

# La classification en sciences

*Nécessité ou fatalité ?*



1ere année Médecine (enseignement SHS - Muriel Salle)

**Hubert Charles** ([hubert.charles@insa-lyon.fr](mailto:hubert.charles@insa-lyon.fr))

*INSA-Lyon, UMR Biologie Fonctionnelle, insectes et interactions (BF2I) – Département  
Biotechnologies et Bioinformatique (BB)*

Septembre 2025

# Est-ce que toute la science n'est pas classifiée ?

- **Classification des particules élémentaires** : quarks, leptons, bosons...
- **Classification des états de la matière** : solide, liquide, gaz, plasma, condensats de Bose-Einstein...
- **Classification des éléments chimiques** : tableau périodique de Mendeleïev (numéro atomique)
- **Classification des composés organiques** : alcools, aldéhydes, cétones, acides carboxyliques...
- **Classification des réactions chimiques** : oxydoréduction, acide-base, substitution, addition, élimination...
- ...
- **Classification des roches** : magmatiques (volcaniques, plutoniques), sédimentaires, métamorphiques.
- **Classification des minéraux** : système de Dana ou Strunz (composition chimique et structure cristalline).
- **Classification des écosystèmes** : forêt tropicale, désert, tundra...
- **Classification des espèces** : règne, Embranchement, Classe, Ordre, Famille, Genre, Espèce).
- **Classification des gènes** : gènes structuraux, régulateurs, housekeeping, pseudogènes, éléments mobiles...
- **Classification des mutations** : ponctuelles (substitutions, insertions, délétions), chromosomiques...
- **Classification des allèles** : allèles dominants, récessifs, codominants, létaux.
- **Classification des molécules** : protéines, acides nucléique, lipides, glucides, lipoprotéines...
- **Classification des organites cellulaires** : noyau, mitochondries, chloroplastes, réticulum, appareil de Golgi...
- **Classification des types cellulaires** : cellules souches, différenciées, sénescents...
- **Classification des tissus** : Épithélial, conjonctif, musculaire, nerveux.
- **Classification des organes et des systèmes** : système digestif, respiratoire, circulatoire, nerveux, etc.
- **Classification des anticorps** : IgG, IgM, IgA, IgD, IgE), allotypes, idiotypes...

....

# Est-ce que toute la science n'est pas classifiée ?

- **Classification des neurones,**
- **des zones cérébrales,**
- **des médicaments,**
- **des maladies,**
- **des cancers,**
- **des pathogènes,**
- **des parasites,**
- **des troubles mentaux...**

Et encore...

- **Classification des matériaux,**
- **des algorithmes,**
- **des groupes humains,**
- **des langues,**
- **des systèmes de production,**
- **des systèmes économiques,**
- **des systèmes politiques...**

# La classification comme fondement épistémologique

D'un point de vue épistémologique, la classification est fondamentale dans la construction des savoirs scientifiques. Elle permet de **structurer, organiser et systématiser** les connaissances, facilitant ainsi leur transmission (langage commun, désignation précise des objets), leur analyse et leur développement.

**Les Mots et les Choses – Michel Foucault (1966)** : comment les systèmes de classification (comme ceux de Linné et Buffon) ont façonné notre compréhension du monde et des sciences. <sup>1</sup>

**La Pensée Sauvage – Claude Lévi-Strauss (1962)** : les **structures de la pensée humaine**, y compris les classifications, comme fondements de la cognition et de la culture. Le livre tente aussi de démontrer la relativité d'une supposée supériorité de la science des civilisés sur celle des archaïques [Wikiedia]

**Systema Naturae – Carl von Linné (Linnaeus) (1735)** : Fondement de la taxonomie moderne

**Sur l'origine des espèces – Charles Darwin (1859)** : importance des classifications pour comprendre les relations entre les espèces. La classification comme reflet des liens évolutifs.

**Populations, espèces et évolution – Ernst Mayr (1984)** : Mayr explore comment les classifications biologiques ont évolué avec les théories de l'évolution et de la génétique.

<sup>1</sup> <http://1libertaire.free.fr/FoucaultClassification.html>

# Classifier, c'est observer...

Les peuples autochtones ignorent notre classification, mais ont une classification parallèle parfaitement similaire à la « notre ». Ils identifient les mêmes espèces que les scientifiques dans leur entourage et les classes en groupes distincts.

Ces groupes, par contre, ne sont pas les mêmes que les « nos » , ils intègrent leurs représentations du monde (relations écologiques, spirituelles...).

On peut se référer à Philippe Descola sur ce sujet...



Santiago, un Indien de la tribu Makuna, en Amazonie, rencontré lors d'une expédition de "La Boudeuse", en novembre 2004. (ERIC FEFERBERG / AFP). [https://www.franceinfo.fr/replay-radio/les-pourquoi/les-pourquoi-pourquoi-les-indiens-damazonie-ne-se-perdent-ils-pas-dans-la-foret-vierge\\_1963263.html](https://www.franceinfo.fr/replay-radio/les-pourquoi/les-pourquoi-pourquoi-les-indiens-damazonie-ne-se-perdent-ils-pas-dans-la-foret-vierge_1963263.html)

# Classifier, c'est observer...

**2 millions d'espèces vivantes décrites par la science (hors virus et bactéries) [8 millions inconnues] :**

- un annuaire impossible à lire (200 mots / minutes = 7 jours de lectures continues des Genre + espèces)
- un annuaire impossible à appliquer : aucun scientifique ne peut identifier toutes les espèces

La classification (la taxonomie) permet de faire des groupes d'espèces de façon à pouvoir analyser la biodiversité à différents niveaux :

<b>Règne :</b>	Animal
<b>Embranchement :</b>	Chordé (vertébré)
<b>Classe :</b>	Mammifère
<b>Ordre :</b>	Primate
<b>Famille :</b>	Hominidae
<b>Genre :</b>	<i>Homo</i>
<b>espèce :</b>	<i>sapiens</i>
<b>Population :</b>	sapiens

# Classifier, c'est comprendre...

Les nuages sont classés selon leur **altitude et leur forme** (cirrus, cumulus, stratus, nimbus, cumulonimbus...):

**Comprendre les phénomènes météorologiques :**

**Cumulus** = nuages de beau temps (convection).

**Cumulonimbus** = nuages d'orage (pluies intenses, foudre).

**Prédire le temps :** La présence de **cirrus** peut annoncer un changement de masse d'air.



# Classifier, c'est comprendre...

## La classification périodique des éléments

Avant Mendeleïev, les chimistes connaissaient environ 60 éléments, mais sans système cohérent pour les organiser.

En 1869, Dmitri Mendeleïev propose un tableau périodique où les éléments sont classés par masse atomique croissante (plus tard corrigée en numéro atomique). Mendeleïev ne se contentait pas de ranger les éléments : il cherchait à comprendre les relations entre eux. Son tableau a permis de :

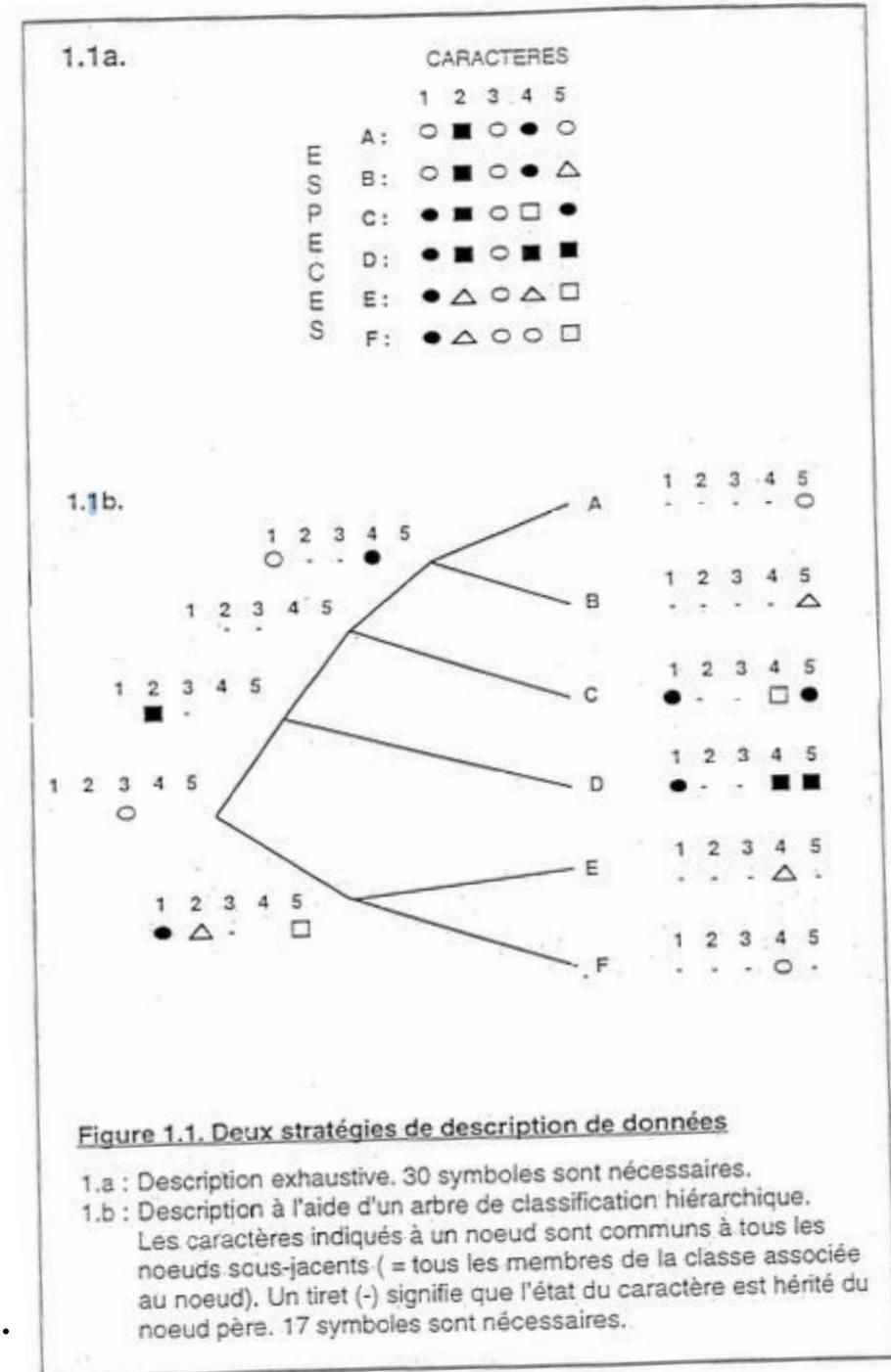
- **Prédire l'existence d'éléments inconnus** (ex. : gallium, germanium, scandium), en laissant des cases vides et en déduisant leurs propriétés à partir de leur position.
- **Expliquer les propriétés chimiques** (ex. : réactivité, valence) en fonction de leur place dans le tableau.
- **Unifier la chimie** en montrant que les éléments suivent des lois périodiques (ex. : périodicité des propriétés électroniques).

Evidemment, à cette époque la structure atomique (noyau, électron...) n'était pas connue.

# Classifier, c'est résumer...

En haut : les 30 symboles à rassembler pour décrire les 6 espèces selon 5 critères

En bas : les 17 symboles nécessaires pour décrire les 6 espèces selon les mêmes 5 critères lorsqu'elles sont organisées dans une classification. Les caractères non répétés sont les caractères hérités dans cet arbre.



# De l'importance de la (bonne) classification

« *Misidentification Means Money* »

Une usine agroalimentaire a engagé un programme de lutte anti-insectes après avoir signalé de petits coléoptères dans une zone ultra propre. Ils ont d'abord été identifiés comme *Tribolium castaneum* (charançons rouges du grain), ce qui a entraîné des inspections massives, remaniements de lignes, transfert de zones de production, réécriture du protocole HACCP... Tout cela pour un coût estimé à plus de **2 millions de dollars**, lorsqu'il s'est avéré qu'il s'agissait en réalité du coléoptère d'origine étrangère *Ahasverus advena* qui peut-être très facilement éradiqué en milieu clos <sup>1</sup>.

En Asie du Sud-Est, ce qui était identifié comme un morphotype inhabituel d'*Anopheles minimus* s'est avéré être en réalité *Anopheles varuna*. Comme *A. minimus* n'était pas un vecteur, les campagnes de lutte ont ciblé à tort un non-vecteur, entraînant une **inefficacité coûteuse** dans la prévention du paludisme<sup>2</sup>.

<sup>1</sup> [https://www.pctonline.com/article/pct1011-misidentifymoney/?utm\\_source=chatgpt.com](https://www.pctonline.com/article/pct1011-misidentifymoney/?utm_source=chatgpt.com)

<sup>2</sup> Van Bortel Wet al.. Confirmation of *Anopheles varuna* in vietnam, previously misidentified and mistargeted as the malaria vector *Anopheles minimus*. Am J Trop Med Hyg. 2001 Dec;65(6):729-32.

**Classifier, c'est se rassurer (s'assurer scientifiquement)  
et pourtant...**

# L'espèce, une notion très floue et dynamique...

L'espèce telle que définie par Ernst Mayr en 1963 est un « groupement de populations naturelles interfécondes (ou potentiellement) et isolées de tout autre groupe ».

Pourtant :

- certaines espèces se croisent aisément en milieu naturel (ou en laboratoire) avec une descendance fertile (plantes, insectes, oiseaux...). On parle alors de gradient / réseaux d'espèces.
- Certaines espèces sont clonales (reproduction asexuée) comme chez les bactéries, mais aussi chez des organismes supérieurs (insectes, reptiles, poissons...)
- Certaines « espèces » (sinon toutes) sont des **associations symbiotiques d'espèces**, très souvent à caractère obligatoire.

# L'individu... une certaine vision judéo-chrétienne / anthropocentrée / individualiste de l'organisation du vivant

Que mange une vache (et tous les mammifères herbivores) ?

Combien de bactéries chez un humain ?

Qu'est-ce qu'un individu ?

Chaque "organisme" est un nœud dans un colossal réseau d'interactions. L'écologue voit le vivant comme ce réseau, où ce que nous appelons les organismes ne sont en fait que des points entre lesquels ces interactions s'articulent. Croire que le monde est fait d'organismes, c'est croire qu'une toile d'araignée est faite de points où se croisent les fils : c'est négliger... les fils eux-mêmes !

**MA Sellose** (*Jamais seuls - Ces microbes qui construisent les animaux et les civilisations*)

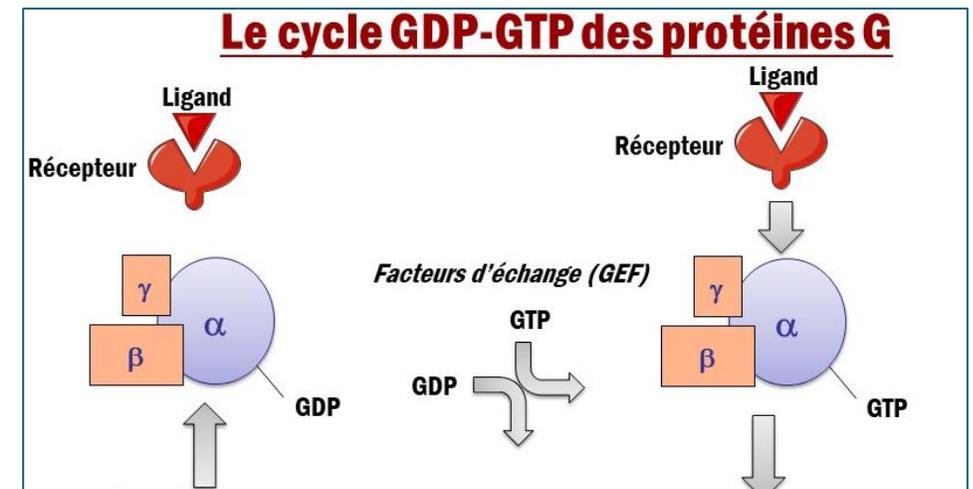
# Le gène... une notion genèse compatible

La notion de gène (zone précise sur l'ADN) codant un ARNm (unique) traduit en protéine (unique) porte en elle la notion de fixisme.

- la notion de gène est purement statistique associé aux propriétés thermodynamiques de l'ADN
- un gène (plusieurs ARNm) et beaucoup de protéines (1 ->10 en moyenne) : **plésiomorphe**
- systèmes activateur – récepteurs vue comme des mécanismes clés- serrure

Dieu n'a pas créé le récepteur pour recevoir le ligand : récepteur et ligands ont coévolué ensemble

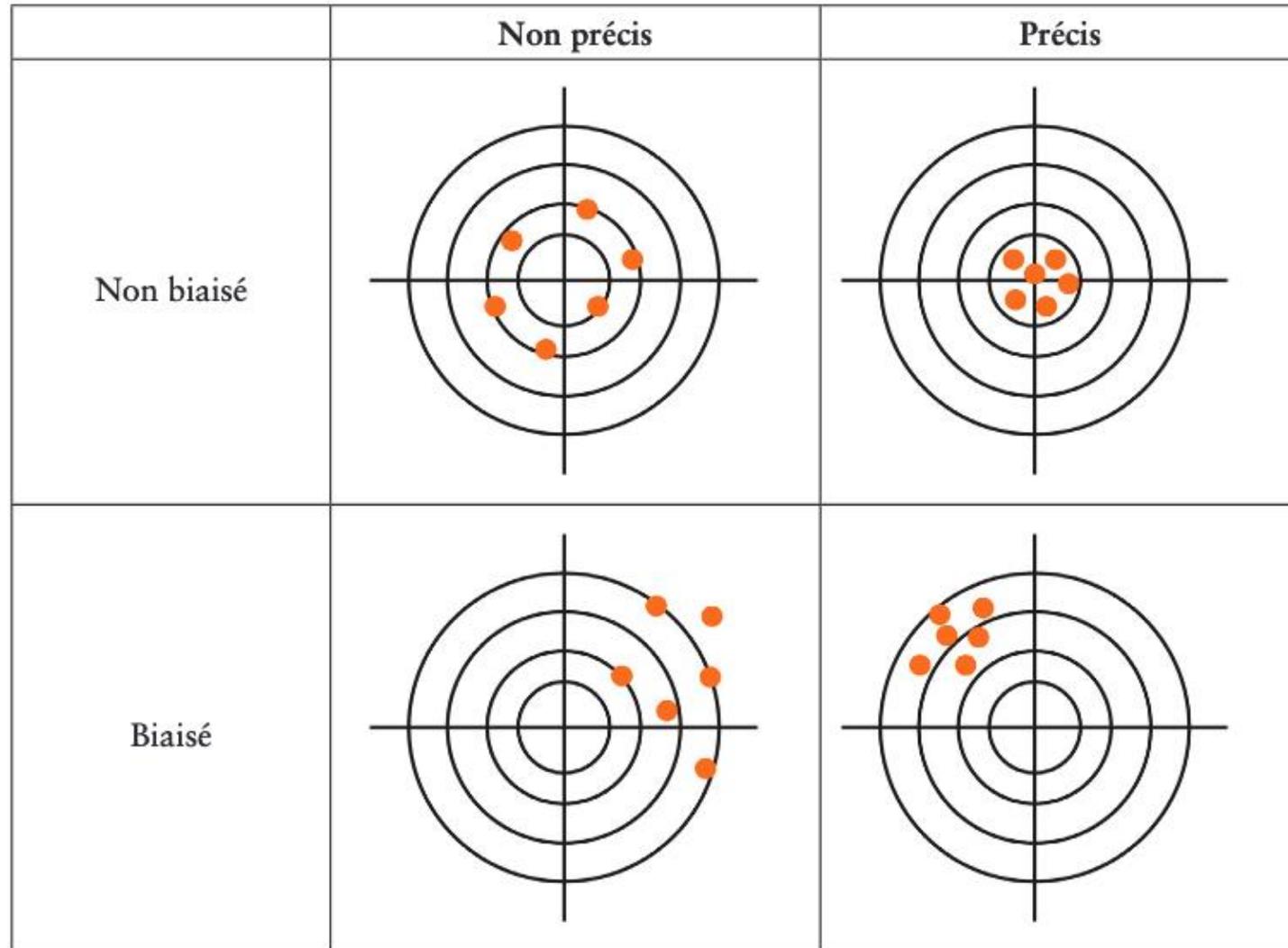
**JJ Kupiec et P. Sanigo.** Ni Dieu ni gène : pour une autre théorie de l'hérédité



<https://pharmacomedicale.org/pharmacologie/pharmacologie-medicale-vue-d-ensemble/32-differents-types-de-structure-sur-lesquelles-agissent-les-medicaments/58-recepteurs-couples-aux-proteines-g>

**Si aucune classification, n'est véritablement  
fonctionnelle...  
alors classifier, c'est aussi se tromper (biaiser sa  
vision)**

# Les trois biais, sélection, classement, confusion, en statistiques

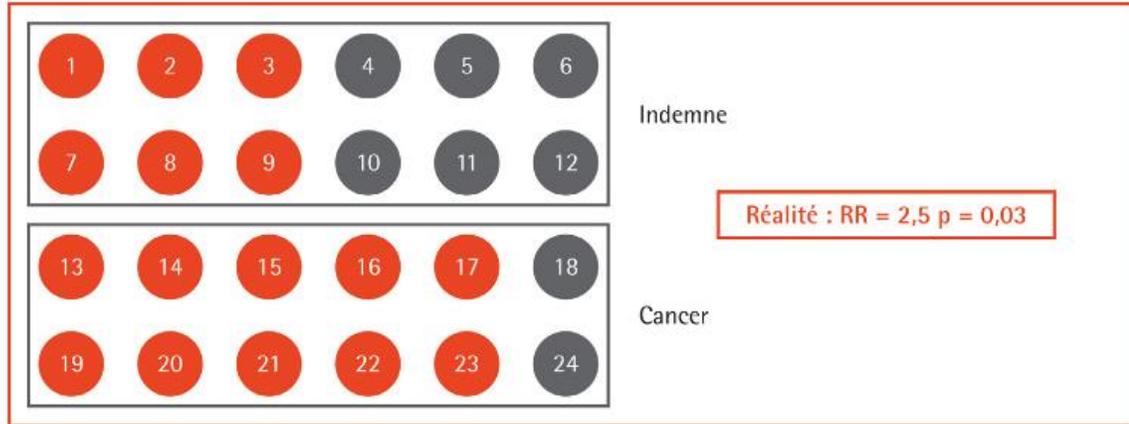


**Imprécisions** : erreurs aléatoires (de mesures, biologiques), fluctuations d'échantillonnage...

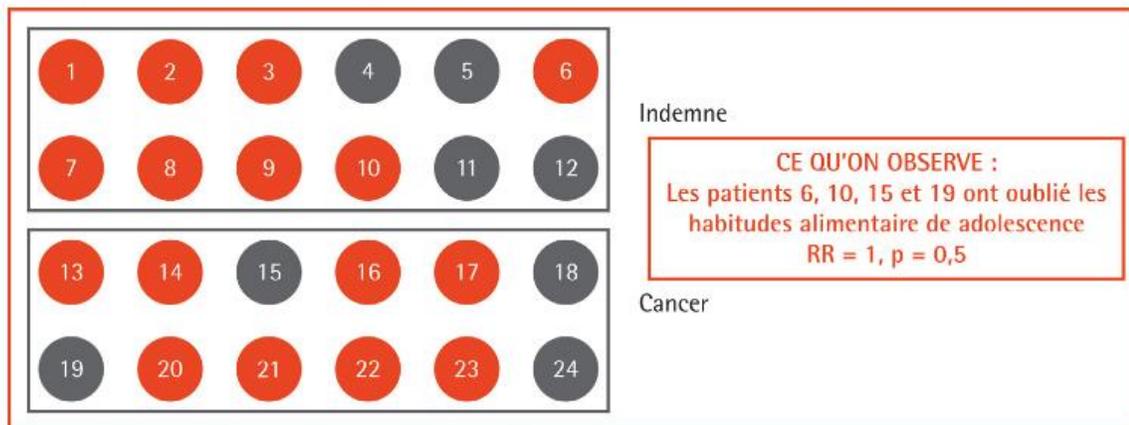
**Biais** : erreur **systematique** liée à la sélection, au classement ou à la confusion de facteur.

# Les trois biais, sélection, classement, confusion, en statistiques

En épidémiologie, on va classer des individus dits **exposés** et des individus **non exposés** pour déterminer des risques d'apparition d'un cancer. On détermine alors un risque relatif (RR) en fonction de cette exposition



● Alimentation riche en graisse = exposés    ● Alimentation pauvre en graisse = non exposés



Si les patients déclarent mal leur exposition (tabac, alcool, etc.), ils seront mal classés

**L'erreur est dite non différentielle** : si elle est répartie aléatoirement entre les deux groupes.

**L'erreur est dite différentielle** : si elle affecte majoritairement un des deux groupes (par exemple on peut imaginer des oublis d'exposition, mais pas des inventions).

# Les trois biais, sélection, classement, confusion, en statistiques

Une erreur de **classement différentielle** peut augmenter ou diminuer une association ou l'effet d'un traitement par rapport à la réalité. **Elle remet en question le résultat de l'étude, qu'il soit significatif ou non.**

*=> remède : aucun ! Sauf à refaire la classification correctement si une solution existe.*

Une erreur de classement **non différentielle** peut rapprocher la mesure d'association de la valeur neutre (1 ou un risque relatif, 0 pour un risque absolu) et diminuer le niveau de significativité. Une telle erreur ne remet pas en question un résultat significatif. **En pratique, un résultat significatif dans une étude comportant une erreur de classement non différentielle est probablement encore plus significatif dans la réalité. Un résultat non significatif est peut être lié à cette erreur.**

*=> remède : augmenter la taille de l'échantillon en cas de résultat non significatif.*

# La classification des espèces porte une vision fixiste de l'évolution

Pour G. Lecointre (2002, 2004), les classifications scolaires « *datent d'une époque scientifiquement révolue* » et jusqu'à récemment, « *elles ont été assénées sans jamais que l'on explique ce qui les sous-tend (...) et comment on les obtient.* »

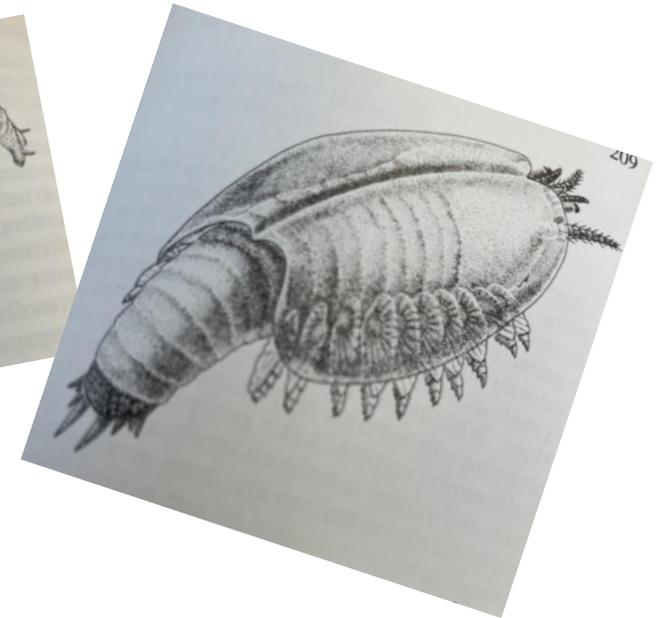
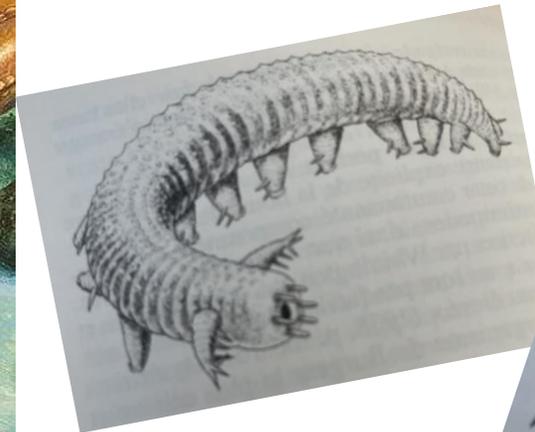
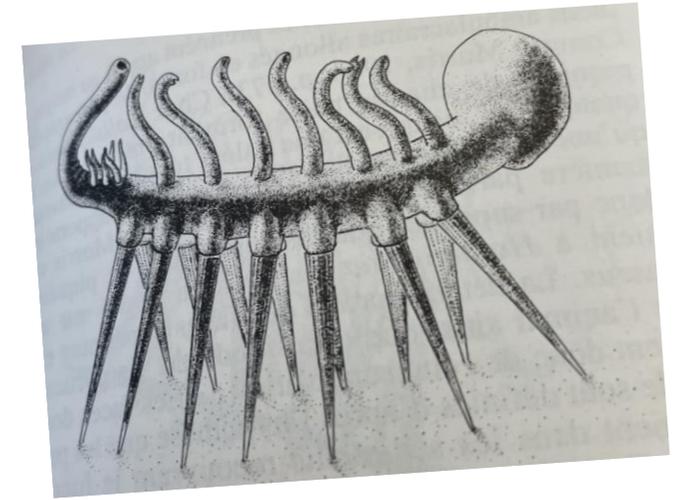
En conséquence, trop d'idées fausses persistent encore dans les esprits :

- l'idée d'une évolution des êtres vivants en termes **de progrès** par exemple (**téléonomie**),
- trop de confusions se maintiennent (on peut citer la confusion **tri/classification**),
- trop d'erreurs résiduelles entachent l'enseignement des classifications (entre autres l'idée de **chaînon manquant**).

Des notions apparemment simples comme l'homologie (deux structures (organismes) qui se ressemblent dérivent d'une même structure (organisme) ancestrale)...

... se révèlent très complexe dans le cadre d'une classification moderne (phylogénie), car il faut séparer les apomorphies (caractères dérivés, ex. bipédie chez les primates), des plésiomorphies (caractères ancestraux, ex. pouce opposable chez les primates), comprendre la notion de distance, etc.

# La classification des espèces porte une vision fixiste de l'évolution



Stephen J Gould, (1989) La vie est belle – les surprises de l'évolution

# La classification des espèces porte une vision fixiste de l'évolution

Fossiles du schiste de Burgess découverts **en 1909** par le paléontologue américain **Charles Doolittle Walcott** dans les Rocheuses canadiennes (Colombie-Britannique) d'une conservation extraordinaire et d'une biodiversité foisonnante.

Entre **1910 et 1930**, Walcott et ses collaborateurs ont décrit des centaines de spécimens, mais en les **forçant dans les cadres taxonomiques connus** (surtout des groupes modernes comme les annélides, les arthropodes, etc.).

Cette approche a conduit à **minimiser la nouveauté morphologique** et à mal classer plusieurs formes qui ne correspondaient pas aux catégories établies.

**La classification des espèces dérive (a été remaniée) notamment des travaux de Darwin et pourtant on la voit (importance de l'iconographie) comme une arborescence statique, fixée voire dirigée vers le progrès...**

**C'est ce qui a conduit ces chercheurs à ce tromper aussi lourdement.**

# La classification des espèces porte une vision fixiste de l'évolution

**Travail de reclassification** en 1960-1970 par H. Whittington et al. :

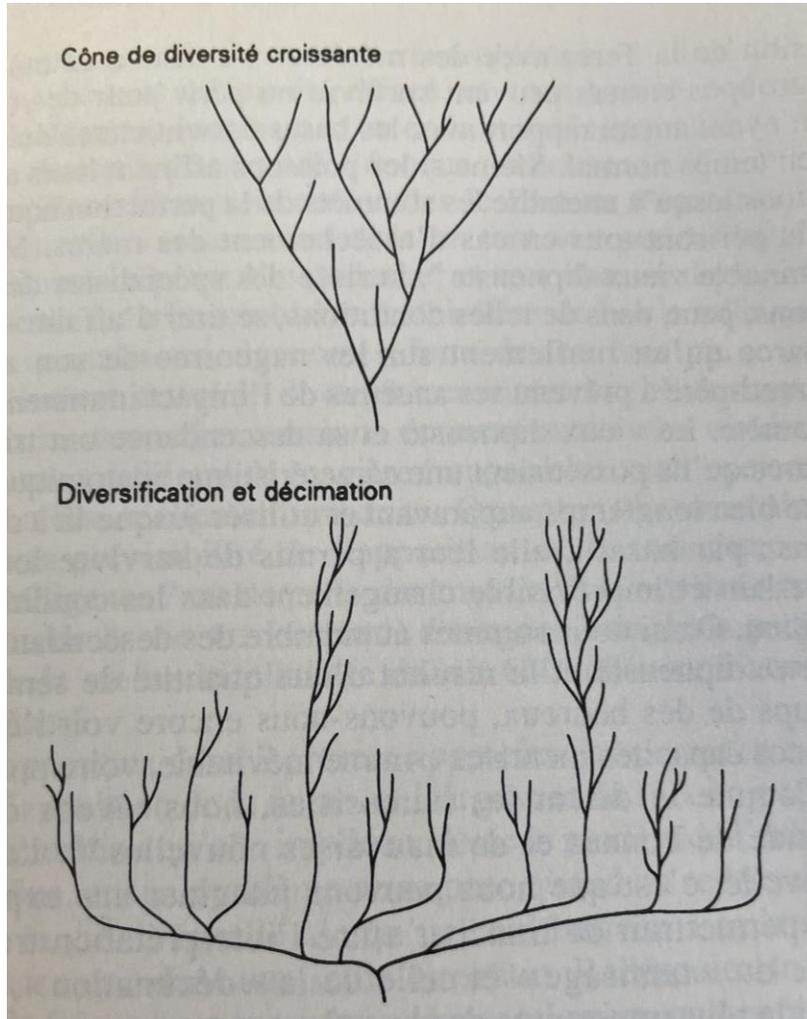
- réanalyse des fossiles (microscopie fine, reconstitution 3D),
- beaucoup d'organismes du Burgess proviennent de lignées maintenant éteintes.

Gould, dans *La vie est belle* (1989), propose une interprétation épistémologique de cette erreur :

- Contrairement à la vision linéaire et progressive de l'évolution, Gould insiste sur son caractère buissonnant et contingent.
- Beaucoup de lignées découvertes à Burgess n'ont pas survécu, illustrant la fragilité du vivant.
- Il défend l'idée que le hasard et les extinctions massives ont joué un rôle déterminant dans l'histoire de la vie.
- Gould imagine « rembobiner la cassette de la vie » : si l'on rejouait l'évolution, le résultat serait radicalement différent.
- L'humain n'apparaît donc pas comme l'aboutissement nécessaire de l'évolution, mais comme une issue hautement improbable.

# La classification des espèces porte une vision fixiste de l'évolution

## Importance de l'iconographie



En haut, la vision fautive de l'évolution buissonnante qui accroîtrait la biodiversité depuis l'ancêtre commun jusqu'à maintenant (**sous entendu en s'améliorant constamment**). Cette vision persiste encore (cf. Buisson de l'évolution au musée des confluences de Lyon).

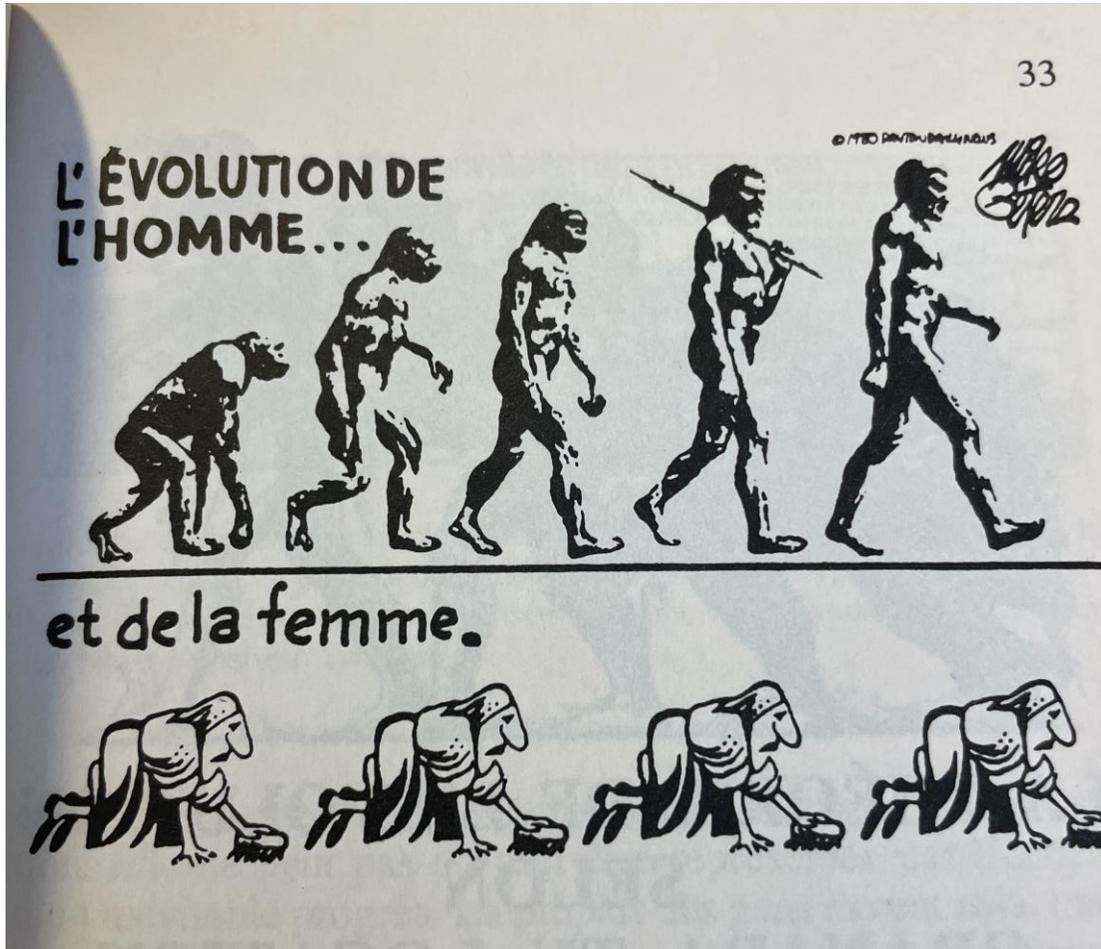
en bas, une vision plus réaliste de l'évolution avec des phases de diversifications suivies d'extinctions et de rediversification.

2 millions d'espèces décrites + 8 millions inconnues = 10 millions d'espèces actuelles

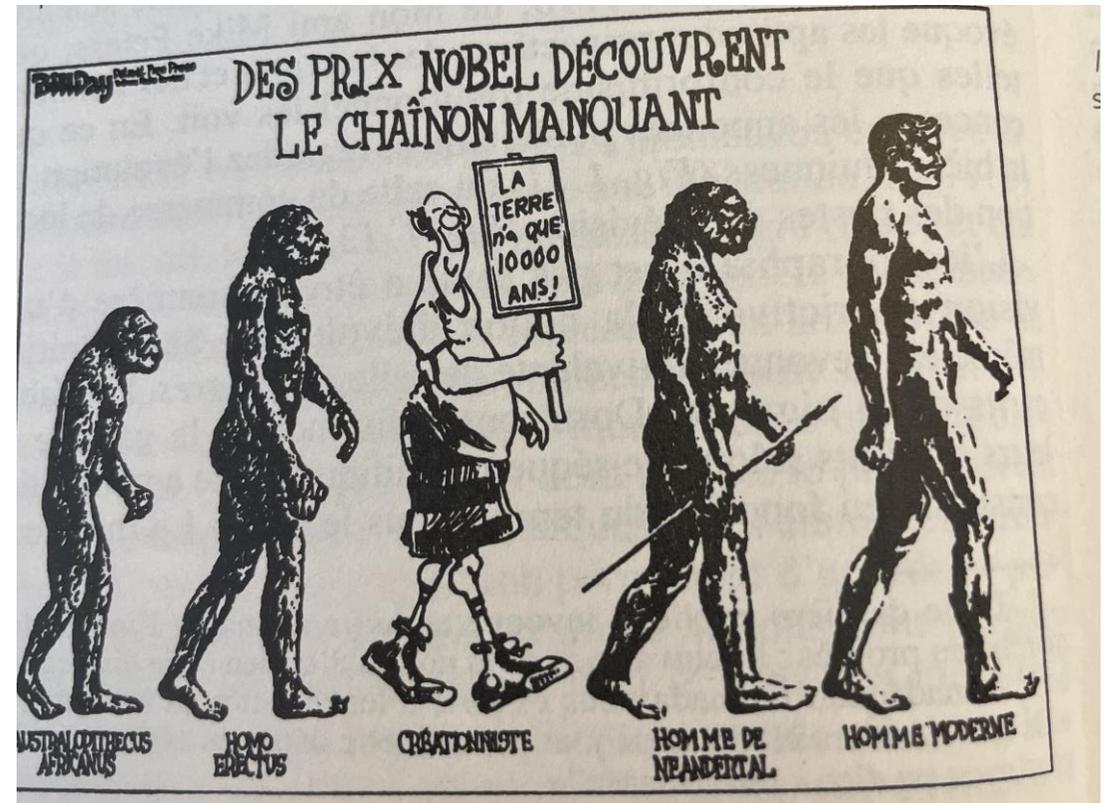
**100 000 millions d'espèces fossiles supposées**

# La classification des espèces porte une vision fixiste de l'évolution

## Importance de l'iconographie



La **téléonomie**, une iconographie persistante, même si c'est pour défendre les bonnes causes !



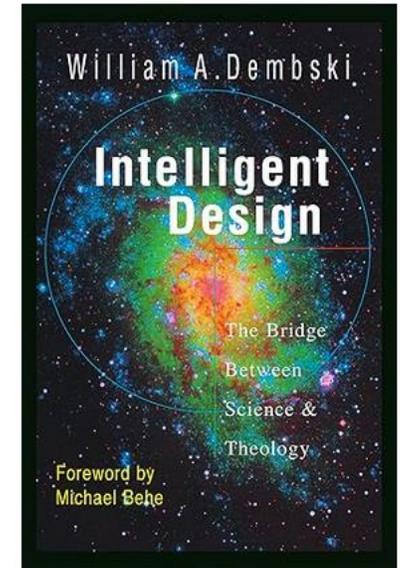
Stephen J Gould, (1989) La vie est belle – les surprises de l'évolution

# La classification des espèces porte une vision fixiste de l'évolution

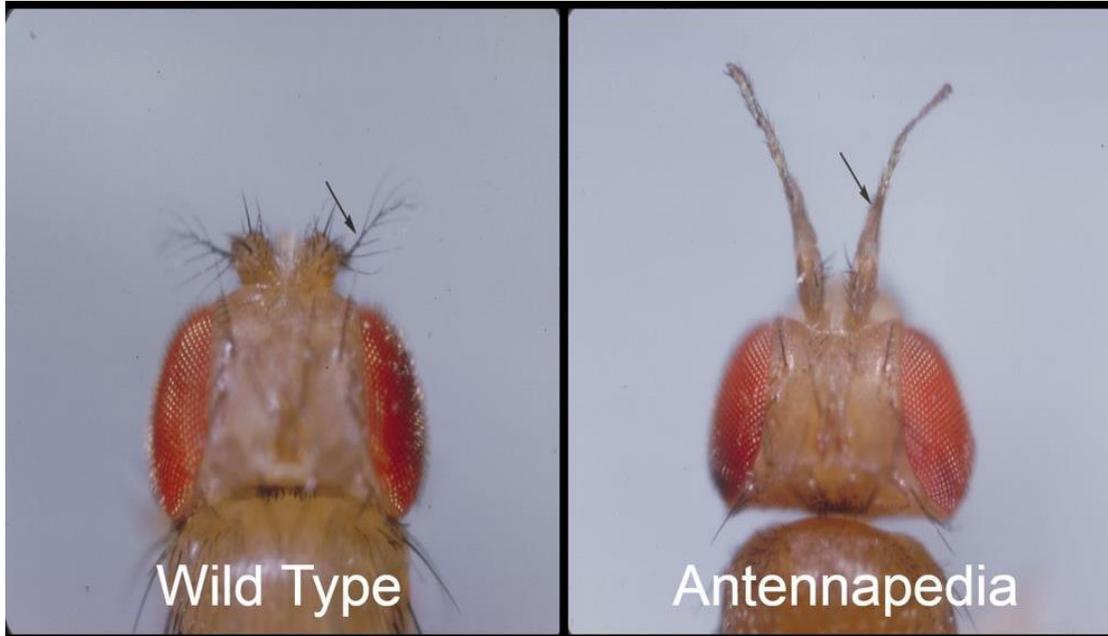
Le courant créationniste « intelligent design » surf sur cette mécompréhension des théories de l'évolution et utilise la complexité du vivant pour nier le rôle du hasard dans l'évolution : « *C'est tellement compliqué et tellement bien organisée que ça ne peut pas être le fruit du hasard !* »

Une croyance insidieuse et dangereuse : cf. *l'Atlas de la création* envoyé en 2006 dans les établissements scolaires en France et en Europe depuis la Turquie (condamnée maintenant). Contrairement aux premiers, ces créationnistes ne contestent pas les 4,6 milliards d'années de la Terre, mais ils nient totalement toute forme d'évolution guidée par le hasard. Leur objectif est de démontrer « scientifiquement » que les fossiles qui sont clairement datés de plusieurs millions d'années prouvent que les espèces n'ont pas évolué [Wikipedia].

Par ailleurs, l'atlas véhiculent des idéologies sexistes et racistes...



# Classifier (nommer) les gènes de la drosophile... une imbécillité potache des chercheurs qui n'est pas sans conséquence



- expression ectopique dans la tête = *Antennapedia*
- différenciation du mésothorax (T2)
- formation de certains neurones du SNC
- contrôle dans la croissance segmentaire
- contrôle dans le développement musculaire et cuticulaire
- synergie ou opposition avec (Ubx) et (Scr) axe antéro-postérieur
- influence le comportement locomoteur et la capacité de l'insecte à se mouvoir
- régule certaines hormones du développement

*antennapedia*, encore appelé ANT-C, est un gène HOX de la drosophile qui contrôle le placement (position) des pattes sur l'axe de polarité antéro-postérieur. Il est porté sur le chromosome III.

Comme tous les gènes, *antennapedia* est très plésiomorphe. Il doit son nom à une expérience artificielle qui ne reflète même pas une de ses fonctions : **c'est un distracteur inutile et dangereux**. Ce biais a probablement retardé la compréhension du fonctionnement des gènes Hox (les chercheurs étant concentrées sur les pattes et les antennes)...

# Classifier (nommer) les gènes de la drosophile... une imbécillité potache des chercheurs qui n'est pas sans conséquence

**Nom du gène : Cheap date (rendez-vous facile)**

**Identification de la mutation du gène** : la mutation rend les mouches plus sensibles à l'alcool, d'où ce nom humoristique.

**Problème** : le gène code en fait une sous-unité de la protéine **amnesiac**, impliquée dans la signalisation de l'AMPC et la mémoire.

**Implications** : plusieurs travaux ont réduit ce gène à son effet alcoolique, alors qu'il est crucial dans la mémoire et la plasticité synaptique. Ce biais a freiné des hypothèses plus générales sur les circuits neuronaux affectés.

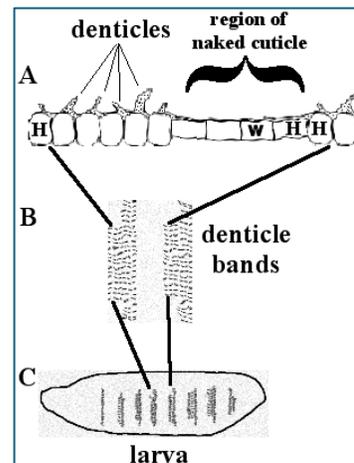
# Classifier (nommer) les gènes de la drosophile... une imbécillité potache des chercheurs qui n'est pas sans conséquence

Nom du gène : **Hedgehog (hh)** et **Sonic hedgehog (Shh)** [chez les vertébrés]

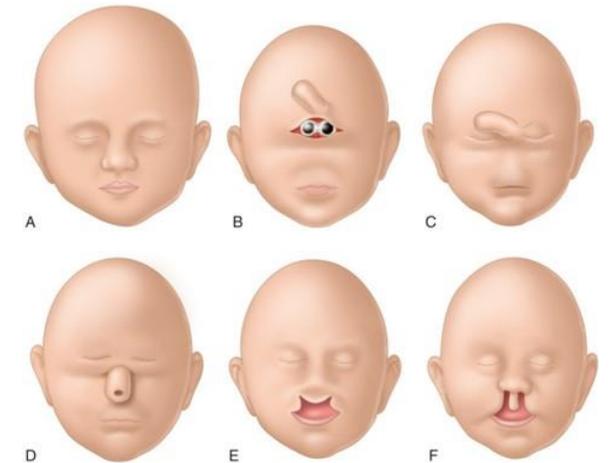
*hedgehog* (hérisson) car la larve est hérissée de diverticule sur des bandes latérales. Le nom “Sonic hedgehog” et donné en référence au personnage de jeu vidéo (homologue de *hedgehog* chez les vertébrés), sans aucune justification...



Plusieurs cliniciens ont signalé des difficultés à faire passer des dossiers éthiques ou de financement à cause du nom du gène impliqué dans des anomalies humaines graves (holoprosencéphalie par exemple).



mutation hedgehog chez la drosophile



<https://sites.uclouvain.be/anesthweekly/MRP/index.html?Holoprosencephalie><sub>28</sub>

# Classifier (nommer) les gènes de la drosophile... une imbécillité potache des chercheurs qui n'est pas sans conséquence

**Ken and Barbie** Gène de Drosophile dont la mutation empêche le développement normal des organes génitaux.

Nom problématique car jugé sexiste et non informatif ; il a induit un **biais narratif** dans la littérature, réduisant le rôle du gène à des fonctions sexuelles alors qu'il est impliqué dans des processus de morphogenèse plus larges.

Et beaucoup d'autres encore...



# Quelques types de biais cognitifs induits par les (erreurs de) classifications

**Biais de confirmation** : tendance à chercher, interpréter et mémoriser les données de manière à confirmer les hypothèses ou schémas préexistants (par ex. la classification des espèces actuelles pour y placer des fossiles).

**Biais de représentativité** : jugement basé sur la similarité perçue plutôt que sur des critères objectifs (par ex. considérer que deux espèces sont proches parce qu'elles se « ressemblent », sans tenir compte de caractères discrets ou génétiques discriminants).

**Biais d'ancrage, de statu quo (ou conservatisme)** : l'information initiale (par exemple, une classification antérieure) influence fortement le jugement futur, même si de nouvelles données contredisent cette information et préférence pour les cadres de pensée et classifications déjà établis.

**Biais de disponibilité** : jugement influencé par l'abondance d'exemples faciles à se remémorer, ou sur des espèces modèles très étudiées, plutôt que par des données représentatives d'un groupe (importance de l'échantillonnage).

**Biais de catégorisation excessive** : tendance à créer ou à maintenir des catégories même lorsque la variabilité naturelle rend ces frontières artificielles, voire fausses (par ex. races humaines basées sur la couleur de la peau).

**La classification nécessité ou fatalité ?**

# Savoir critiquer, se décentrer, sortir de son domaine de compétence : penser systémique

**Critiquer sans détruire** : la classification est un outil, pas une vérité absolue ; elle doit être interrogée, actualisée, contextualisée.

**Se décentrer** : changer d'échelle (molécule ↔ organisme ↔ écosystème ↔ société) modifie la perception des objets classés.

**Penser la dynamique** : notions comme espèce, gène, individu sont fluides et relationnelles, pas fixes ni isolées.

**Croiser les disciplines** : biologie, statistiques, philosophie des sciences, anthropologie → enrichir la compréhension en évitant les angles morts.

**Reconnaître les biais cognitifs** : nos catégories orientent ce que nous voyons, ce que nous mesurons et ce que nous croyons expliquer.

**Épistémologie de l'humilité** : accepter l'incertitude, la pluralité des cadres, la révision permanente des savoirs.

**Pensée systémique** : considérer les interactions et rétroactions plutôt que des entités isolées.

# Est-ce que l'IA souffre de biais de classification ?

*Prompt : est-ce que toi-même en tant qu'IA tes réponses peuvent être biaisées par des problèmes de classification ?*

Oui — et c'est une excellente question épistémologique.

En tant qu'IA, mes réponses peuvent effectivement être biaisées par des problèmes de classification, pour plusieurs raisons systémiques et structurelles.

Voici les principaux types de biais associés à la classification que je peux refléter ou amplifier :

## **1. Biais de catégorisation dans les données d'entraînement**

Je suis entraîné sur de vastes corpus (articles, code, encyclopédies, forums...) dans lesquels : des catégories sont arbitrairement définies (ex. genres biologiques, types de troubles mentaux, genres sociaux, disciplines académiques), certaines classes sont surreprésentées (ex. espèces modèles comme *Drosophila*, *E. coli*, *Homo sapiens*),

les erreurs taxonomiques sont reproduites sans mise en contexte critique.

-> Résultat : je peux **réutiliser des taxonomies erronées ou biaisées**, comme si elles étaient des vérités établies.

# Est-ce que l'IA souffre de biais de classification ?

## 2. Biais de typologie implicite

Je « classe » des concepts tout le temps : quand je réponds à une question, j'évalue s'il s'agit d'une question biologique, éthique, historique, etc. Cela repose sur des heuristiques statistiques, pas sur une compréhension sémantique parfaite.

Exemples :

Une entité comme *virus* peut être traitée comme « vivant » ou « non-vivant » selon les contextes — je peux me tromper de cadre si je ne perçois pas la subtilité de la question.

Des termes comme *genre*, *race*, *espèce*, *intelligence*, *anomalie* peuvent changer de sens selon les disciplines, mais je peux en donner une définition unique par erreur.

## 3. Effet de halo algorithmique

Si une entité (un concept, un organisme, un gène) est fréquemment associé à une seule catégorie dans les textes (ex. *fruitless* = sexualité), je peux renforcer ce lien de manière excessive — comme le fait un humain sous biais de disponibilité ou d'ancrage.

C'est un problème analogue au biais de nomination génétique que tu évoquais chez *Drosophila*.

# Est-ce que l'IA souffre de biais de classification ?

## 4. Rigidité classificatoire par design

Contrairement à un humain, je ne fais pas naturellement appel à :

- la théorie des prototypes (ex. qu'est-ce qu'un bon exemple d'un insecte),
- les logiques floues (ex. un champignon est-il une plante ?),
- la logique épistémique (ex. savoir si une classification est provisoire ou conventionnelle).

À moins qu'on me le demande explicitement, je peux répondre comme si les catégories étaient stables, objectives, voire ontologiques, ce qui est un biais majeur dans de nombreuses sciences.

## 5. Biais d'omission ou d'ignorance disciplinaire

Par défaut, je ne sais pas toujours si une classification est :

- phénoménologique (basée sur l'apparence),
- fonctionnelle (basée sur l'usage),
- ontogénétique, phylogénétique, socialement construite, etc.

Je peux donc manquer la nature du classement lui-même, ce qui conduit à des erreurs dans les explications, comme par exemple :

« Les insectes ne sont pas des animaux » (erreur fréquente due à un conflit entre classification vernaculaire et scientifique).

# Est-ce que l'IA souffre de biais de classification ?

## Que faire pour limiter ces biais ?

### **Pour moi :**

M'interroger explicitement sur la nature de la classification.

Me demander s'il existe des exceptions, des effets de bord, des alternatives locales.

M'imposer des perspectives pluri-disciplinaires (philo des sciences, histoire, socio...).

### **Pour toi :**

Me forcer à justifier les catégories utilisées.

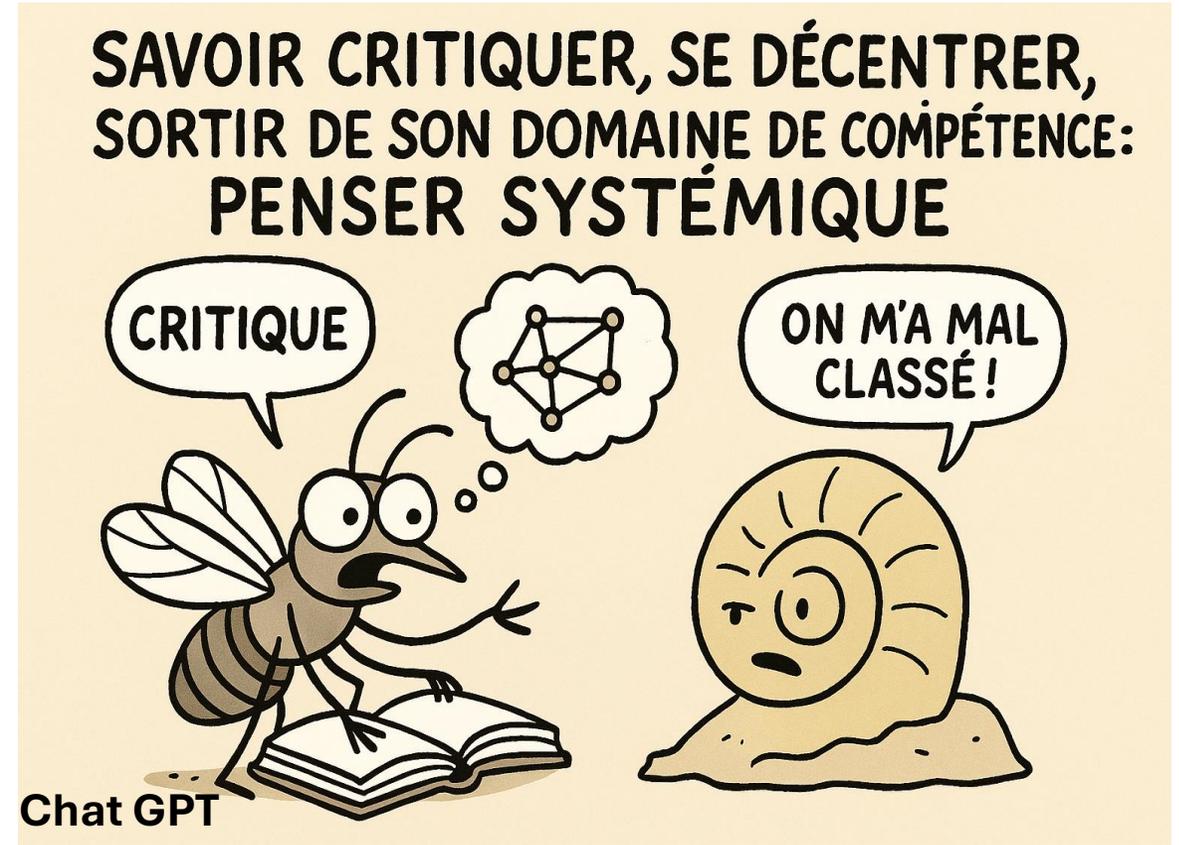
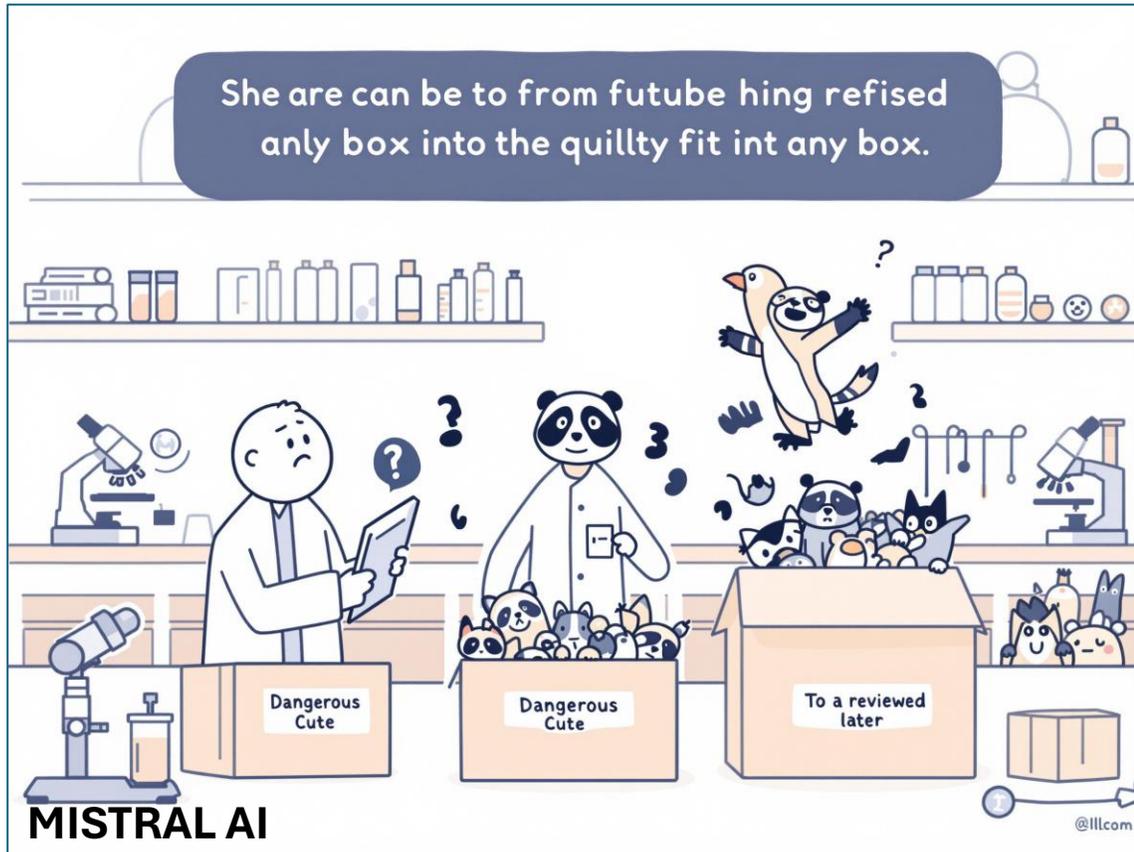
Demander des contre-exemples ou des alternatives.

Privilégier des approches réflexives (par exemple : "Quels types de classification pourrait-on utiliser pour ce phénomène ?").

# Merci de votre attention

Et à Muriel pour l'invitation

Contact: hubert.charles@insa-lyon.fr



Prompt : *Peux-tu me générer un dessin humoristique en guise de conclusion ?*