BEeS

The LifeWatch ERIC Biodiversity & Ecosystem eScience Conference

Heraklion, 30 June - 3 July 2025

30 June 2025 | 15:00

BEeS Heraklion
30 June - 3 July 2025

LifeWatch ERIC

hcmr ΕΛΚΕΘΕ

iMBBC ΙΘΑΒΒΥΚ

Session: Taxonomy: Identifying the units of diversity in life

1 July 2025 | 14:30-16:30

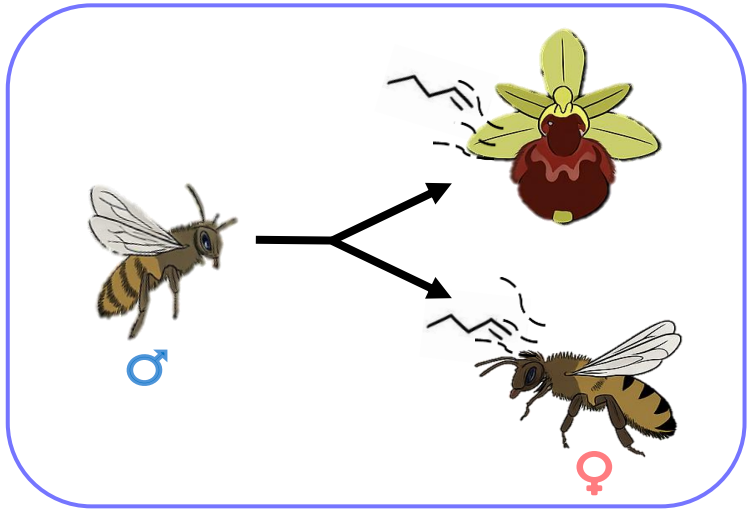
Waking up from a taxonomist's nightmare: *Ophrys* species delimitation meets AI in the phylogenomic era

Presenter: Elisa Addonizio

"Species are groups of actually or potentially interbreeding natural populations, which are reproductively isolated from other such groups"

– Ernst Mayr, 1942, *Systematics and the Origin of Species*

Case study: *Ophrys araneola/sphegodes* group



- Pollination in the orchid genus *Ophrys* works by **sexual deception**

- This genus generates deep controversy over species delimitation; estimates range from 9 to over 350 taxa, often based on subtle differences and presumed pollinator specificity

MAIN AIM

Develop and validate an automatic classification system based on deep learning to distinguish morphologically related species of the genus *Ophrys* from flower images

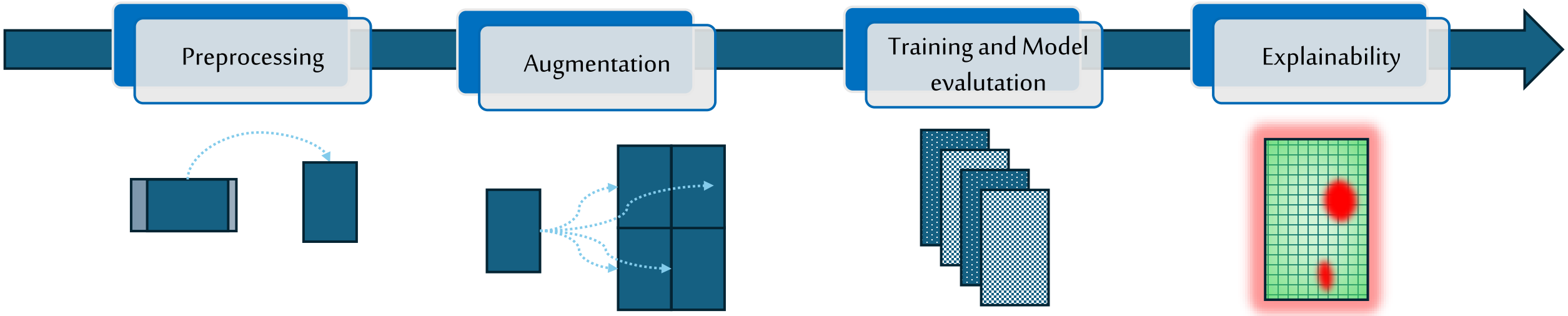
1. Can AI accurately discriminate morphologically similar *Ophrys* species from flower images?
2. How do model architecture, training strategy, and dataset composition influence classification performance?
3. Is it possible to interpret and visualize the morphological features guiding the neural network's decisions?

Materials and methods

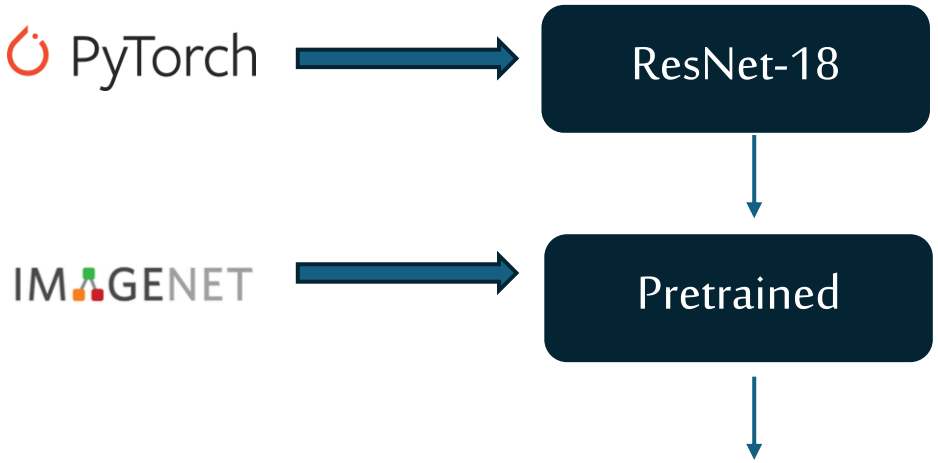
1. Dataset acquisition and image labelling

O. exaltata	*O. garganica*	*O. incubacea*	*O. sphegodes*
			
283 images	251 images	156 images	100 images
218 images	319 images	210 images	218 images

2. Design of a complete classification pipeline



3. Model training



M1 Trained exclusively on the private dataset, excluding *O. majellensis*

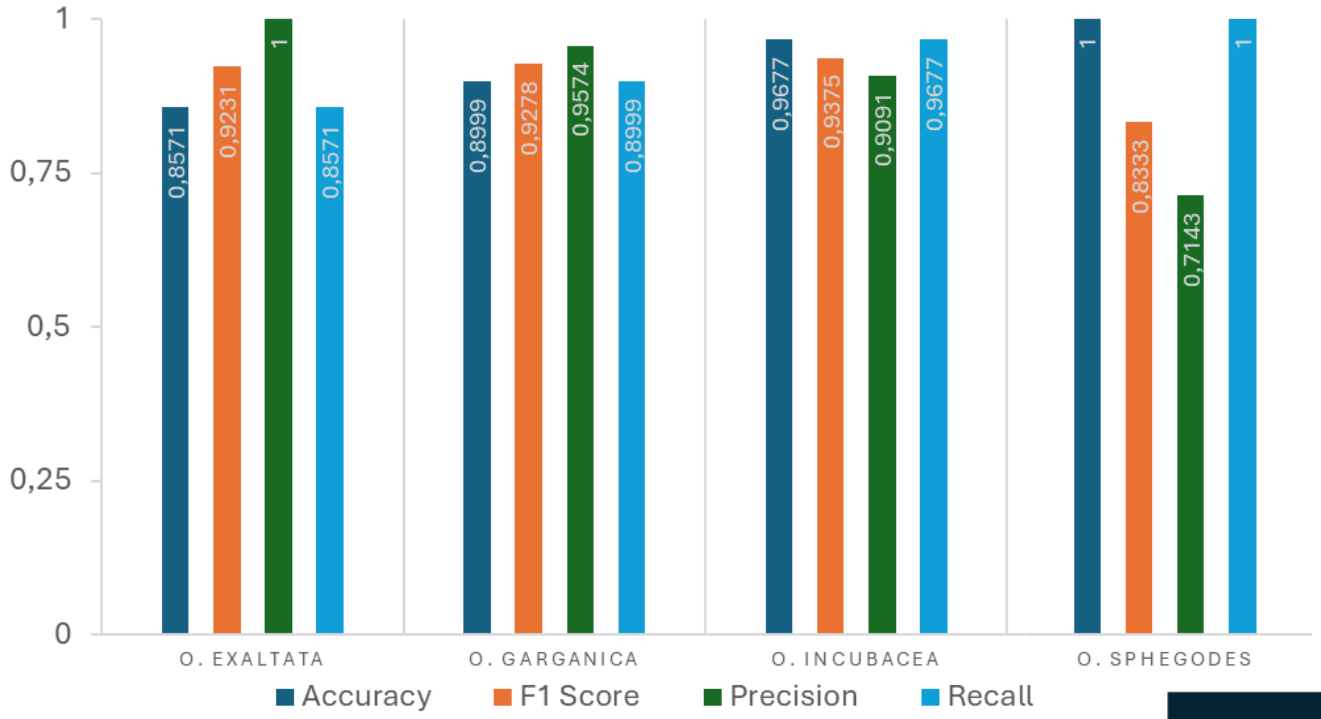
M2 It also includes *O. majellensis*

M3 Extension of training with a mixed dataset (private + images from the internet)

M4 One-vs-All approach, with training on private images only

1 Can AI accurately discriminate morphologically similar *Ophrys* species from flower images?

Metrics for the model (private dataset)



- This model achieves high performance, confirming the effectiveness of the automated morphological approach in distinguishing *Ophrys* species

- The results confirm strong class separation and high accuracy across all categories

Confusion Matrix for model (private dataset)

Label\Predicted	*O. exaltata*	*O. garganica*	*O. incubacea*	*O. sphegodes*
O. exaltata	48	0	0	8
O. garganica	0	45	3	0
O. incubacea	0	1	30	0
O. sphegodes	0	0	0	20

2 How do model architecture, training strategy, and dataset composition influence classification performance?

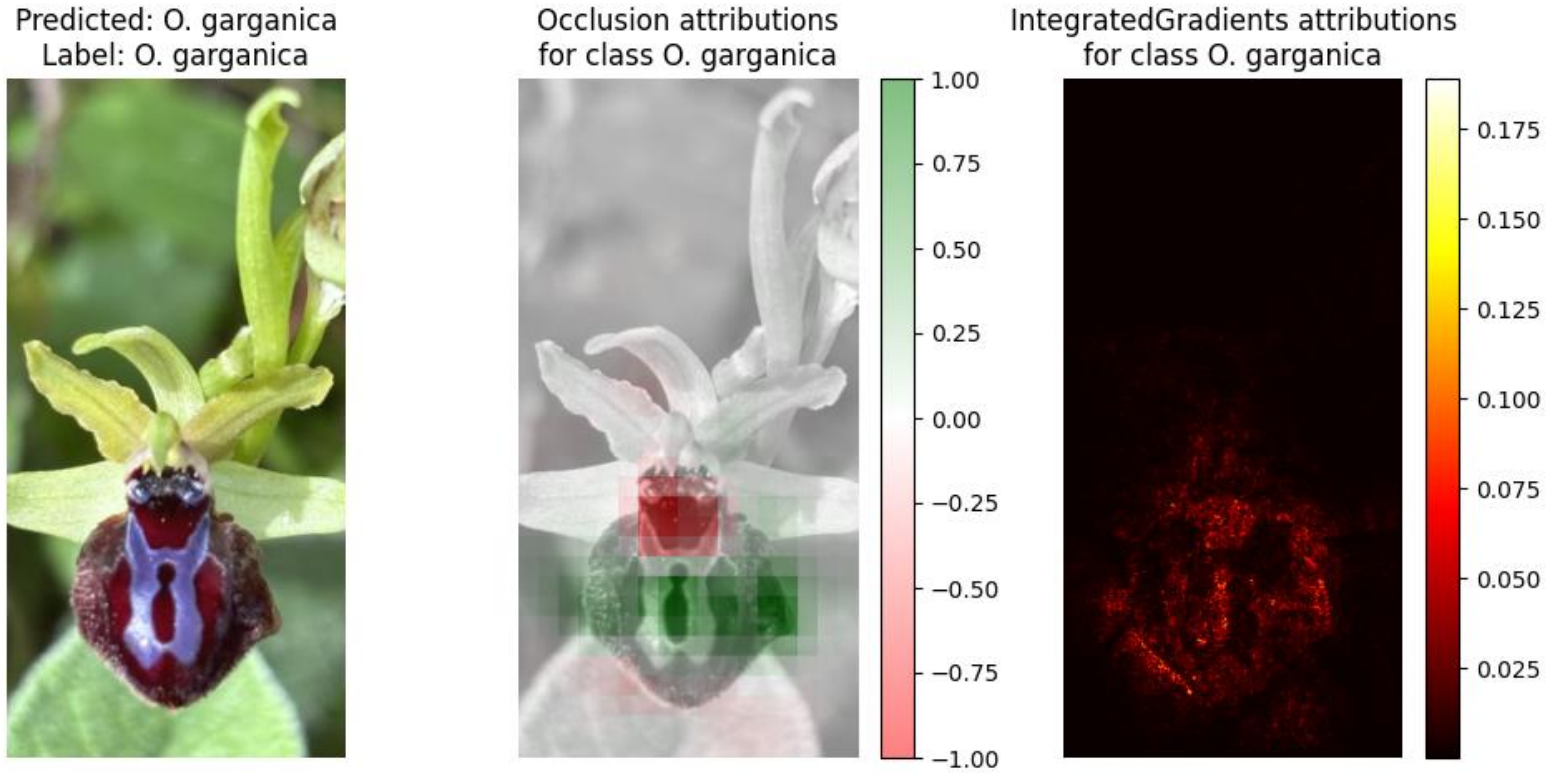
Metrics

Macro and Micro model comparison

➢ All models show high performance; expanding to 6 classes and adding online data increases variability and generalization, while One-vs-All allows detection of unknown classes

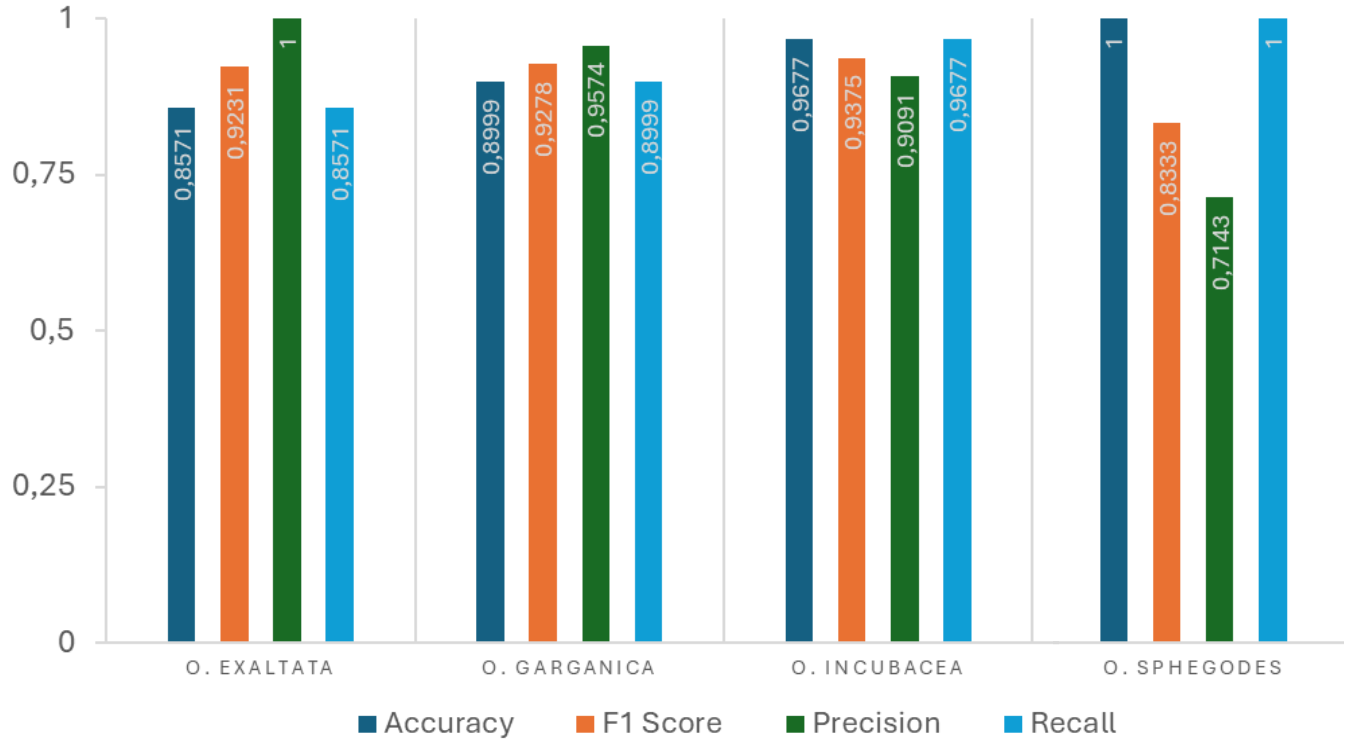
3 Is it possible to interpret and visualize the morphological features guiding the neural network's decisions?

Explainability



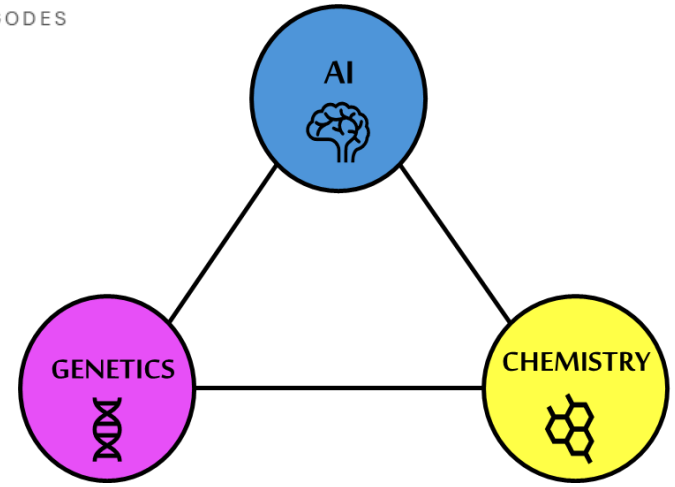
➢ Explainability techniques such as Occlusion and Integrated Gradients reveal that, across all analysed images, the most informative regions for classification consistently correspond to the labellum, confirming that the model relies on key morphological traits consistent with the biology of the *Ophrys* genus

The AI model successfully distinguished different *Ophrys* species, with several F1-scores above 0.93



Does the model's high classification accuracy reflect underlying

genetic and **chemical** differences?



Case study application: *O. sphegodes* and *O. majellensis* in Majella National Park

Ophrys sphegodes

❖ Flowering period: **April**

Ophrys majellensis

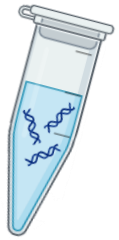
❖ Flowering period: **June**

They share the lawn, but flower at different times

GENETICS

METHODS

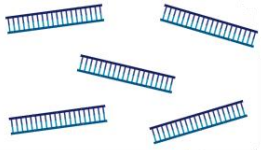
DNA extraction



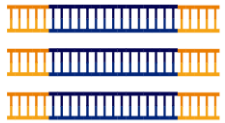
Digestion by

Double Restriction enzymes

ddRADSeq

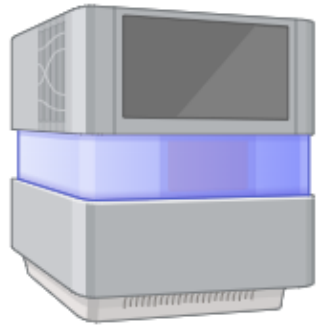


DNA fragments

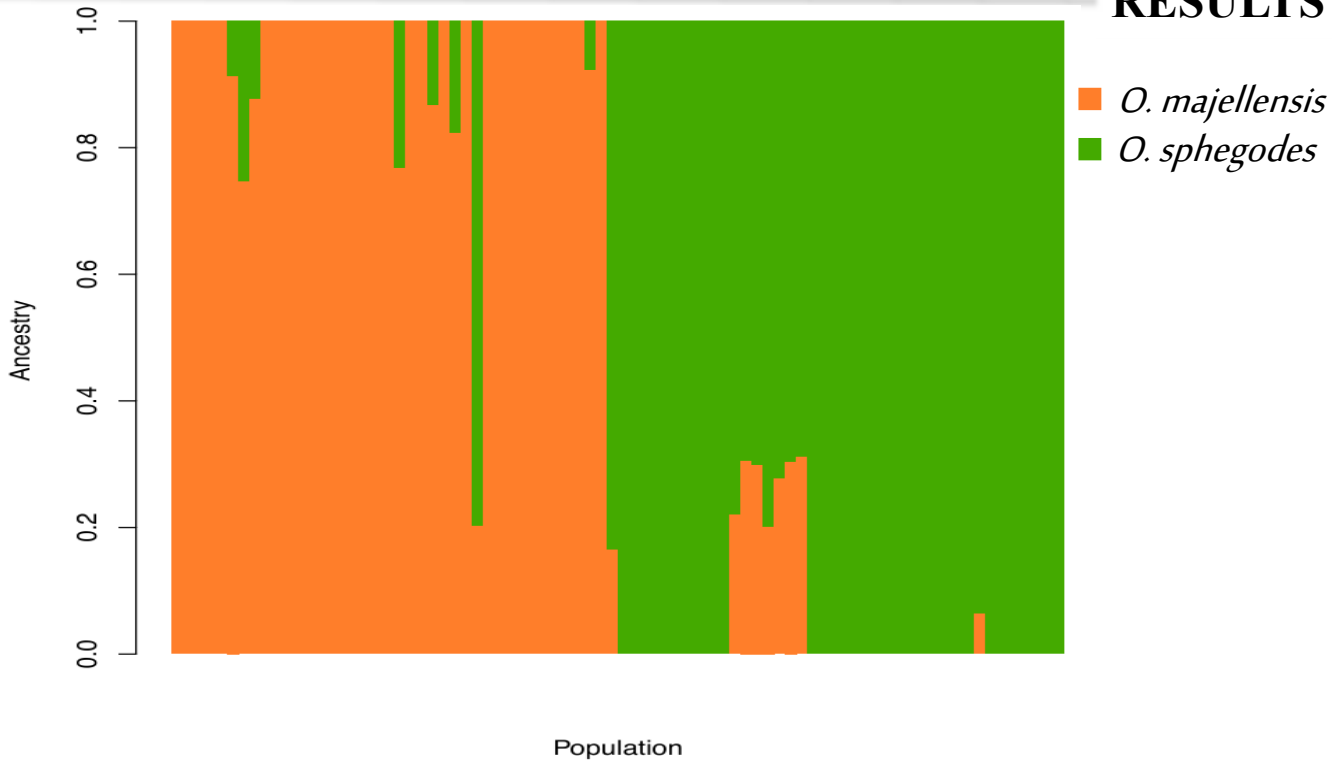


Adapter ligation

Sequencing



RESULTS

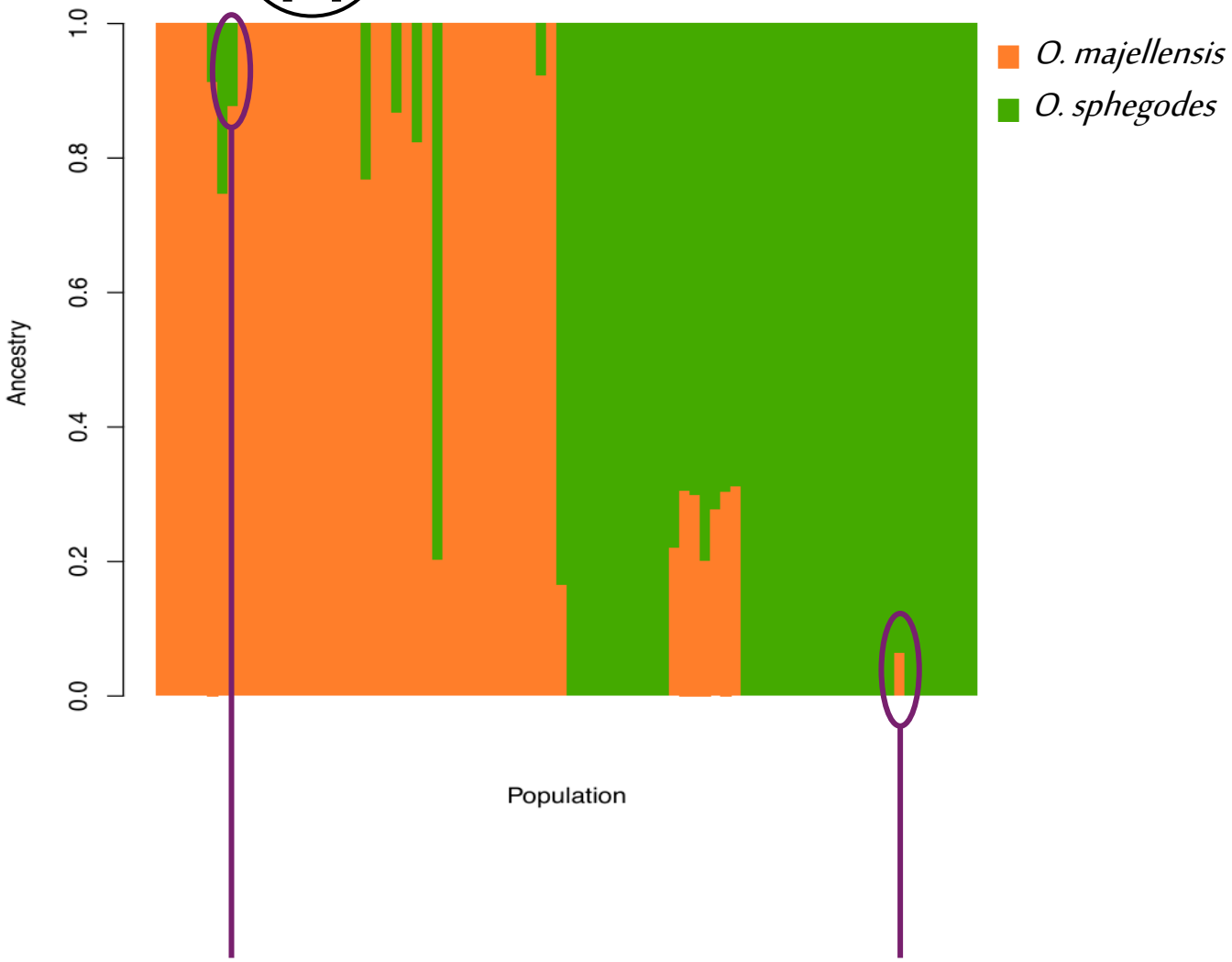


O. majellensis

O. sphegodes

Some individuals show intermediate genetic assignment: in these cases, what do the morphological and chemical data indicate?

GENETICS

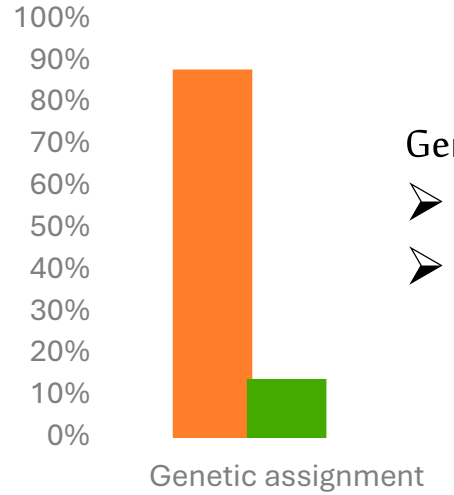


INDIVIDUAL: «***DM32***»

Ophrys majellensis

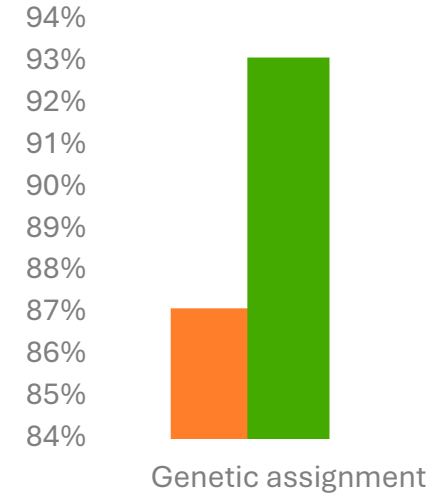
INDIVIDUAL: «***SC31***»

Ophrys sphegodes



Genetic assignment for **DM32**:

- 87% *O. mejellensis*
- 13% *O. sphegodes*

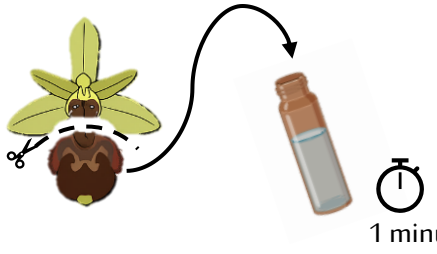


Genetic assignment for **SC31**:

- 93% *O. mejellensis*
- 7% *O. sphegodes*

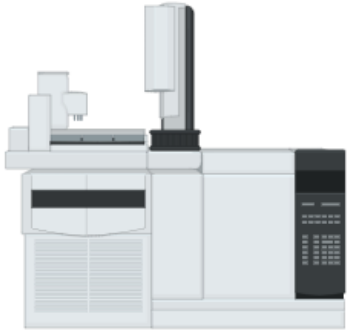
CHEMISTRY

METHODS

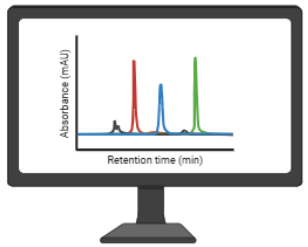


1 minute in agitation

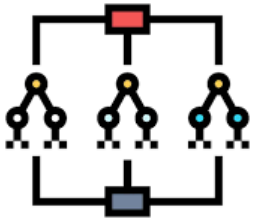
Sampling of labella in hexane



GC-MS analysis

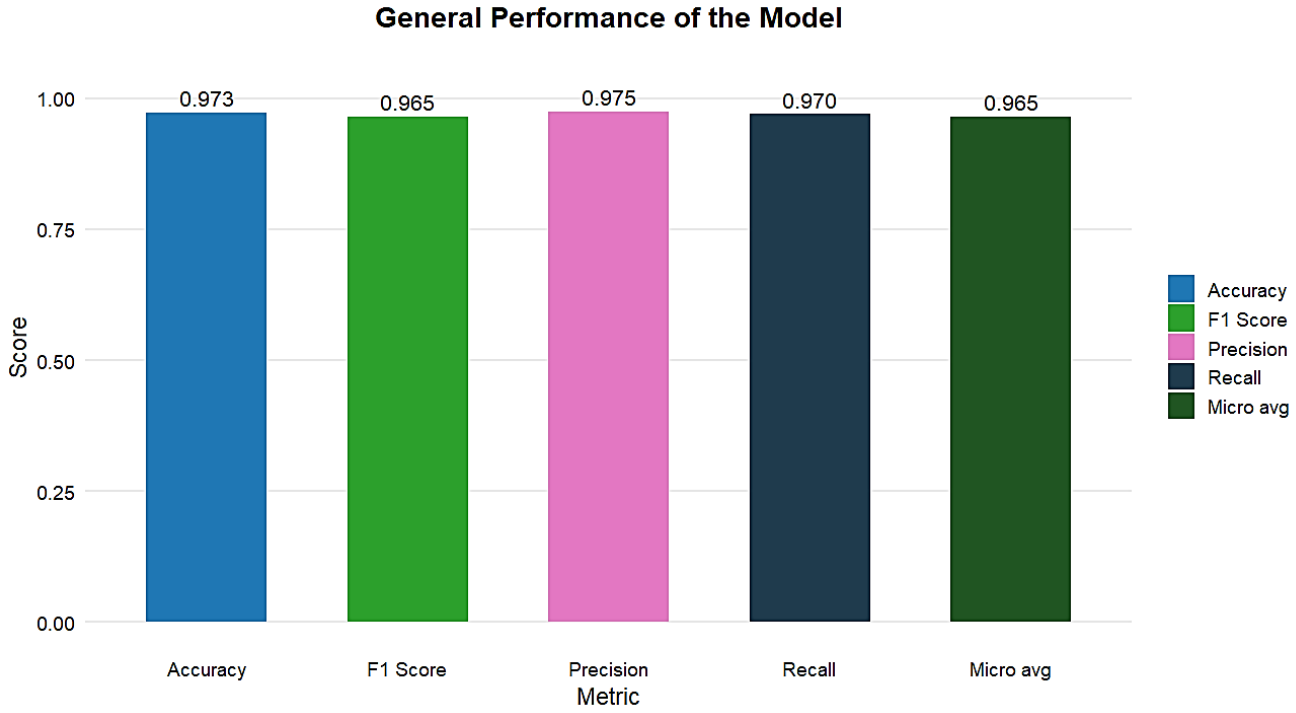


Chromatogram analysis and peak assignment



Species classification using VOCs and Random Forest

RESULTS

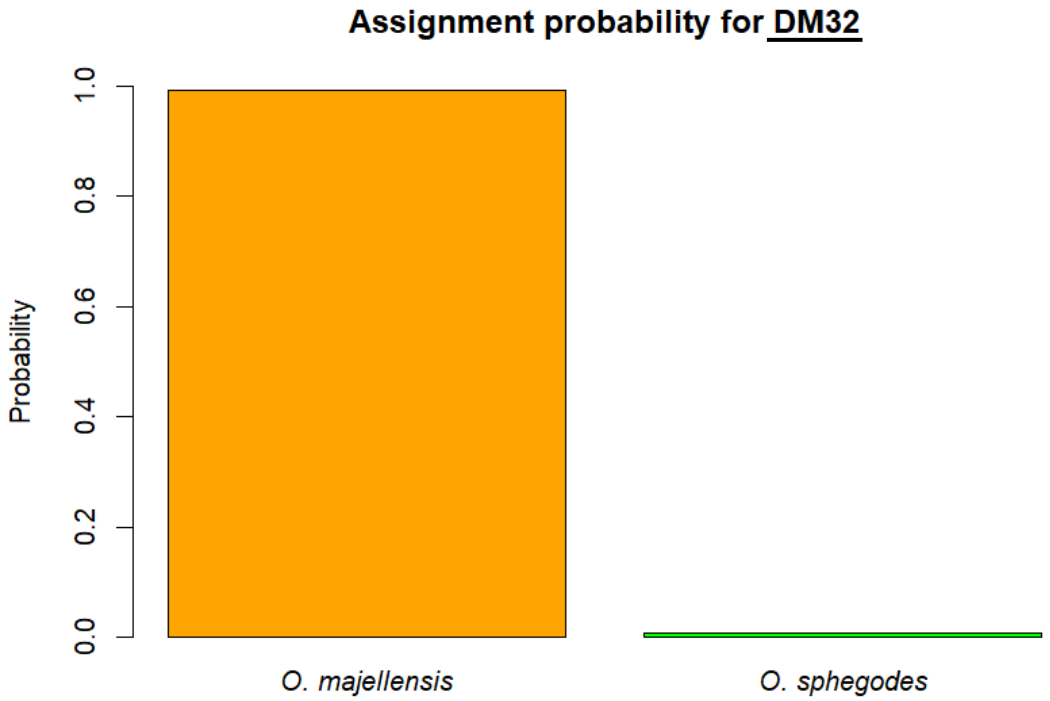


Label\Predicted	*O. majellensis*	*O. sphegodes*
O. majellensis	38	2
O. sphegodes	0	33

VOC-based Random Forest classification supports a distinct chemical profile for *O. majellensis* and *O. sphegodes*

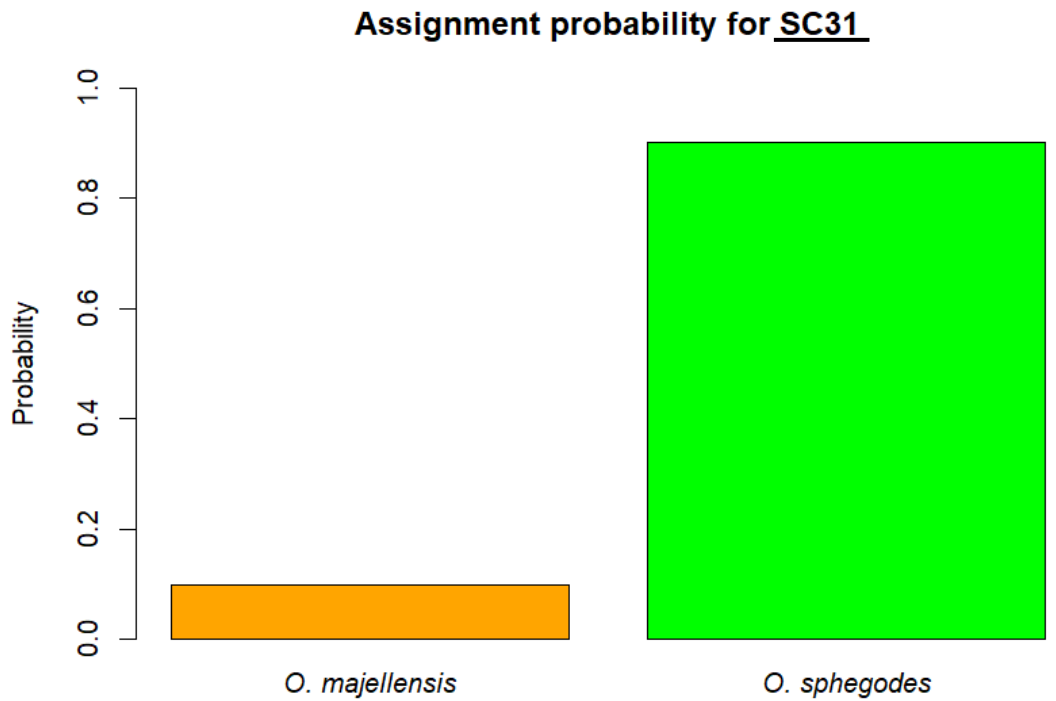
CHEMISTRY

Species assignment based on volatile compound profiles



O.majellensis: 99.3%

O.sphegodes: 0.007



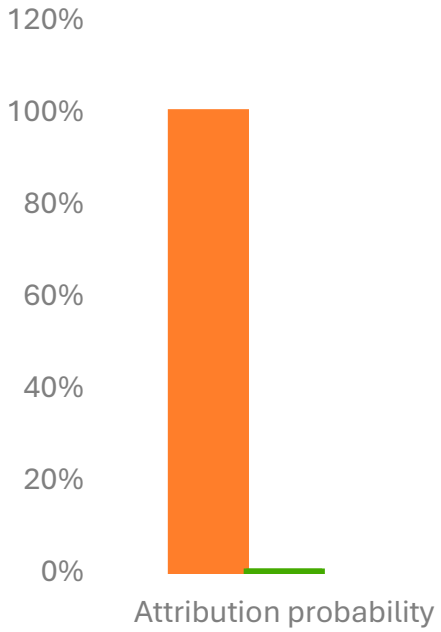
O.majellensis: 9.9%

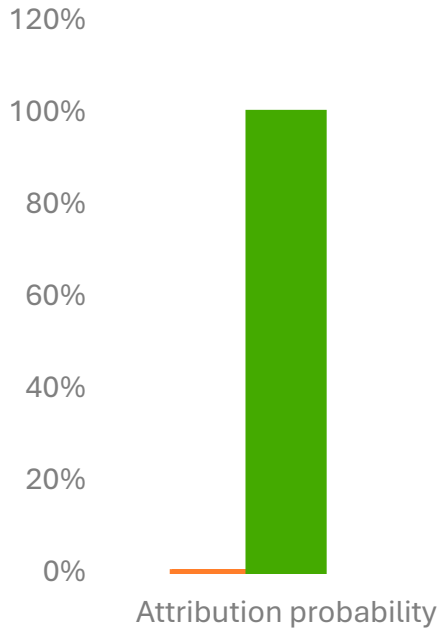
O.sphegodes: 90.1%

The percentages represent classification probabilities generated by a Random Forest model

ARTIFICIAL INTELLIGENCE

Species attribution obtained through automatic classification of the flower image using a convolutional neural network





Predicted class probability for «**DM32**»:

O. majellensis: **99.98%**

O. sphegodes: 0.02%

Predicted class probability for«**SC31**»:

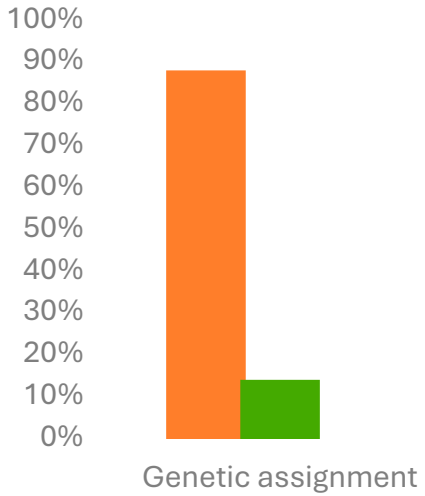
O. majellensis:0.01%

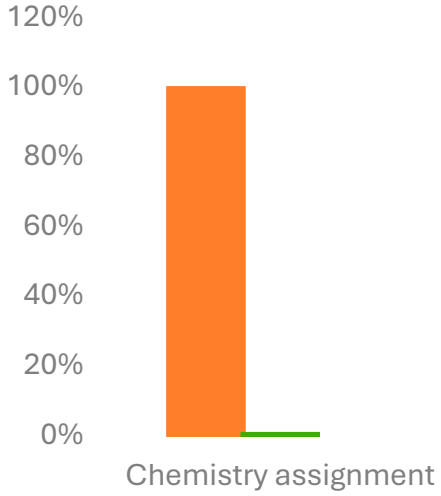
O. sphegodes: **99.99%**

The percentages represent the probabilities of assignment predicted by the model for each species

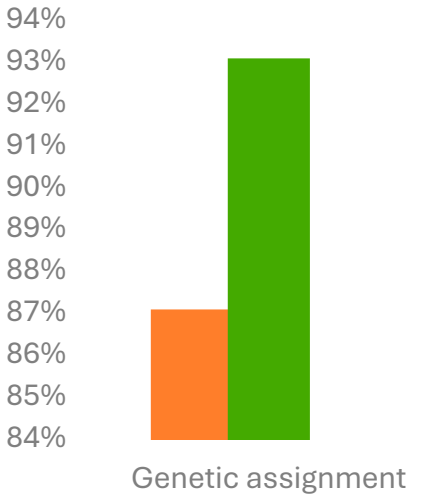
CONCLUSION

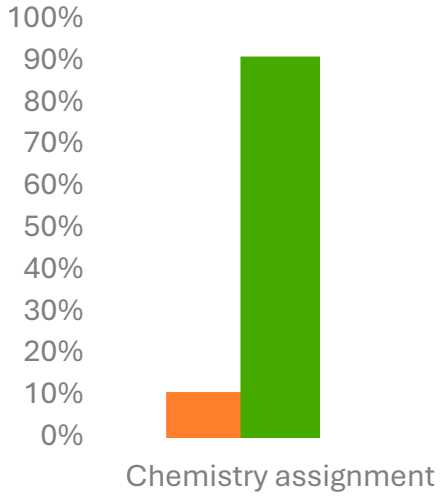
DM32





SC31

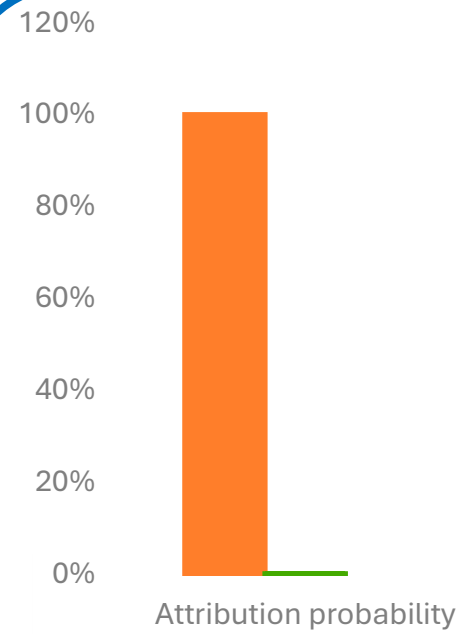


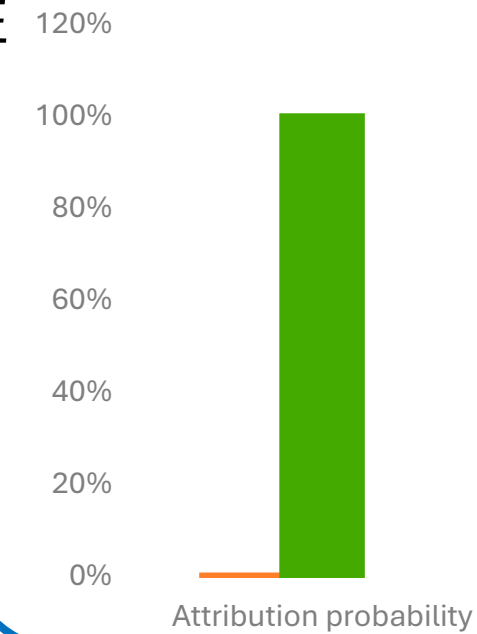


GENETIC

CHEMISTRY

ARTIFICIAL INTELLIGENCE





CONCLUSION

- The results highlight how AI techniques achieve high accuracy rates

- Samples with less defined genetic assignments show high assignment with chemical data and AI

- AI-based classification system is able to capture complex morphological and phenotypic nuances that escape genetic analysis alone, thus offering a promising and more comprehensive tool for the classification of species in *Ophrys*

Thank you!

Questions?
elisa.addonizio@unina.it

BEeS
The LifeWatch ERIC Biodiversity & Ecosystem eScience Conference

Heraklion, 30 June - 3 July 2025