noms** : protogues (lien vers la source, document en pdf, traduction...), types connus, liens vers les scans de ces types, basionymes et synonymes nomenclaturaux ;
- **les liens vers les sites Internet de référence** mentionnant les noms (dont l'IPNI) et vers les citations dans nos propres référentiels taxinomiques.

Résolution du problème des variantes d'écriture

Comme nous l'avons vu précédemment, la compilation des noms des différents référentiels taxinomiques met en évidence des variantes d'écriture qui ne permettent pas une fusion automatique des noms similaires.

Pour résoudre ce problème, nous suggérons de ne pas “fusionner” les noms identiques ou similaires, mais de les associer au sein d'une même “**entité lexicale**” qui regroupe tous les noms similaires, en gardant pour chacun le lien vers son référentiel source.

Ainsi, **Coleostephus myconis (L.) Cass. ex Rchb.f.** (BNFbdtx-nn-1858) issu de la BDNFF et **Coleostephus myconis (L.) Rchb. f.** (isfan-nn-137180) issu du référentiel d'Afrique, sont regroupés dans la même **entité lexicale**, avec la mention de leur source.

De ce fait, lors d'une mise à jour d'un référentiel taxinomique, si un nom a été modifié, il suffit de le remplacer par le nouveau en se référant à son identifiant, sans rien modifier par ailleurs.

Il conviendra par contre de réaliser une procédure informatique qui sera capable de créer les “entités lexicales” en :

- identifiant automatiquement les noms qui sont identiques ;
- repérant ceux qui sont “presque identiques” (fuzzy search) afin qu'un opérateur valide ou non leur appartenance au même ensemble lexical.

Solution pratique envisagée

Si l'on se place maintenant d'un point de vue informatique, ou plus exactement dans le formalisme des bases de données relationnelles, tout ce qui vient d'être dit se résout de façon étonnamment simple ! Voici comment.

On concatène dans une table unique tous les noms contenus dans tous les référentiels taxinomiques utilisés en leur associant les index suivants :

1. un code définissant la **source du référentiel** (BDNFF, TaxRef, AFD...) ;
2. un index attaché de façon univoque à chaque nom. Ce **numéro nomenclatural** (“num_nom”) est propre à chaque référentiel. D'où la clé unique pour l'ensemble des noms composée du code référentiel et de cet index nomenclatural ;
3. un index donnant pour chaque référentiel la **référence du nom retenu**. Cet index est le même que son “num_nom” si le nom est lui-même retenu, différent dans le cas des synonymes. Il est ainsi facile de regrouper tous les synonymes d'un taxon pour un référentiel donné ;
4. un index donnant la **référence du basionyme**. Cet index permet de regrouper tous les synonymes nomenclaturaux dans un référentiel. Il est normalement identique pour tous les référentiels, mais pas nécessairement, permettant ainsi de faire des vérifications ;
5. un **index lexical** donnant la référence de l'entité lexicale (index arbitraire) à laquelle appartient chaque nom permettant de relier toutes données d'ordre nomenclatural. C'est le plus dur et le plus long à réaliser !

Ainsi, cette table constitue un référentiel nomenclatural unique en intégrant tous les référentiels taxinomiques de Tela Botanica. D'autre part, on conserve une table par référentiel permettant de gérer les informations taxinomiques non contenues dans le référentiel nomenclatural unique, telles que la référence au taxon de rang supérieur ou les autres informations décrites ci-dessous. Ces bases seront globalement identiques à celles des fournisseurs de données, mises au format utilisé par Tela Botanica.

Rattachement des connaissances aux référentiels

Les référentiels constituent l'épine dorsale autour de laquelle s'agrègent toutes les autres informations dont nous disposons sur la flore des contrées qui sont référencées. Ces informations sont de deux types : nomenclaturales et taxinomiques :

- **les information nomenclaturales** sont celles qui sont attachées strictement aux noms. Ce sont : les synonymes lexicaux (écritures dans les différents réf. taxinomiques), les synonymes nomenclaturaux, les basionymes, les références bibliographiques, les liens vers les protologues, les liens vers les scans de types (planches d'herbier numérisées), les liens vers l'IPNI et vers différents sites Internet citant les noms... Elles sont partagées par l'ensemble des noms appartenant à la même entité lexicale, indépendamment des référentiels taxinomiques d'où elles ont pu être extraites. Elles sont accessibles à partir du référentiel nomenclatural.
- **les informations taxinomiques** sont celles qui sont attachées aux taxons. Elles sont donc intimement liées aux différents référentiels taxinomiques et ne peuvent pas être partagées entre eux. C'est le cas des synonymes taxinomiques, de la chorologie, de l'écologie, des observations de terrain, des photographies... propres aux plantes citées dans chaque référentiel. Elles sont accessibles en consultant indépendamment chaque référentiel taxonomique. Par ailleurs, les taxons n'existant pas "concrètement" dans la base de données, ces informations ne peuvent être attachées qu'aux noms (retenus et synonymes) qui sont mentionnés dans chaque référentiel et qui composent un taxon. Ainsi pour un taxon donné, dans un référentiel donné, la chorologie pourrait être attachée au nom retenu, son écologie à l'un de ses synonymes et les illustrations à plusieurs synonymes différents... Toutes ces informations étant alors présentées simultanément lorsque l'on affiche le nom du taxon dans le référentiel.
- Nota : en cas de regroupement de deux taxons, de séparation d'un taxon en deux ou du changement du rattachement d'un nom à un taxon, les informations liées à ce nom suivent celui-ci quelle que soit sa destination.

Conclusion : six grands principes adoptés par Tela Botanica

Ce que nous venons de présenter se résume sur le plan opérationnel à retenir six principes fondamentaux :

Principe 1 : les référentiels taxinomiques sont indépendants les uns des autres.

Chaque référentiel suit sa propre vie (mises à jour, corrections), indépendante de celle des autres. Les référentiels taxinomiques gardent intégralement leur identité (pas de fusion taxinomique).

Principe 2 : toutes les informations liées aux référentiels sont toujours rattachées à un nom et un seul

Qu'elles soient de nature nomenclaturale (protologue par exemple) ou taxinomique (illustration), ces données sont toujours liées à un nom et un seul via son numéro nomenclatural ("num_nom"). Peu importe que ce nom soit correct ou synonyme.

Principe 3 : les connaissances qui sont de type taxinomique sont propres à un référentiel et à lui seul

Elles ne peuvent pas être partagées par les autres référentiels. C'est le cas des illustrations, de la répartition, de l'écologie...

Principe 4 : les connaissances taxinomiques relatives à un taxon dans un référentiel donné sont obtenues en agrégeant l'ensemble des connaissances de type taxinomique attachées à chacun des noms relatifs à ce taxon dans ce référentiel (son nom retenu et ses synonymes). Elles sont regroupées sous le nom correct du taxon.

Principe 5 : les combinaisons nomenclaturales des différents référentiels relevant du même nom gardent leur originalité orthographique, mais sont regroupées dans une “entité lexicale” commune

Notamment, on ne corrige pas les erreurs d'écriture des noms, cette action relevant de la responsabilité du fournisseur du référentiel. On peut, si on le souhaite, décréter que l'une des écritures est "conforme" à des règles qui nous sont propres (basée sur Flora Gallica par exemple), tout en respectant les autres écritures. Mais rien ne nous y oblige.

Principe 6 : les connaissances qui sont de type nomenclatural et issues des différents référentiels sont agrégées et partagées par tous les noms appartenant à la même entité lexicale

C'est le cas notamment des protologues, des basionymes, des liens vers les types, la bibliographie, etc.