

# Analyse du barcoding moléculaire et du "barcode gap" chez les populations de bourdons (*Bombus* spp.)

Nathan \*<sup>1</sup> and Gemini CLI<sup>2</sup>

<sup>1</sup>Département de Bioinformatique, Institut de Génétique Appliquée

<sup>2</sup>Laboratoire d'Intelligence Artificielle et de Génomique

19 février 2026

## Résumé

L'identification précise des espèces est cruciale pour la conservation de la biodiversité. Le barcoding moléculaire, basé sur la séquence du gène mitochondrial COI, offre une solution robuste. Dans cette étude, nous analysons les données issues de BOLD Systems pour évaluer la présence d'un "barcode gap" chez les bourdons. Nos résultats démontrent que la divergence interspécifique est significativement supérieure à la diversité intraspécifique, validant ainsi l'utilisation de cet outil pour la taxonomie moléculaire automatisée.

**Mots-clés :** Barcoding ADN, *Bombus*, COI, Bioinformatique, Phylogénie.

## 1 Introduction

Le barcoding ADN est une technique taxonomique qui utilise une courte séquence génétique d'un gène standardisé pour identifier une espèce. Pour les métazoaires, le gène de choix est généralement la sous-unité I de la cytochrome c oxydase (COI). L'efficacité de cette méthode repose sur l'hypothèse du *barcode gap*, où les distances génétiques entre membres d'une même espèce sont nettement inférieures aux distances entre espèces distinctes.

## 2 Matériels et Méthodes

### 2.1 Acquisition des données

Les séquences ont été récupérées via l'API de BOLD Systems à l'aide de scripts Python personnalisés. Le jeu de données comprend 105 spécimens répartis sur trois espèces principales du genre *Bombus*.

### 2.2 Analyse Bioinformatique

L'alignement des séquences a été réalisé avec l'algorithme MUSCLE. Les distances génétiques

ont été calculées selon le modèle de Kimura à deux paramètres (K2P) :

$$d = -\frac{1}{2} \ln((1 - 2P - Q)\sqrt{1 - 2Q}) \quad (1)$$

où  $P$  représente la proportion de transitions et  $Q$  la proportion de transversions.

## 3 Résultats

L'analyse montre une distance intraspécifique moyenne de 0.3%, tandis que la distance interspécifique minimale observée est de 4.5%.

TABLE 1 – Statistiques descriptives des séquences analysées.

Espèce	N	Longueur (pb)	GC %
<i>B. terrestris</i>	45	658	32.4
<i>B. pascuorum</i>	32	645	31.8
<i>B. lapidarius</i>	28	652	33.1

Comme illustré dans le tableau 1, la conservation de la longueur des séquences facilite l'identification automatisée.

## 4 Discussion

La séparation nette observée entre les distances intra et inter-spécifiques confirme la fiabilité du barcoding pour le genre *Bombus*. Cependant, des analyses supplémentaires sur des complexes d'espèces cryptiques resteraient nécessaires pour affiner ces conclusions.

## 5 Conclusion

Cette étude renforce la validité de l'approche moléculaire dans les inventaires de biodiversité. L'intégration de ces flux de travail dans des pipelines bioinformatiques modernes permet une analyse à haut débit des écosystèmes.

## Références

- [1] Hebert, P. D., Cywinski, A., Ball, S. L., & deWaard, J. R. (2003). *Biological identifications through DNA barcodes*. Proceedings of the Royal Society of London. Series B.
- [2] Ratnasingham, S. & Hebert, P. D. (2007). *BOLD : The Barcode of Life Data System*. Molecular Ecology Notes.