



Landrace gap analysis Introduction

November 10, 2025

Alliance Bioversity International & CIAT

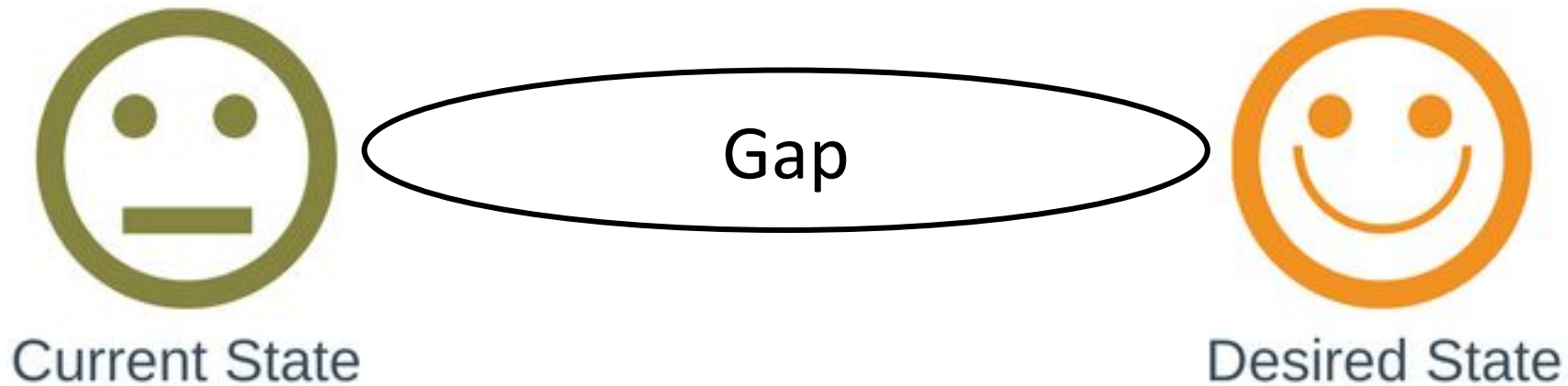
Our goal is to answer these questions

- How well does my collection of accessions represent the region's biodiversity?
- Where can I find landraces, and which places or areas should I prioritize for collecting them within my region?

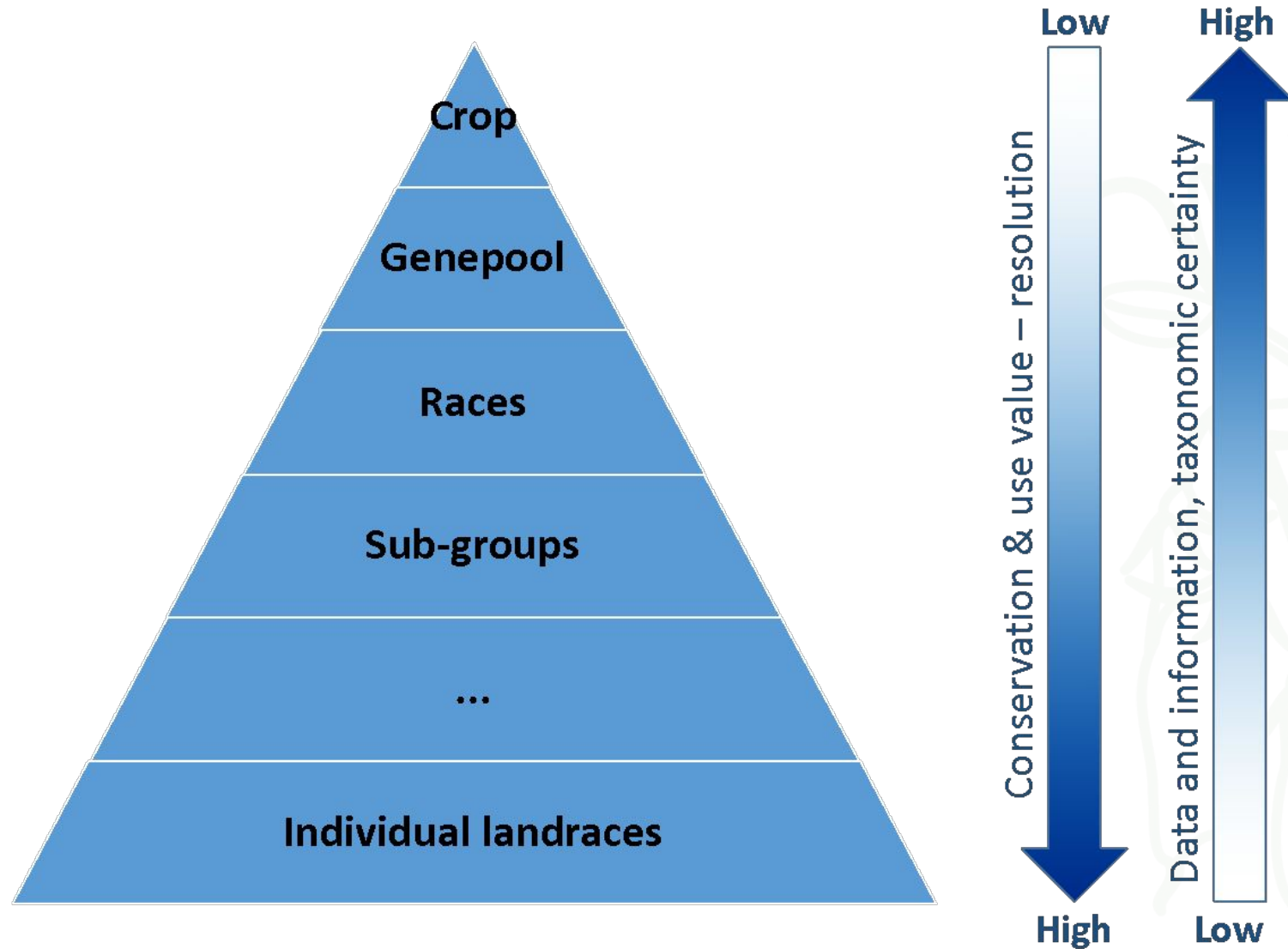
What is a gap?

Something missing

This "something" typically takes us from an undesired current state to a desired future state.



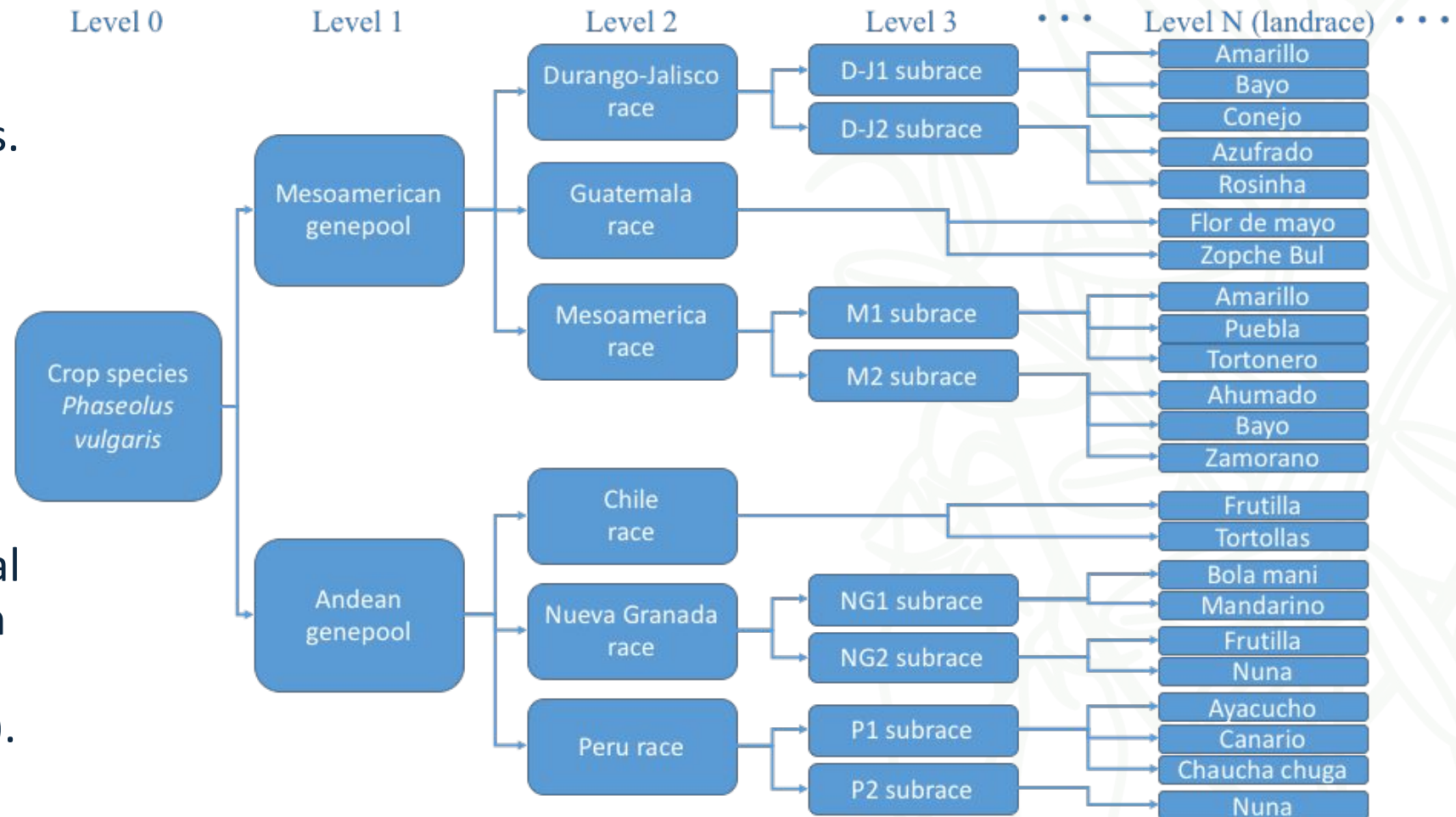
First, we would like to understand the structure of the landrace accession using what is known about its diversity.



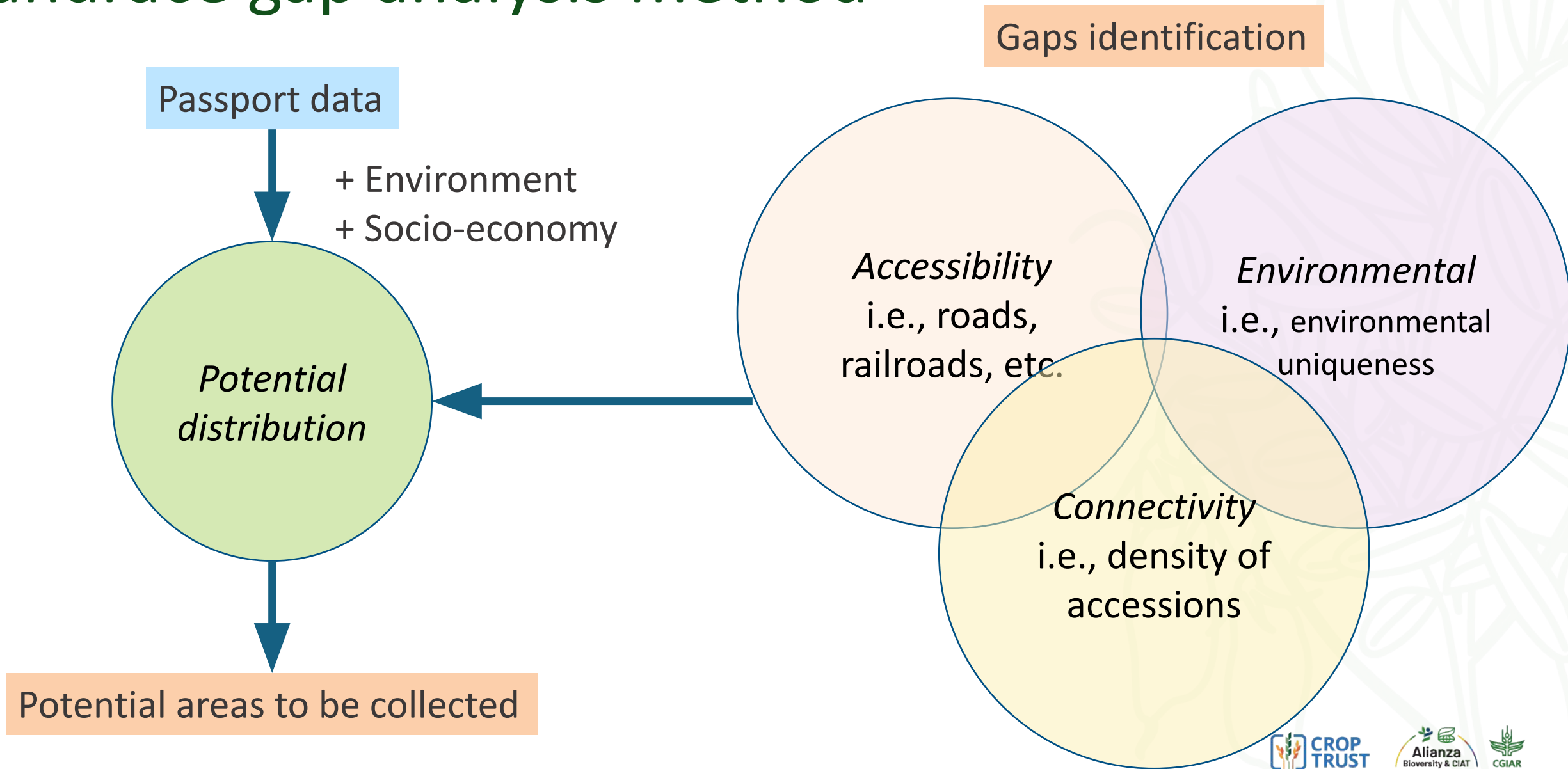
Determine landrace group

Why? Geographic distribution of landraces is influenced by anthropogenic factors

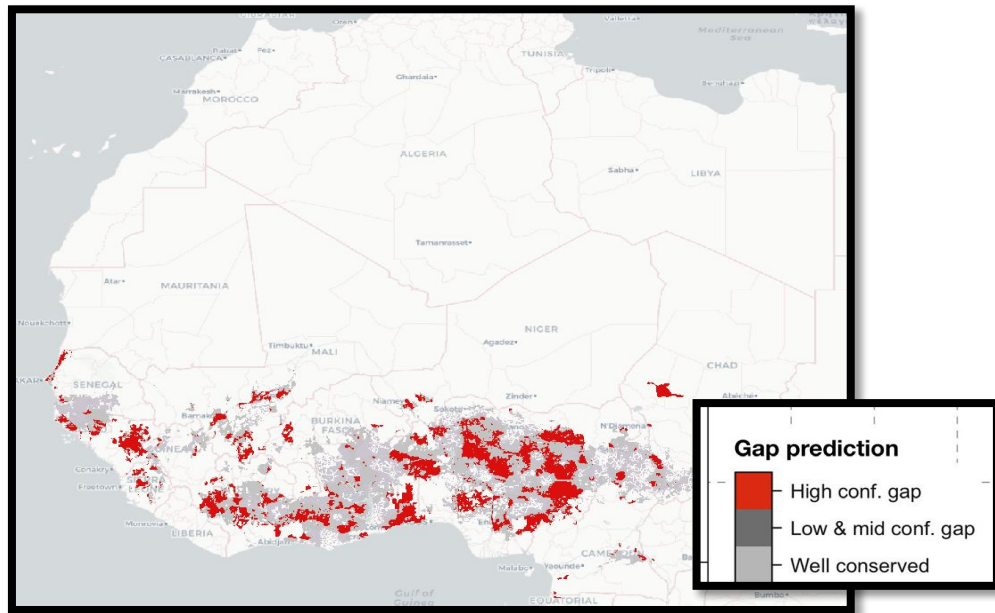
- Literature review, passport data used to group landraces.
- Groups are tested for their association with spatial predictors. The goal is to identify a set of "good" spatial predictors.
- The groups relevant to spatial analysis varied depending on the crop (e.g., taxonomic, morphological, or genotypic).



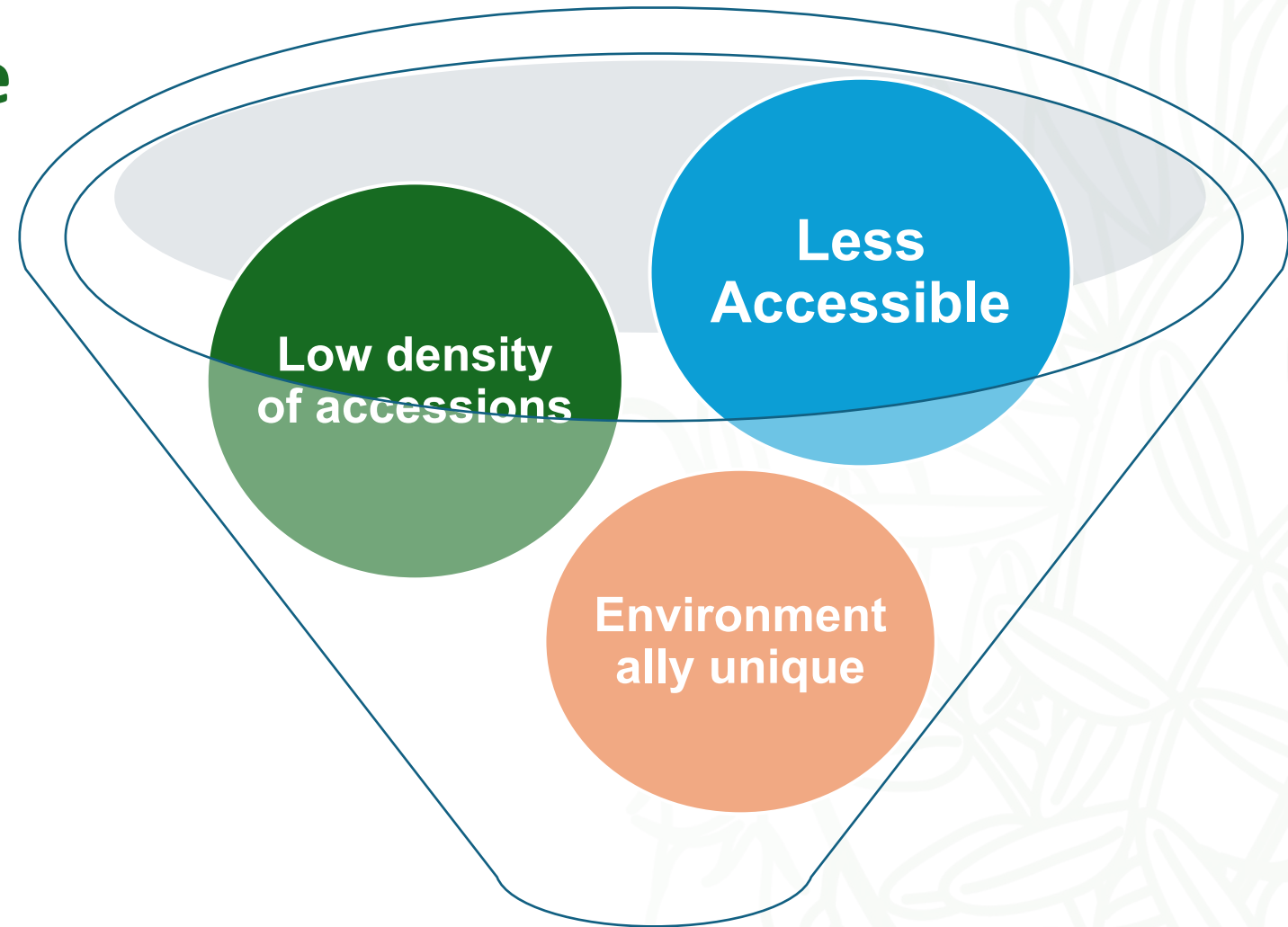
Landrace gap analysis method



We would like to determine where the gaps are in their species' distribution.

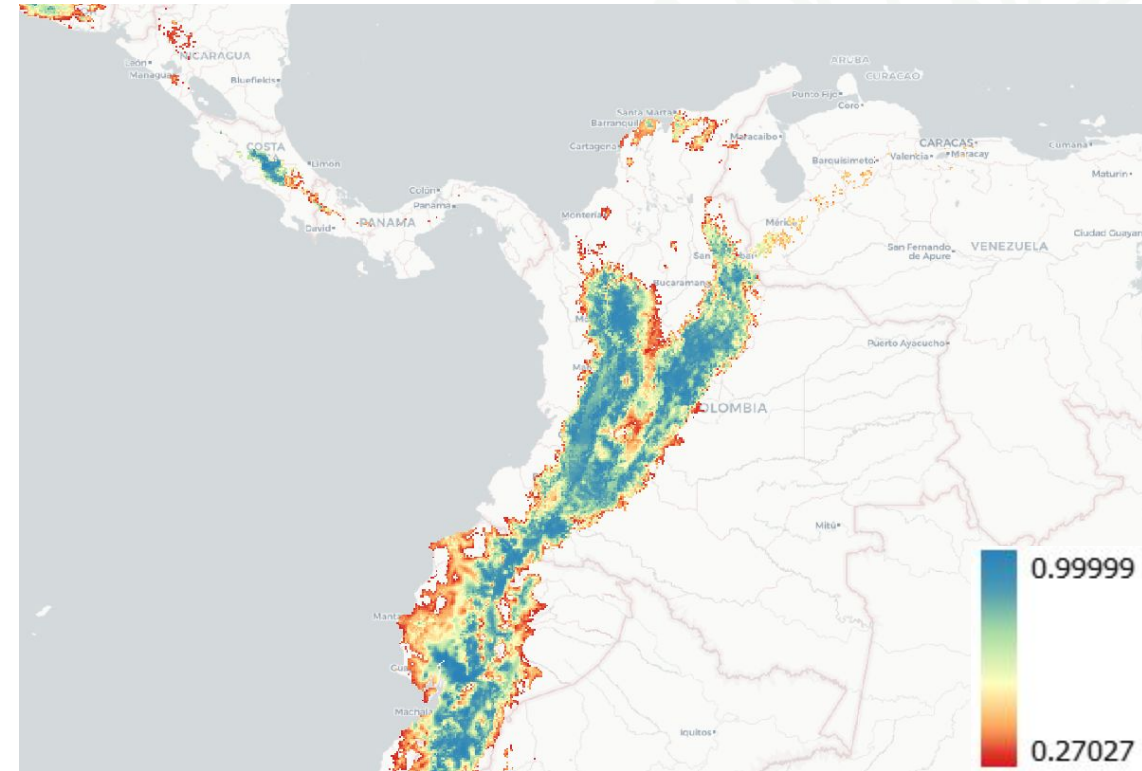
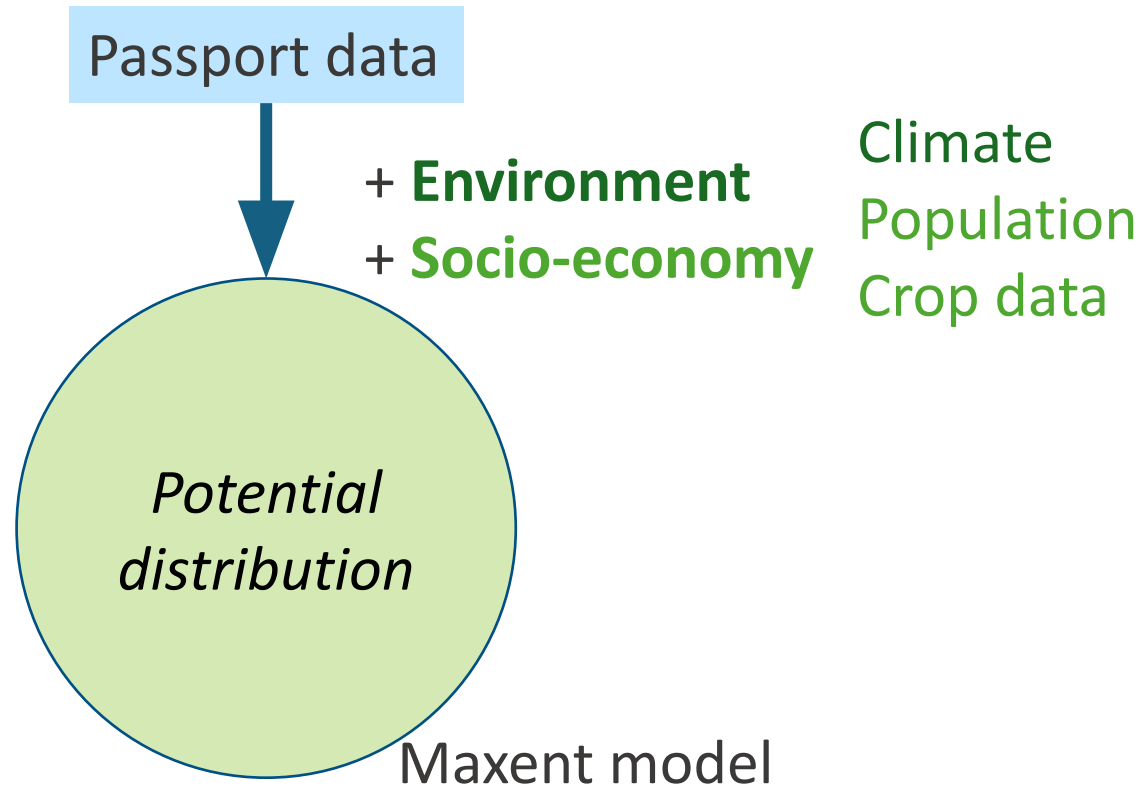


High confidence gaps occur when all three criteria are met.



High Confidence gap

Potential distribution



0: Low chance to find a landrace presence
1: High chance to find a landrace presence

* https://biodiversityinformatics.amnh.org/open_source/maxent/

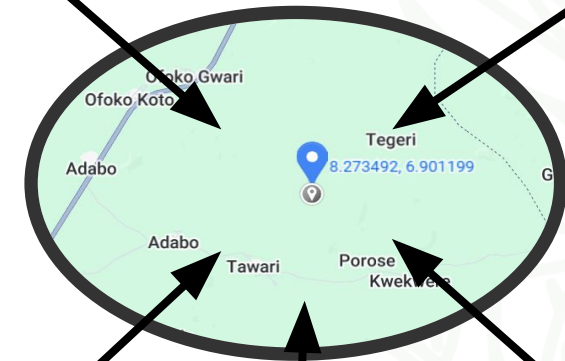
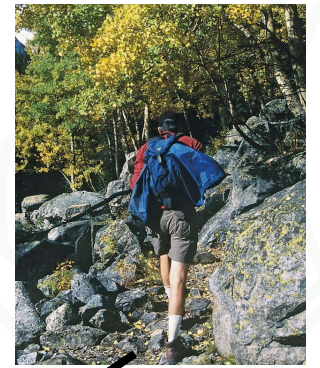
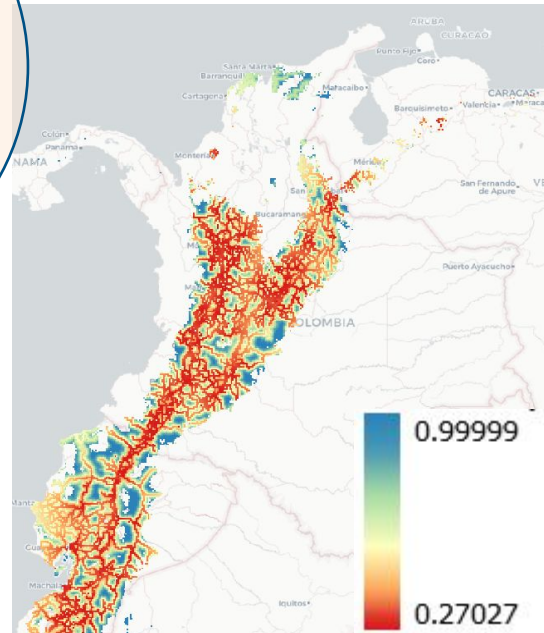
Accessibility gap score

+ Friction surface
+ Potential distribution

Passport data

Accessibility
i.e., roads,
terrain type,
etc.

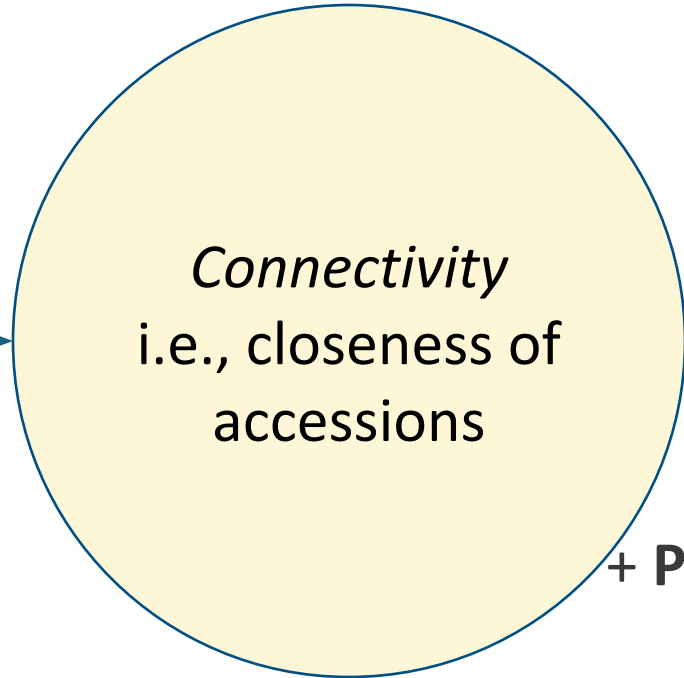
How accessible a location is, based on
the terrain type ?



0: The lower the chance to find an uncollected landrace
1: The higher the chance to find an uncollected landrace

Connectivity gap score

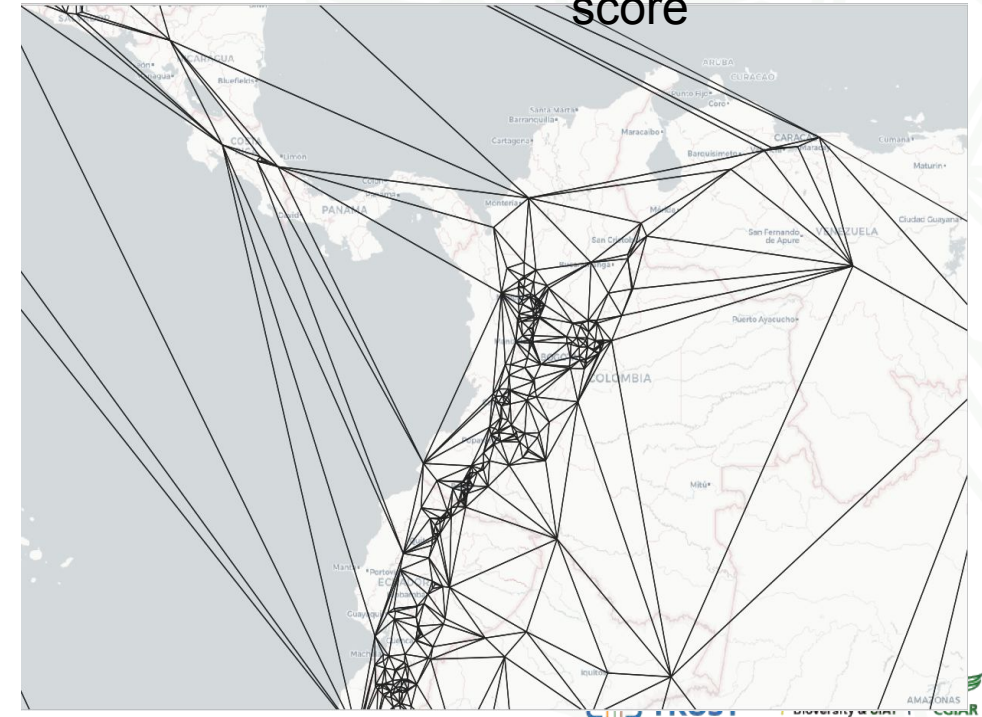
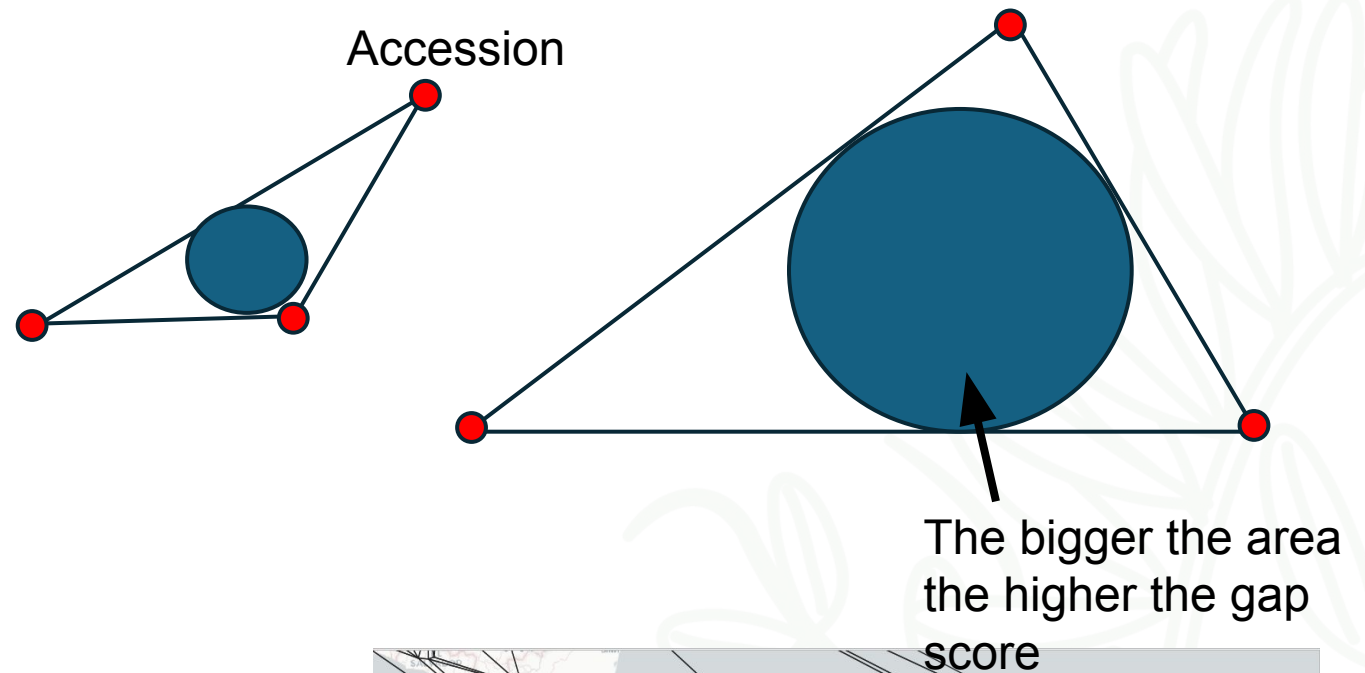
Passport data



Connectivity
i.e., closeness of
accessions

+ **Potential distribution**

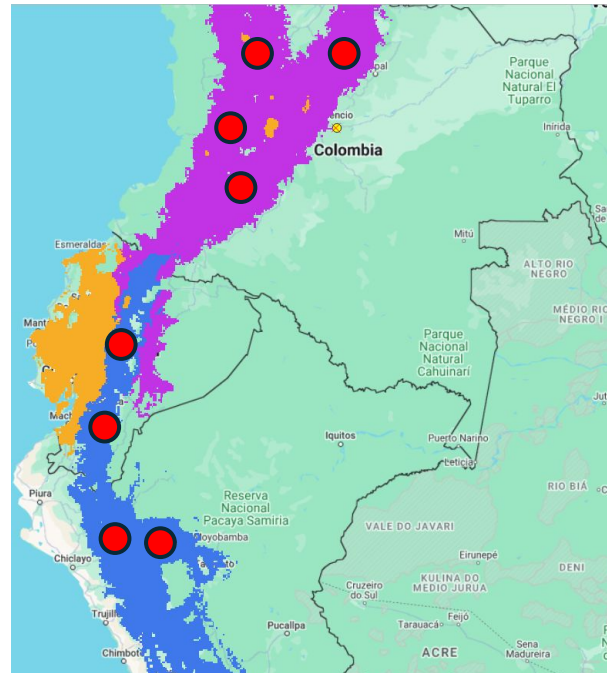
Method to identify high density areas
where accessions are already collected
(low chance to find anything new)



Environmental gap score

Passport data + Potential distribution
+ Environmental data

Environmental
i.e., environmental
uniqueness

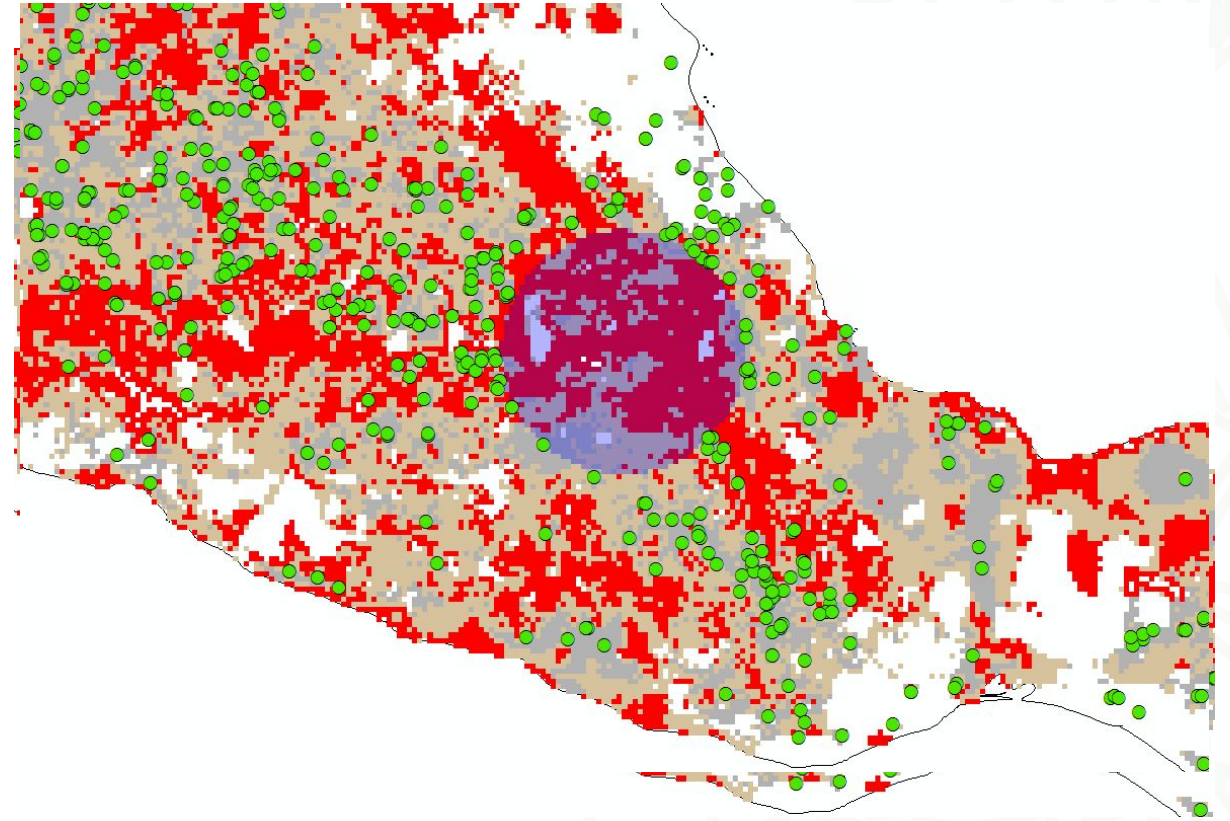
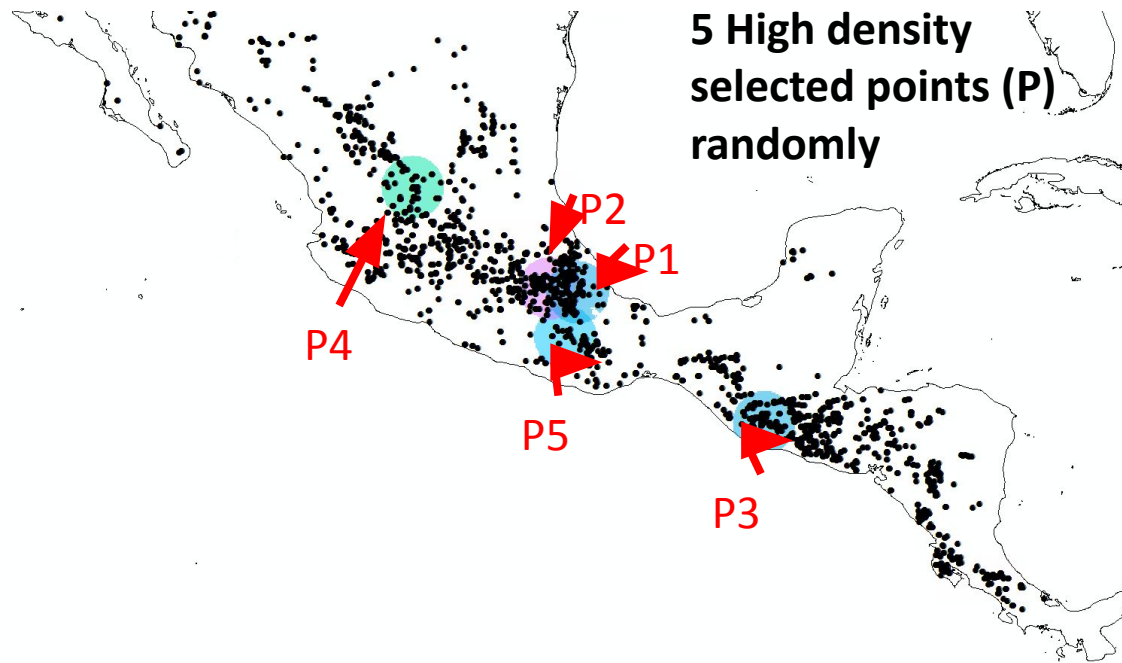


Env. Cluster	Temp. avg (C°)	Prec. avg (mm)
CLUSTER 1	18	25
CLUSTER 2	13	11
CLUSTER 3	25	5

Cluster 1 and **Cluster 2** are well represented but **Cluster 3** it is not, and it has some environmental unique conditions.

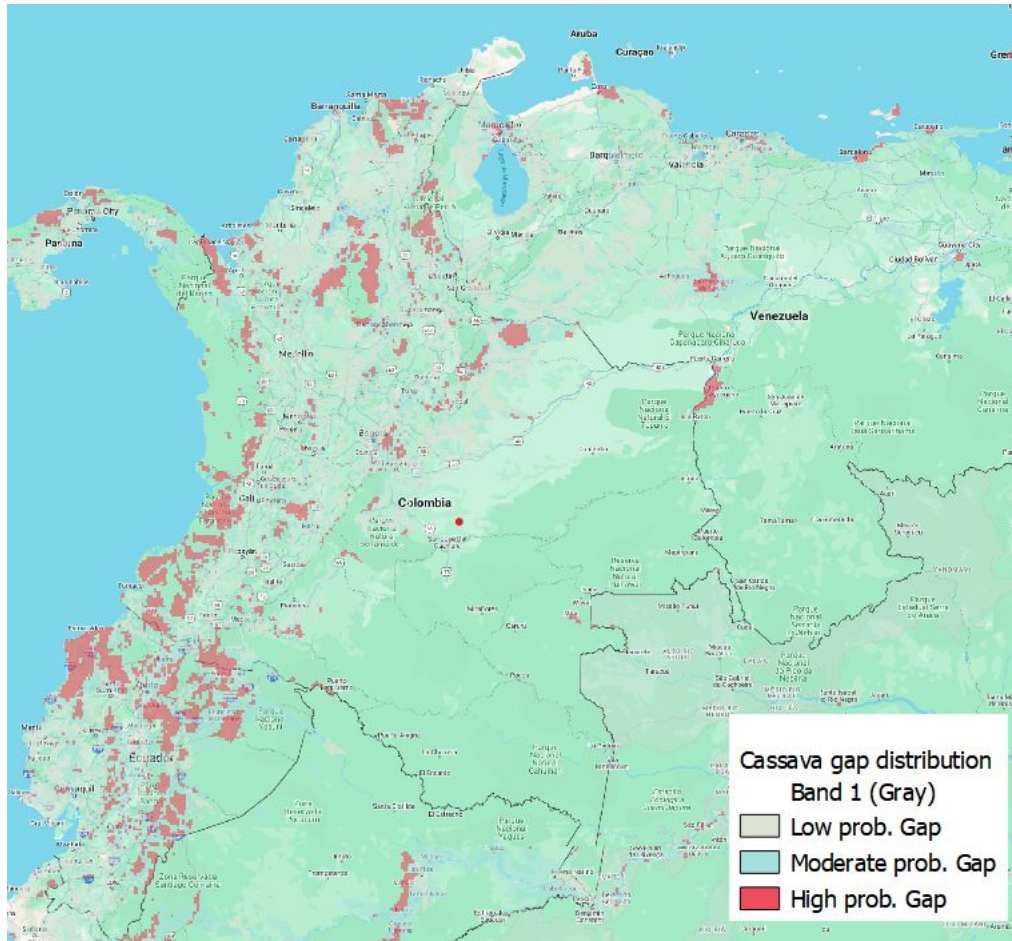
Identify geographical areas with low environmental representativeness in the collection

Gap validation



By introducing synthetic gaps, the validation process allow us to know the prediction capacity of the gap.

What insights can we gain from the results?



Metrics	Value
Coverage	68%
Total Gap Area	23.000 Km ²
Average Elevation	635 Mts

- **Coverage:** The germplasm collection has covered 68% of the species distribution. This means that 32% remains missing. Is this good or not?
- The total area marked as a gap is **23,000 km²** (which is impressive considering the entire country is more than 1,142,000 km²).
- On average, the gap areas are at an elevation of **635 meters above sea level**. Is this elevation suitable for growing crops?

Considerations

- The final gap maps do not filter out protected areas.
- The final gap maps are not filtered by land use coverage.
- The spatial model did not use the soil property variables (a possible enhancement).
- Conduct an expert inspection of the results to further refine the gap maps.



Questions?