

# Une approche RAD-seq pour délimiter des espèces dans le complexe *Bactrocera dorsalis* (Diptera, Tephritidae)

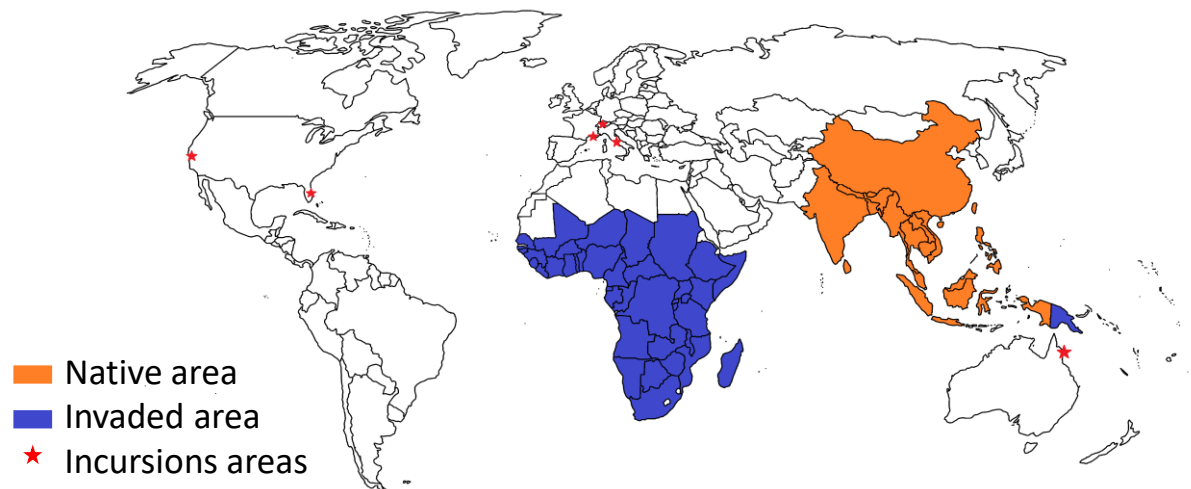
---

CHARBONNEL Emeline



## La mouche orientale des fruits : *Bactrocera dorsalis*

- Ravageuse polyphage des fruits et légumes tropicaux et subtropicaux
  - Hôte principal : Mangue
- Native d'Asie
- Invasive en Afrique sub-Saharienne, Océan Indien et Océanie



## Le complexe *Bactrocera dorsalis*

- 89 espèces
- 39 attirées au Méthyleugénol (leurre chimique mâle utilisé pour l'échantillonnage et la surveillance)
- 21 bien référencées dans la littérature



*B. arecae*

*B. bivittata*

*B. cacuminata*

*B. carambolae*

*B. caryeae*

*B. dorsalis*

*B. endiandrae*

*B. indonesiae*

*B. irvingiae*

*B. kanchanaburi*

*B. kandiensis*

*B. latilineola*

*B. musae*

*B. occipitalis*

*B. ochroma*

*B. opiliae*

*B. pallida*

*B. paraarecae*

*B. raiensis*

*B. unimacula*

*B. verbascifoliae*

## Le complexe *Bactrocera dorsalis*

- 89 espèces
- 39 attirées au Méthyleugénol
- 21 bien référencées dans la littérature
- 6 ravageuses sympatriques avec *B. dorsalis*



*B. arecae*

*B. bivittata*

***B. cacuminata***

***B. carambolae***

***B. caryeae***

***B. dorsalis***

*B. endiandrae*

*B. indonesiae*

*B. irvingiae*

*B. kanchanaburi*

***B. kandiensis***

*B. latilineola*

*B. musae*

***B. occipitalis***

*B. ochroma*

*B. opiliae*

*B. pallida*

*B. paraarecae*

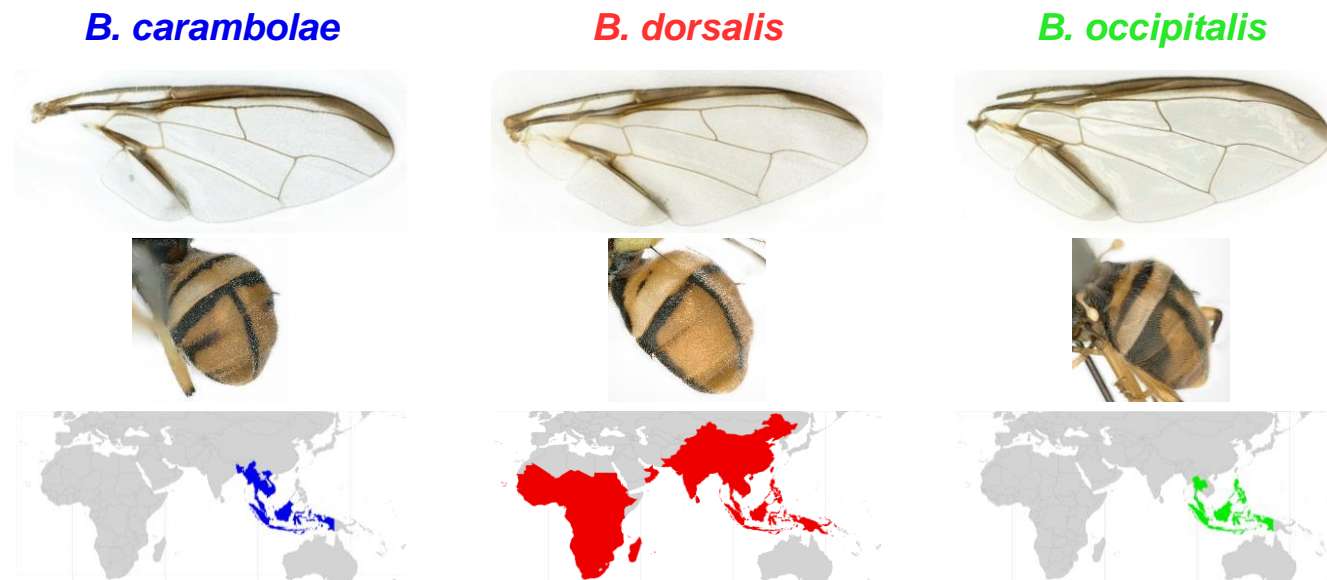
*B. raiensis*

*B. unimacula*

*B. verbascifoliae*

## Le complexe *Bactrocera dorsalis*

- Dont 2 espèces morphologiquement difficiles à différencier avec *B. dorsalis*
  - 45% dans les pays insulaires d'Asie du Sud-Est (Indonésie, Malaisie, Philippines) (Charbonnel et al. submitted)



- Diagnostique moléculaire actuel limité (Charbonnel et al. submitted)

## Le complexe *Bactrocera dorsalis*

➤ Des cas d'hybridation référencés :

*B. dorsalis* – *B. carambolae*

*B. dorsalis* – *B. occipitalis*

- Introgression mitochondriale
- Introgression nucléaire ?

*B. dorsalis* – *B. raiensis*

*B. dorsalis* – *B. kandiensis*

└─→ 23,5% dans l'aire sympatrique (Inde, Sri Lanka)  
(Charbonnel et al. submitted)

➤ Définition de certaines espèces confuse

*B. dorsalis* – *B. invadens* – *B. papayae* – *B. philippinensis* → Synonymes ?!

(Schutze et al. 2015)

## Objectif de l'étude

- Mieux comprendre la **délimitation des espèces** au sein du complexe *B. dorsalis*
- Rechercher la présence d'**hybridation** et d'introgression nucléaire
- Déterminer des **marqueurs diagnostics** des espèces

→ Potentiel des **NGS** pour accéder à l'information génomique

Chez les mouches des fruits : Catullo et al. 2019, Doorenweerd et al. 2022, Zhang et al. 2021

→ Utilisation du RAD-seq (Baird et al. 2008) :

- Chaque séquence est associée à un individu
- Plusieurs milliers de loci homologues dispersés le long du génome
- Peut être utilisé pour les taxons dont les connaissances préliminaires sont limitées

### Echantillonnage dans le complexe *B. dorsalis*

- 13 espèces sur les 21 espèces cibles
  - Rareté des autres espèces dans les collections
  - Mauvaise qualité de l'ADN des spécimens de collection à sec
  
- 66 spécimens
  - 64 identifiés (morphologie + COI + EIF3L; Charbonnel et al. submitted)
  - 2 spécimens non identifiés
  
- + 8 spécimens dupliqués

<i>Bactrocera cacuminata</i>	2
<i>Bactrocera carambolae</i>	8
<i>Bactrocera dorsalis</i>	23
<i>Bactrocera endiandrae</i>	6
<i>Bactrocera kandiensis</i>	3
<i>Bactrocera latilineola</i>	4
<i>Bactrocera musae</i>	2
<i>Bactrocera occipitalis</i>	5
<i>Bactrocera opiliae</i>	2
<i>Bactrocera pallida</i>	2
<i>Bactrocera raiensis</i>	3
<i>Bactrocera unimacula</i>	2
<i>Bactrocera verbascifoliae</i>	2
<b>13 espèces</b>	<b>64</b>



## Production de données NGS à l'échelle du génome

- Construction des librairies par **approche BestRAD** (Ali et al. 2016)
  - Hybridation sur billes de Streptavidine → diminuer les taux de dupliques PCR → ~27%
  - **Optimisation** du protocole (Benoit et al. 2023, protocols.io)
    - Robuste à des ADN de faibles quantité et qualité → min = 3ng/μL / 5kb
    - Diminution du taux d'adaptateurs libres et des séquences chimériques

🔒 A FAIR protocol of the Best-RAD sequencing approach 👥

Laure Benoit<sup>1</sup>, Sabine Nidelet<sup>2</sup>, Emeline Charbonnel<sup>1</sup>,  
Marie-Pierre Chapuis<sup>1</sup>

- Séquençage des marqueurs RAD → **Illumina NovaSeq**

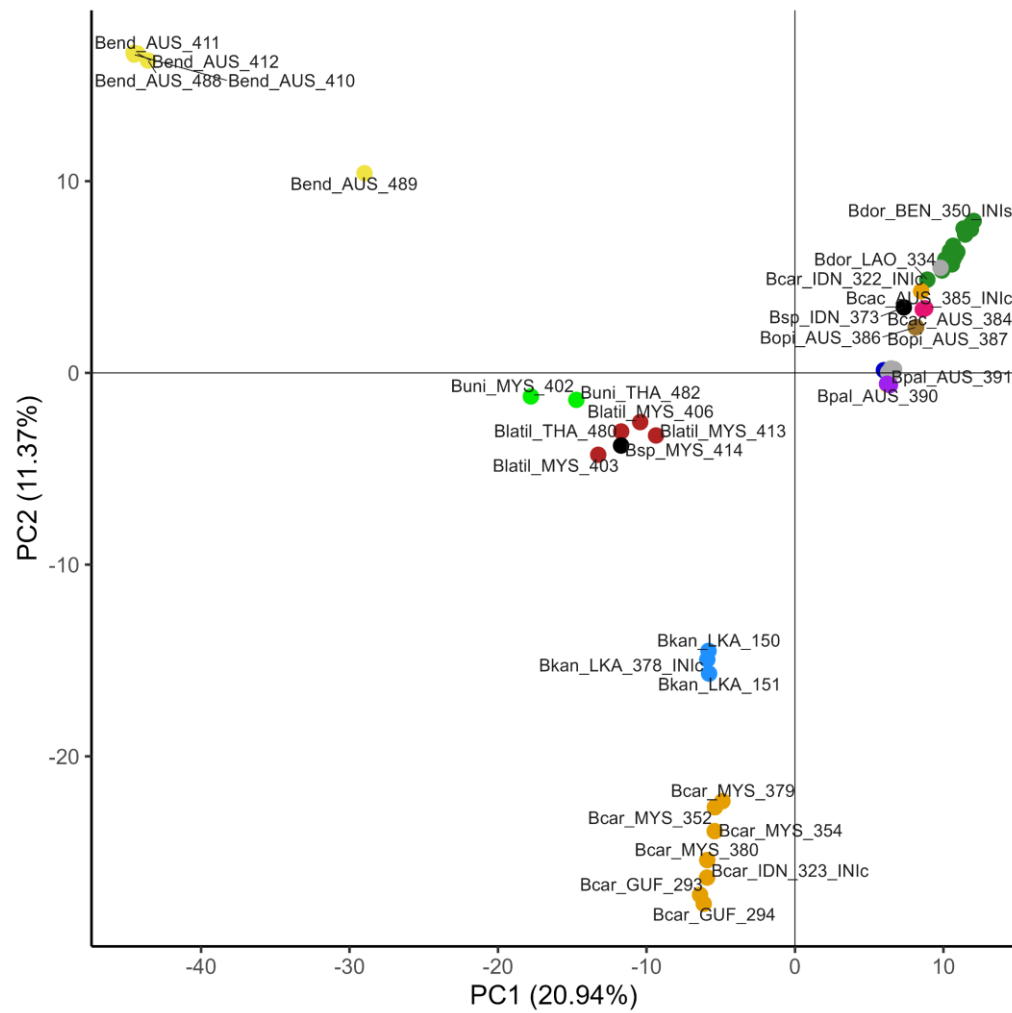
## Analyses bio-informatiques

- Démultiplexage + Denoising (Stacks2)
  
- Construction du catalogue de loci (Stacks2)
  - Alignement sur génome de référence de *B. dorsalis* masqué des éléments transposables (49%)
  - En *de novo*
  
- Filtres sur les SNPs bi-alléliques (SNPfiltR) :
  - Conserver les SNP avec une Maf > 0,05 et profondeur min de 6X
  - Eliminer les paralogues et chimères
  - Conserver les génotypes de qualité
  - Eliminer les SNP/individus avec 80%/50% de données manquantes
  
- Sélection des SNPs indépendants

## Production de données NGS à l'échelle du génome

- 2 216 351 910 reads totaux
- 47% de mapping
- 6 individus retirés (> 50% de données manquantes) → Perte de 2 espèces : *B. raiensis* et *B. verbascifoliae*  
→ 11 espèces étudiées
- 150 930 SNPs
- Couverture moyenne = 34,2
- Taux de données manquantes = 0,076
- 0,16% d'erreur de génotypage moyen
- 14 633 SNPs indépendants

## Les 11 espèces sont discriminées les unes des autres



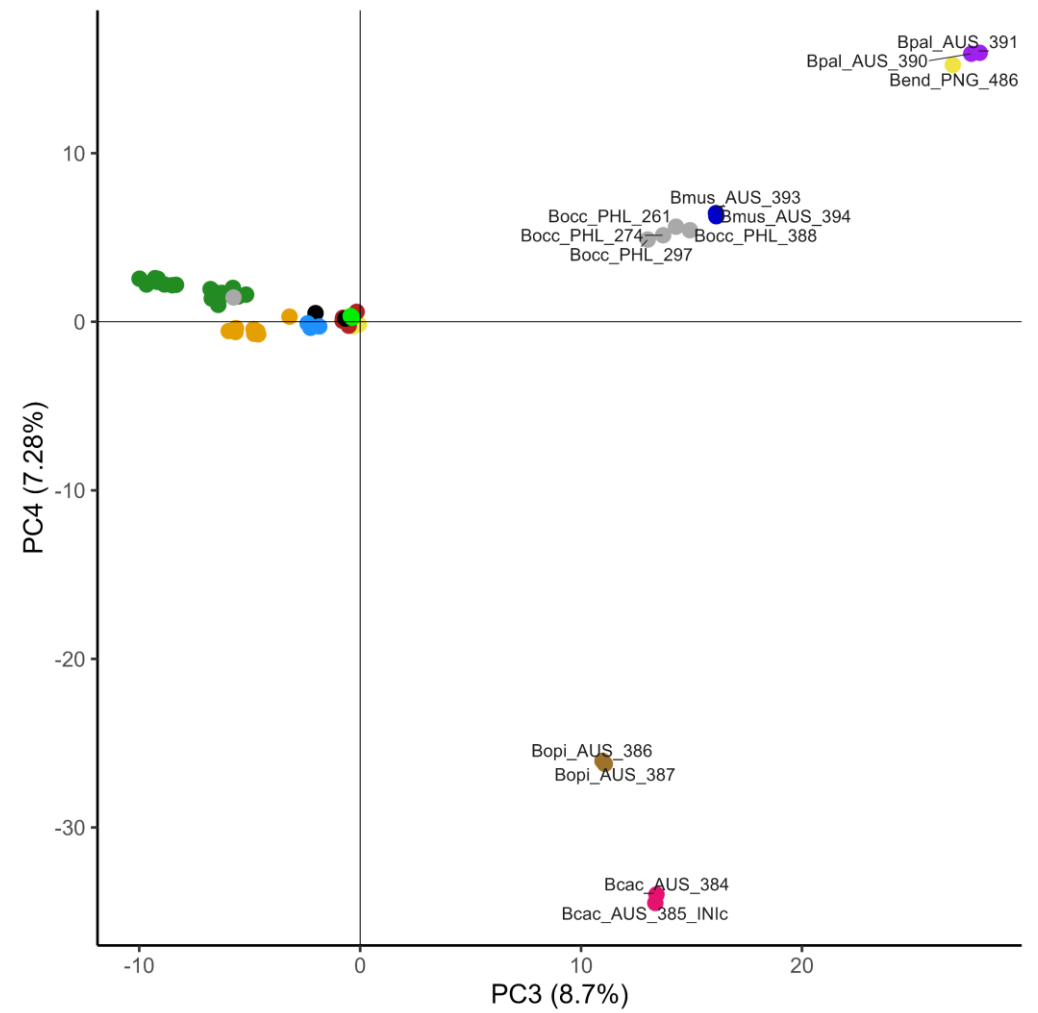
*Bactrocera carambolae*

*Bactrocera endiandrae*

*Bactrocera kandiensis*

*Bactrocera latilineola*

*Bactrocera unimacula*



*Bactrocera cacuminata*

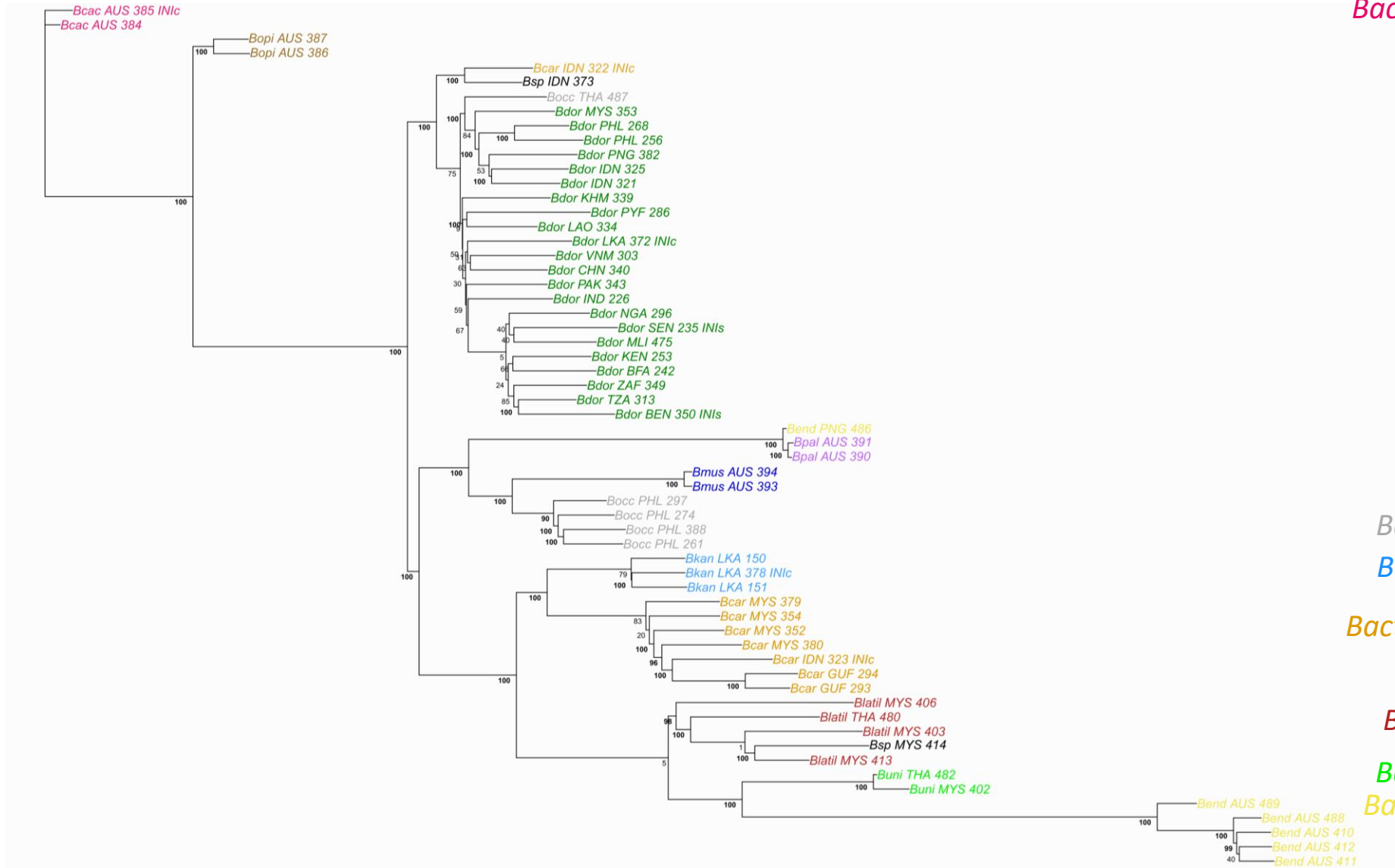
*Bactrocera musae*

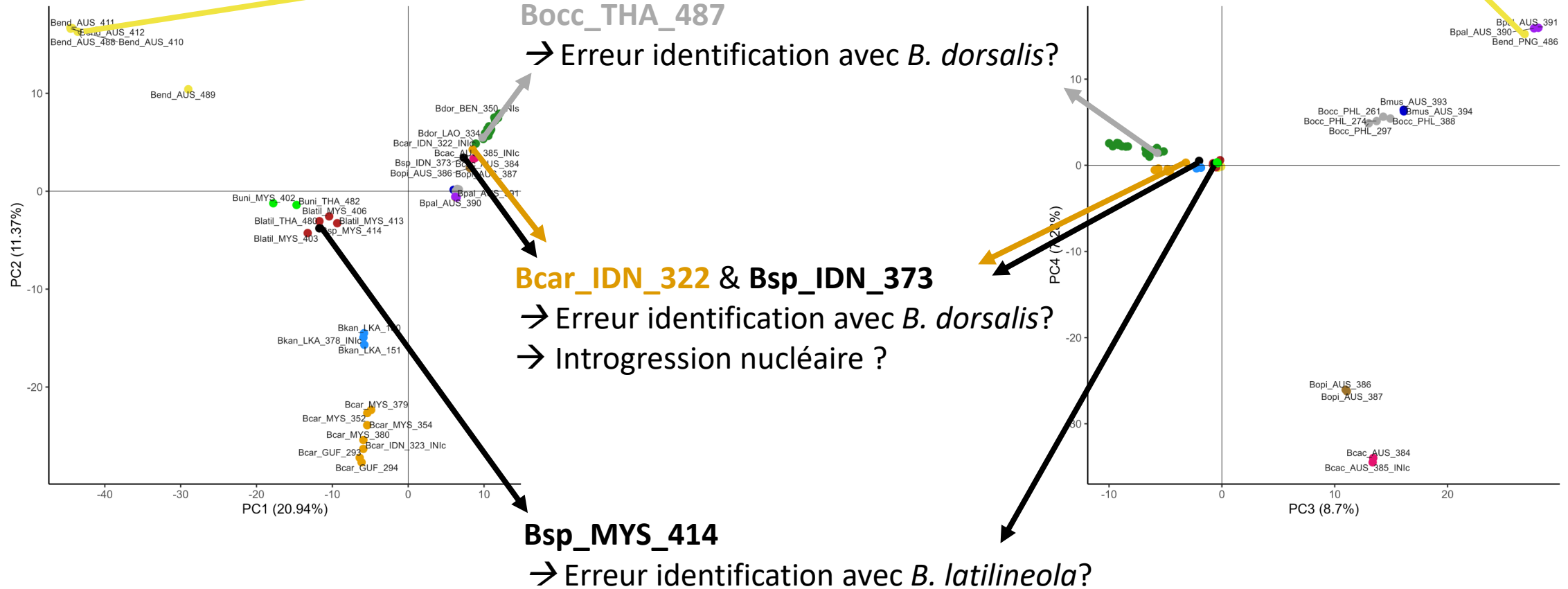
*Bactrocera occipitalis*

*Bactrocera opiliae*

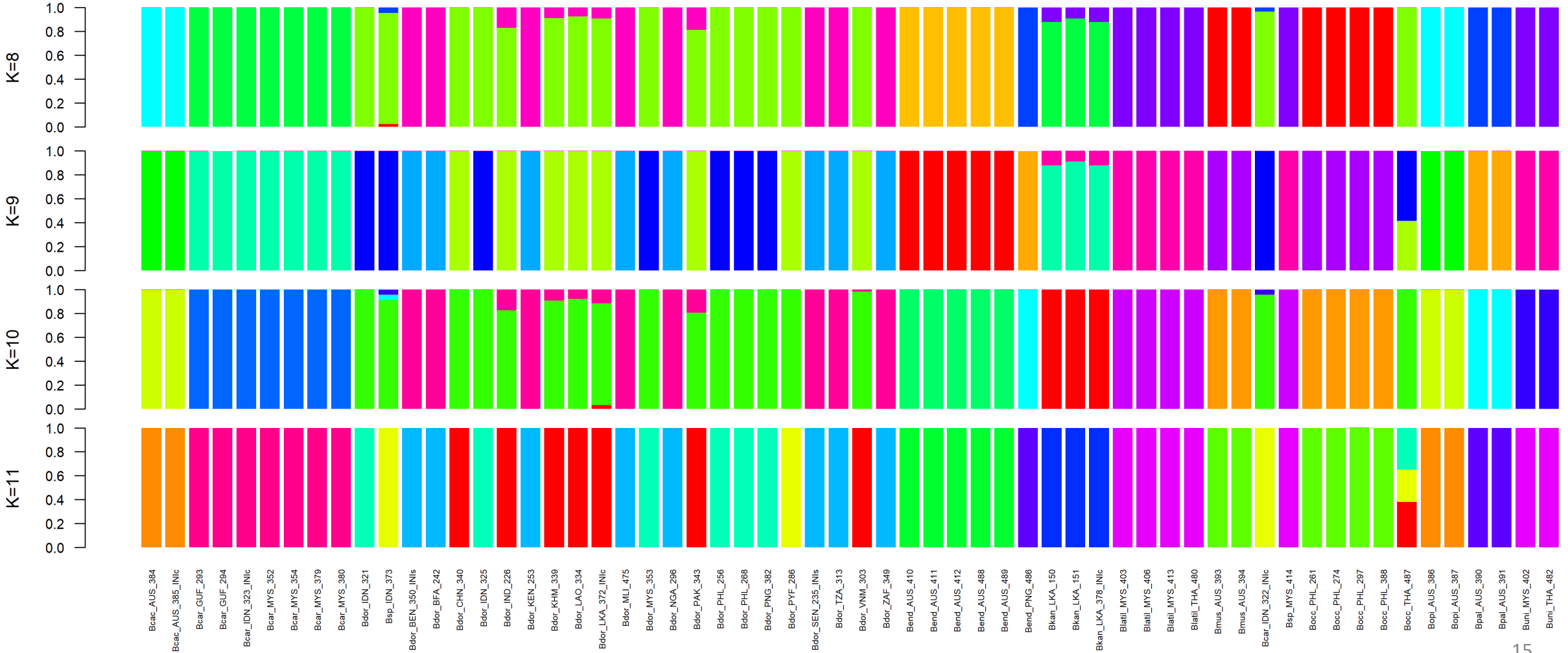
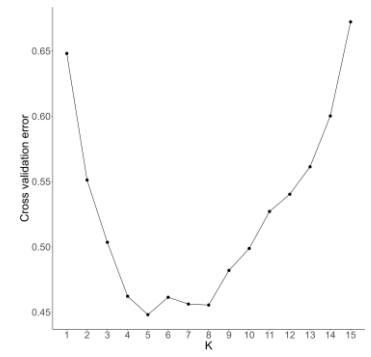
*Bactrocera pallida*

## Les 11 espèces sont discriminées les unes des autres

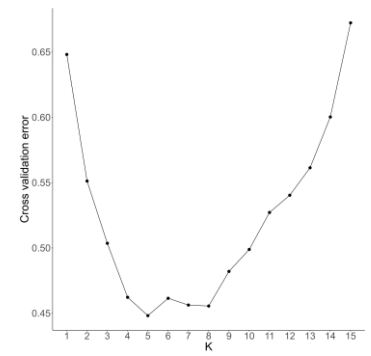


Erreurs d'identification ou introgression nucléaire ?**Bend\_PNG\_486**→ Introgression nucléaire entre *B. endiandrae* et *B. pallida*?

Erreurs d'identification ou introgression nucléaire ?



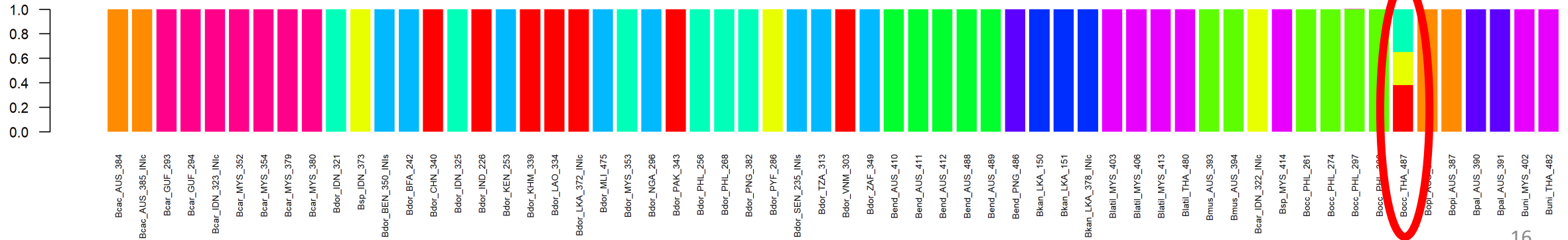
Erreurs d'identification ou introgression nucléaire ?



Bocc\_THA\_487

→ Erreur identification avec *B. dorsalis*?

K=11





➤ **Délimitation des 11 espèces soutenue par l'ensemble des analyses**

- Permet d'**identifier** les spécimens au sein du complexe ou de **repérer les erreurs d'identification**

Bsp\_MYS\_414 / Bocc\_THA\_487 / Bcar\_IDN\_322 / Bsp\_IDN\_373

- ➔ Identification d'un nombre minimal de SNPs diagnostiques des espèces
- ➔ Développement d'un outil moléculaire de diagnostique

- **Pas d'évidence forte d'introgession nucléaire**

Quid Bend\_PNG\_486 ? Bocc\_THA\_487 ?

- ➔ Approche phylogénétique pour étudier la relation entre espèces (arbre en Maximum de vraisemblance + 16 espèces outgroup (22 spécimens))

- Analyse de l'**information mitochondriale**

- ➔ Mitofinder (Allio et al. 2020) + Mapping sur génome mitochondrial de référence de *B. dorsalis*

*Merci de votre attention !*

et merci à ...



CHAPUIS Marie-Pierre  
BIOS-CIRAD



OUVRARD David  
LSV-ANSES



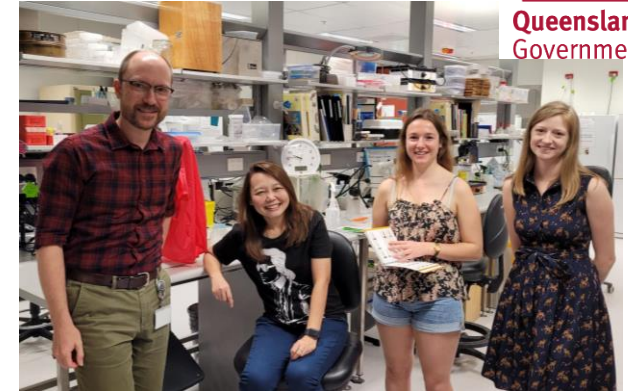
BENOIT Laure  
BIOS-CIRAD



TADDEI Andrea  
LSV-ANSES



MOUTTET Raphaëlle  
LSV-ANSES



SCHUTZE Mark  
QDAF

STARKIE Melissa  
QDAF

