



Les Rencontres du  
**Végétal**

**9<sup>e</sup>**  
édition

**2017**  
**16 & 17 janvier**

**AGROCAMPUS OUEST**  
ANGERS, FRANCE

**RECHERCHE**  
**EXPÉRIMENTATION**  
**INNOVATION**

Fruits  
Légumes  
Ornement  
Plantes aromatiques  
et médicinales  
Semences  
Cidriculture  
Viticulture  
Paysage

*sensométrie*  
*chimiométrie*



## Caractérisation variétale de graines de légumineuses par Imagerie Hyperspectrale Proche Infrarouge

**Session : Diversité au sein des systèmes de  
Production**

**Benoît JAILLAIS**

Ingénieur de Recherches, StatSC INRA/ONIRIS

**Eloïse Lancelot**

Ingénieur de Recherches, Gepea, ONIRIS

**Dominique Bertrand**

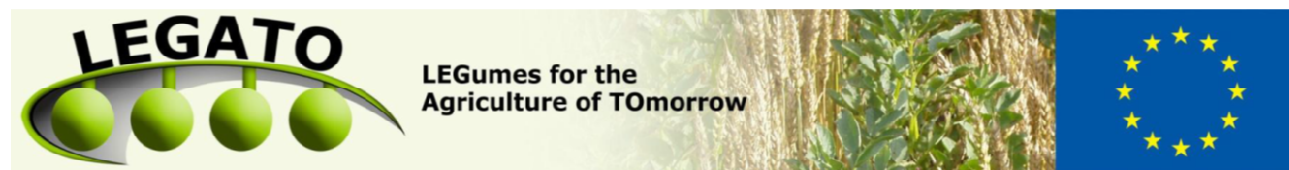
data\_frame, Nantes



**INRA**  
SCIENCE & IMPACT

**Oniris**  
École Nationale  
Vétérinaire, Agroalimentaire et de l'Alimentation  
Nantes Atlantique

# LEGATO Project



LEGATO (LEGumes for the Agriculture of TOMorrow) est un projet de 7,000,000 € pour une durée de 4 ans et financé par l'Union Européenne\*. Il a pour but d'accroître la part de protéines végétales dans le régime alimentaire et donc de développer la culture des légumineuses en Europe.

Parmi les légumineuses étudiées (pois, féveroles, lentilles, pois chiche), les recherches entreprises dans le projet sont focalisées sur les deux espèces majeures (pois et féveroles) et portent sur l'identification et l'évaluation de nouvelles lignées ayant des apports nutritionnels pour le consommateur, et sur le développement d'outils rapides de screening basés sur des modèles spectroscopiques.

L'imagerie hyperspectrale proche infrarouge (NIR-HSI) a été choisie car elle combine simultanément les avantages de la spectroscopie proche infrarouge (simple d'utilisation, précision, reproductibilité, analyse multiparamétrique) et ceux de l'imagerie, et fournit une information spectrale et spatiale pour un échantillon.

\* "The research leading to these results has received funding from the European Community's Seventh Framework Programme (FP7/ 2007-2013) under the grant agreement n°FP7-613551, LEGATO project"

1 January 2014 – 31 December 2017

# Echantillons

## Graines de légumineuses broyées:

ESPECES	Nombre	Code (Nom)
féveroles	92	VF801-VF892
Pois	105	PS201-PS305
Pois chiche	86	CA601-CA731

**Différents pays :** Espagne, Portugal, Bulgarie, France...

**Différentes sous-espèces :** major, equina Pers, paucijuga Beck

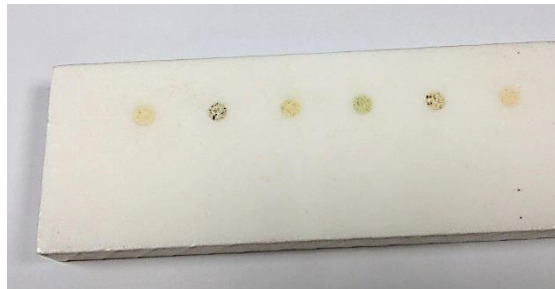
**Différents pays :** Inde, Australie, République Tchèque...

**Différentes conditions de culture:** Landraces, Cultivar, Breeding lines, Wild...

**Pas d'information**

# Matériel et méthode

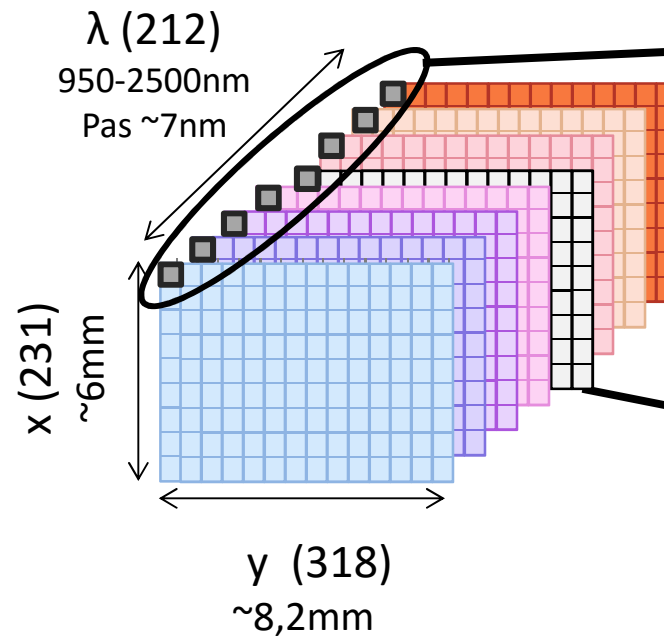
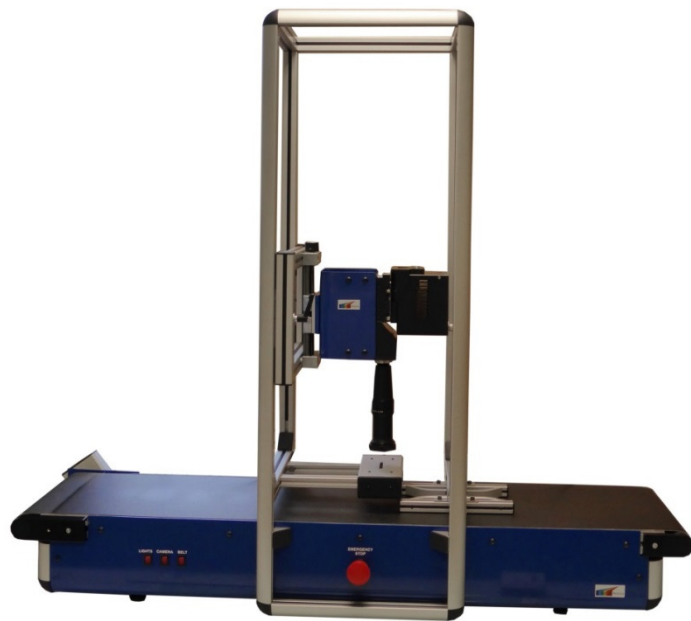
Les broyats de légumineuses sont placés dans des trous d'un porte-échantillons en téflon, usiné au laboratoire.



La surface de chaque poudre est automatiquement imagée par le système NIR-HSI.

# Système d'imagerie hyperspectrale NIR

## HyperPro (BurgerMetrics)



Spectre à un pixel donné xy

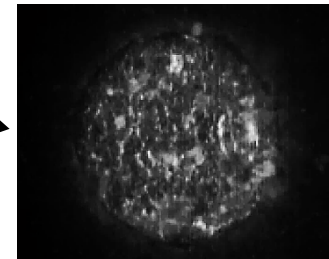
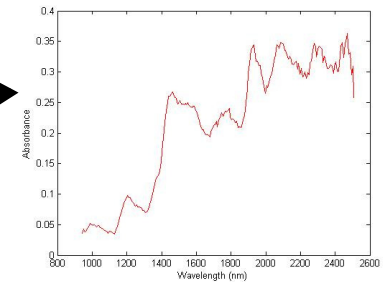


Image obtenue à  
une longueur  
d'onde  
sélectionnée

**1min20s/image**

Système d'imagerie hyperspectrale "Pushbroom" :

- Acquisition d'une ligne entière (résolution spatiale 26 $\mu$ m )
- Le convoyeur avance selon la direction y et une nouvelle ligne est acquise

# Objectifs du projet

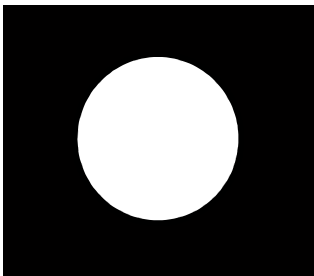
La première étape consiste à effectuer des analyses exploratoires pour faire ressortir les différences majeures entre les échantillons et créer des groupes d'échantillons similaires ayant la même signature spectrale. Nous présentons ici les résultats obtenus par Analyse en composantes principales (ACP), par ACP concaténée et par la méthode des K-means.

Ensuite, des modèles de régression supervisée vont être développés pour relier les images hyperspectrales NIR à des données chimiques et physiques (l'acquisition de ces données par des partenaires du projet est en cours).

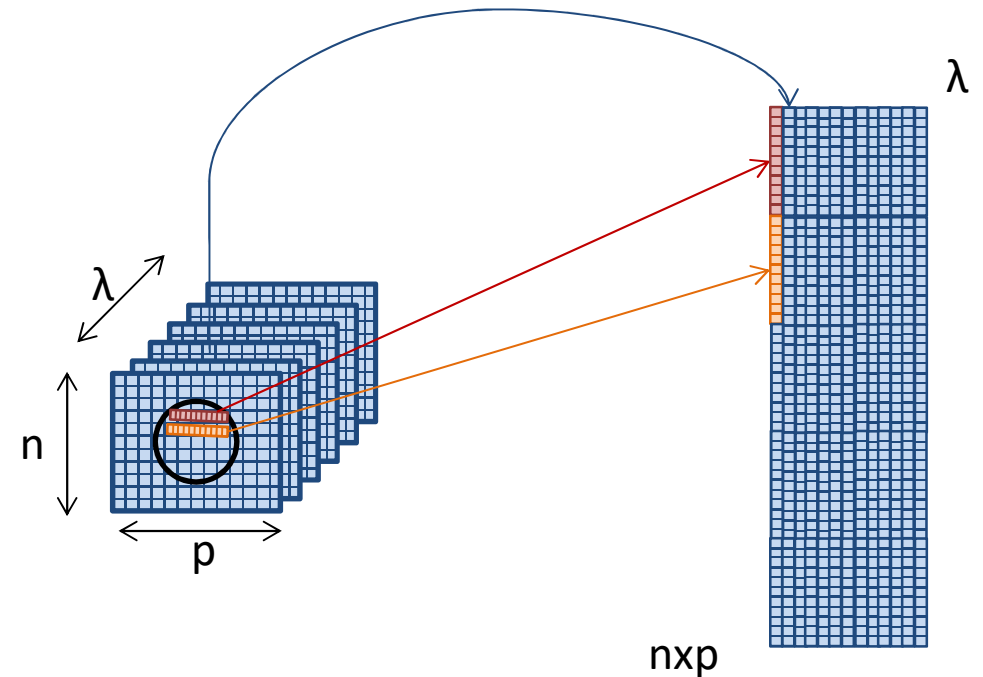
# Prétraitement des données

## Sélection d'une ROI :

Une segmentation manuelle est réalisée sur une image moyenne de chaque cube de données, conduisant à une image binaire (valeurs des pixels égales à 1 pour les pixels dans la ROI et à 0 pour les pixels du fond).

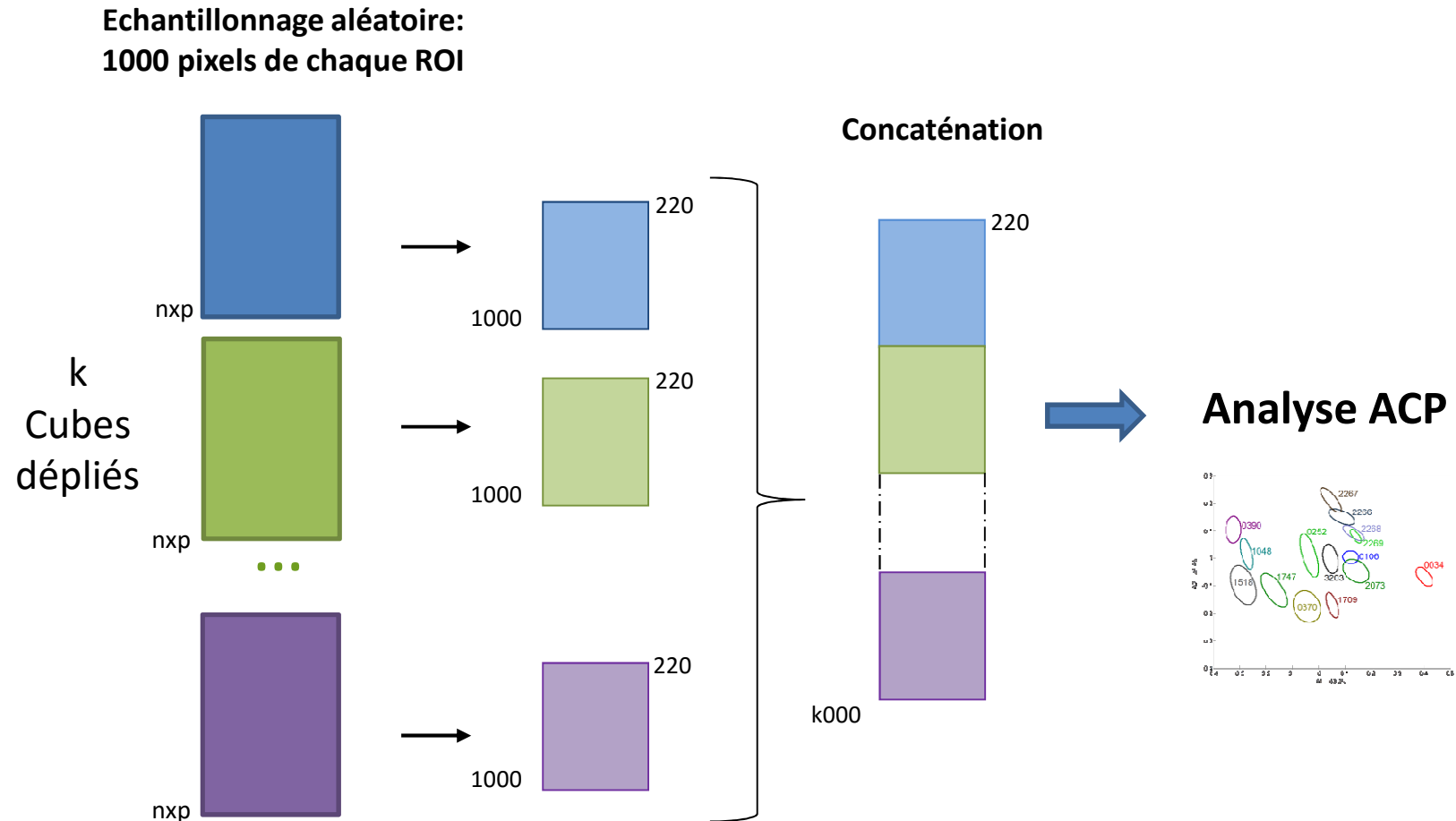


## Dépliage de l'hypercube :



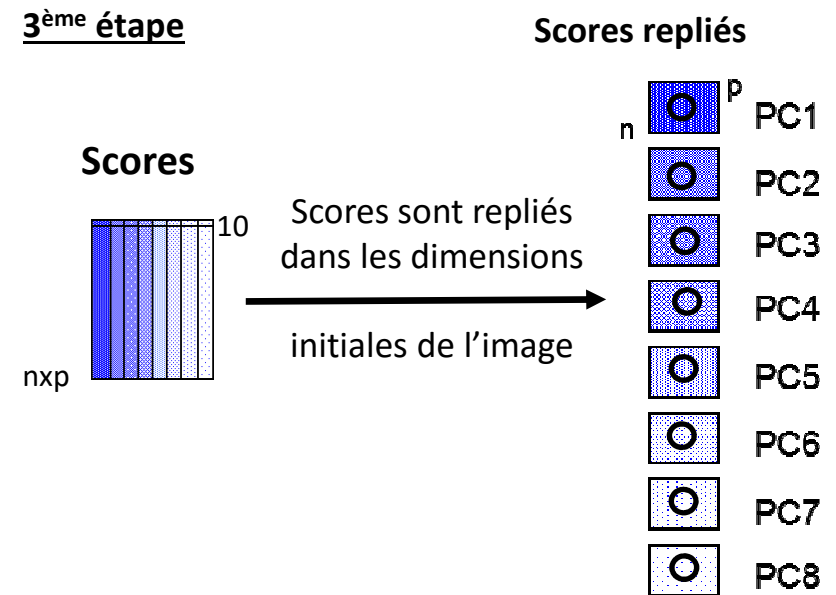
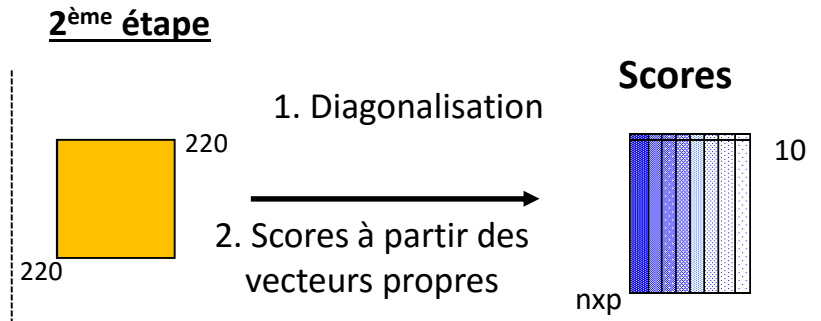
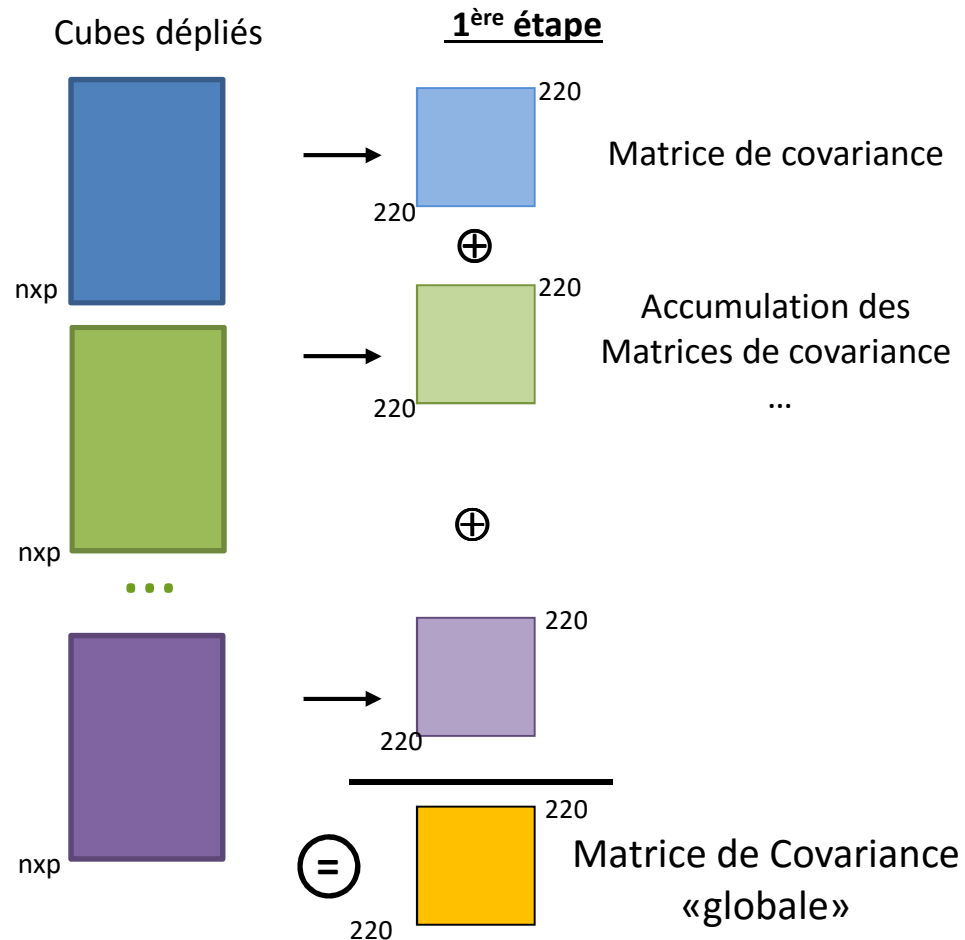
## Smoothing+SNV

# Analyse en composantes principales (ACP)

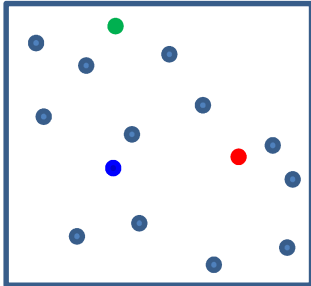




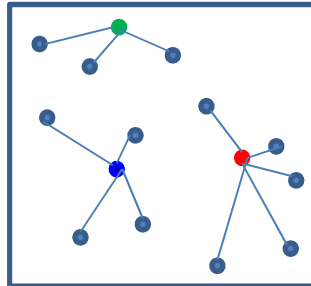
# ACP concaténée



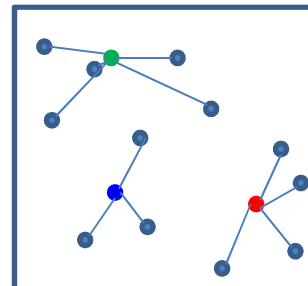
# K-means (1)



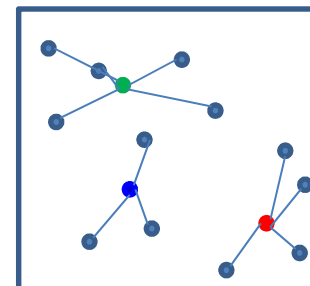
Sélection arbitraire  
de k centres  $\Rightarrow$  k  
groupes



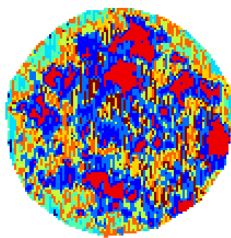
Chaque pixel est  
assigné à un groupe



Les centres de chaque groupe  
sont bougés vers la moyenne  
des points du groupe et les  
pixels assignés aux nouveaux  
groupes



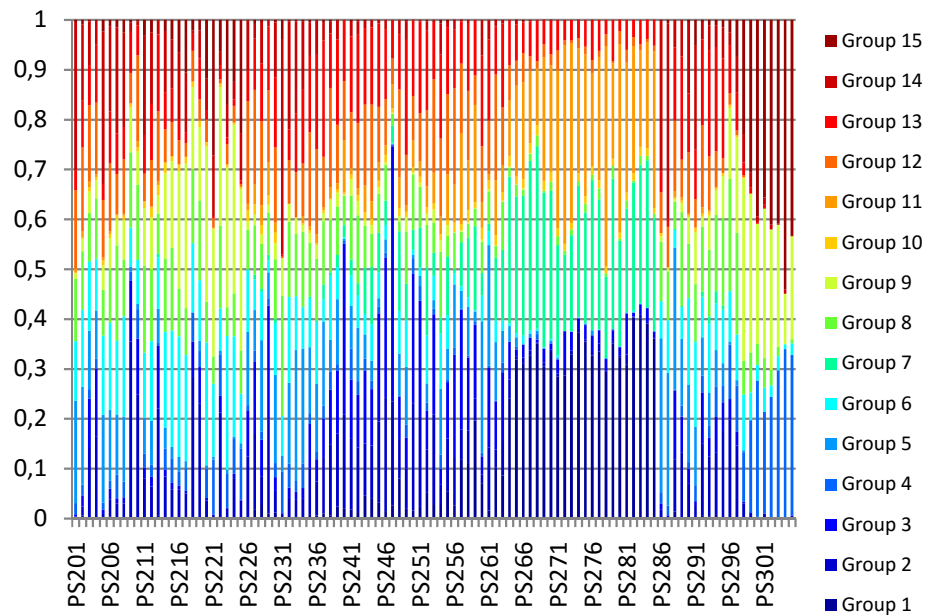
Les centres sont  
encore bougés à la  
moyenne des points



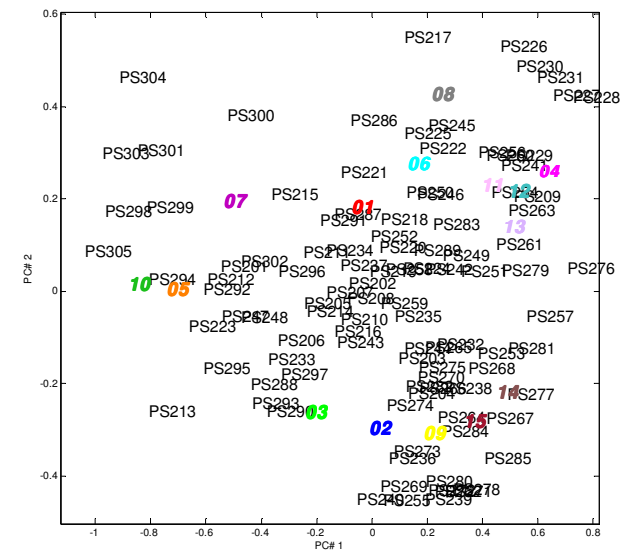
Une image en fausses couleurs peut être recréée en  
assignant une couleur arbitraire à chaque groupe.

# K-means (2)

Les proportions des groupes pour chaque image sont estimées, et une analyse des correspondances est effectuée sur les proportions dans les groupes, de façon à mettre en évidence les relations entre les échantillons et les groupes.



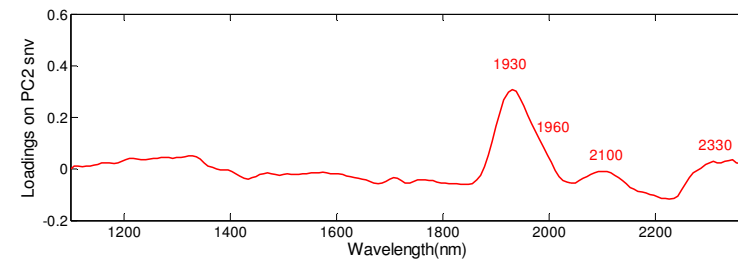
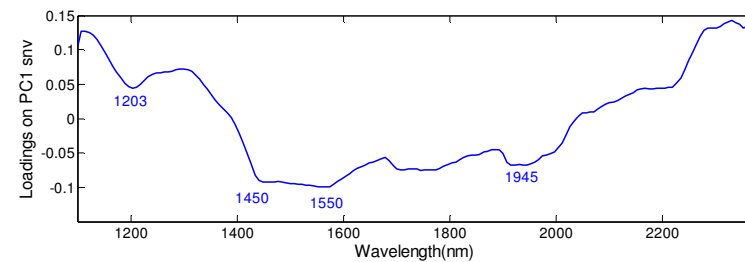
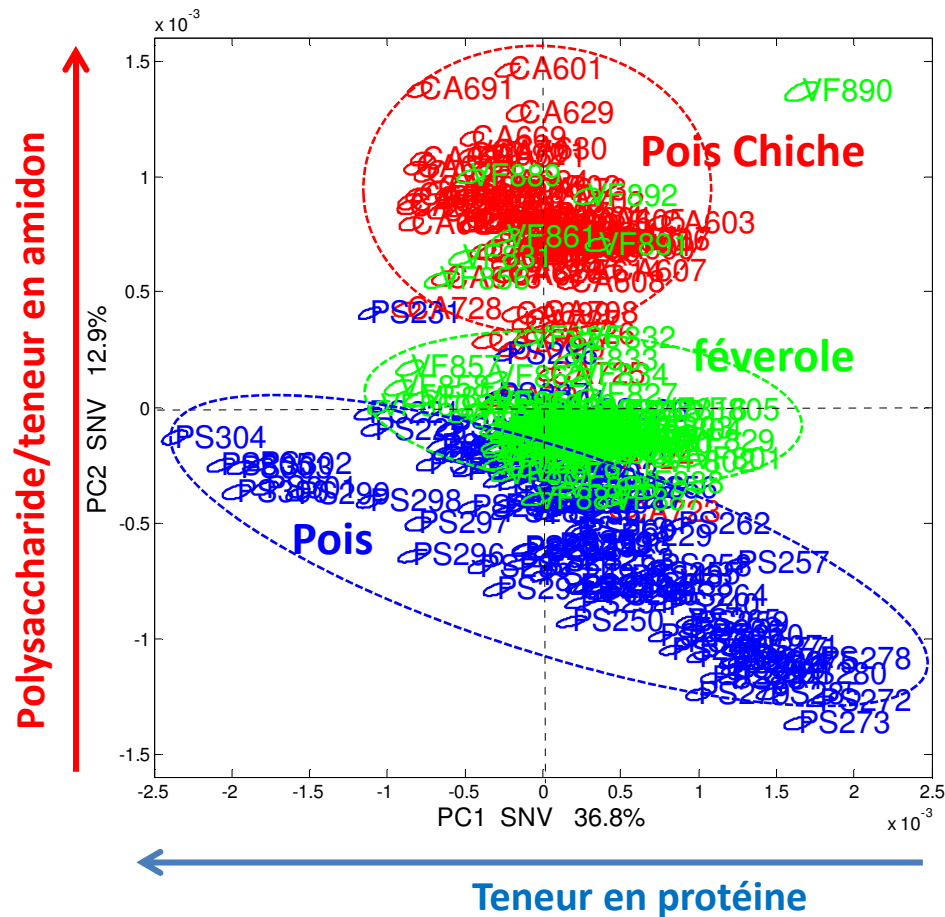
Proportions des groupes pour chaque image



Analyse des correspondances

# Résultats ACP (1)

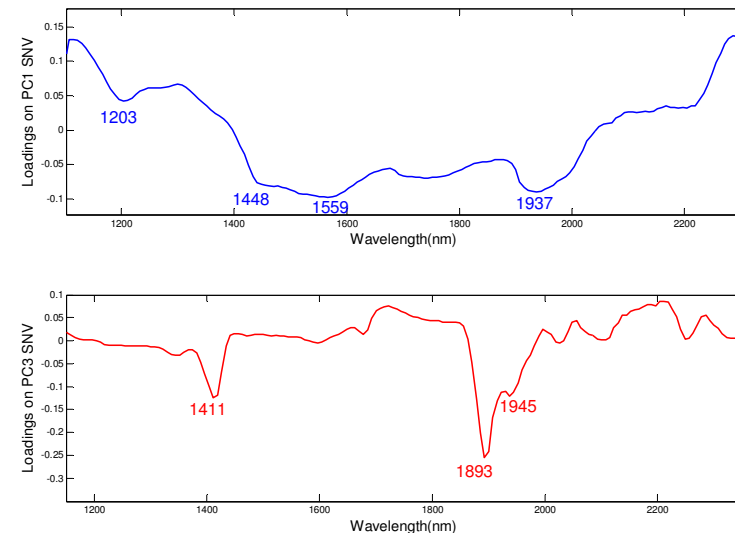
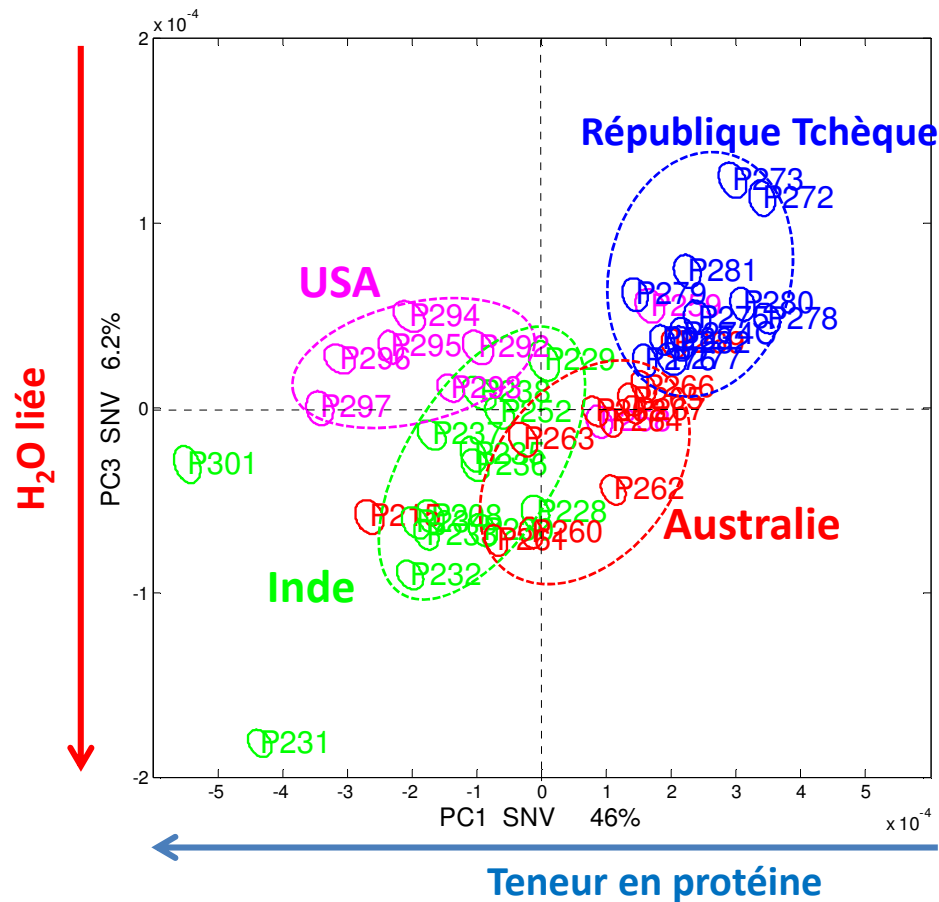
ACP réalisée sur 3 espèces : Pois, féverole et Pois chiche



(à confirmer avec les partenaires en charge des déterminations chimiques)

# Résultats ACP (2)

ACP réalisée sur les échantillons de pois en regard de leur origine

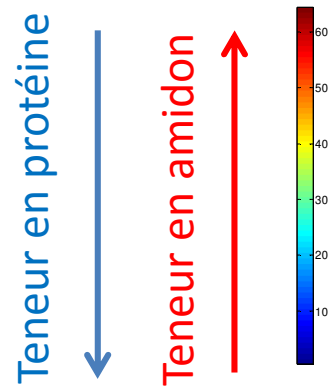


(à confirmer avec les partenaires en charge des déterminations chimiques)

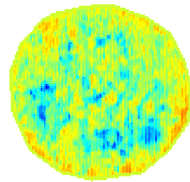
# Résultats ACP concaténée

ACP concaténée réalisée sur les échantillons de pois en regard de leur origine

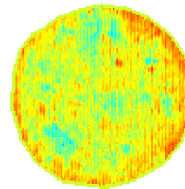
## Scores PC1



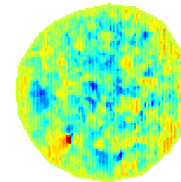
Inde



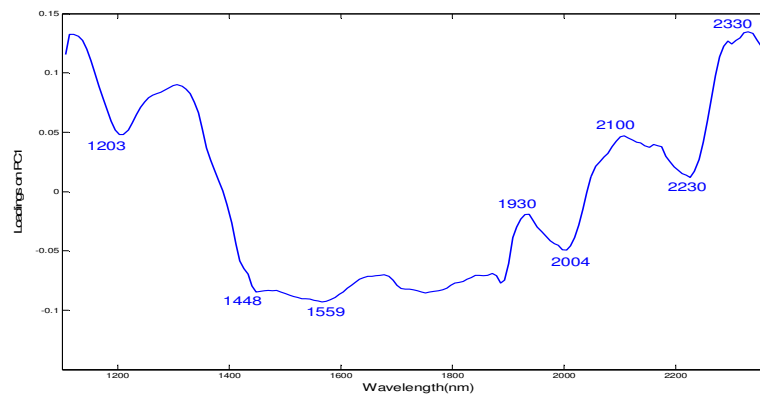
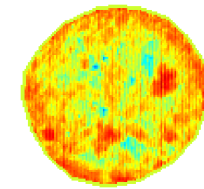
Australie



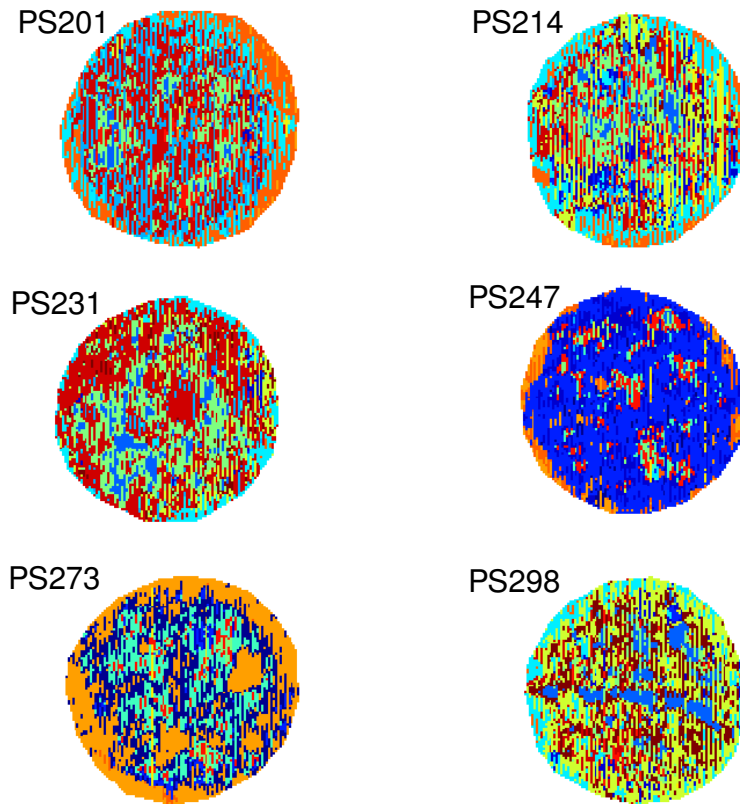
USA



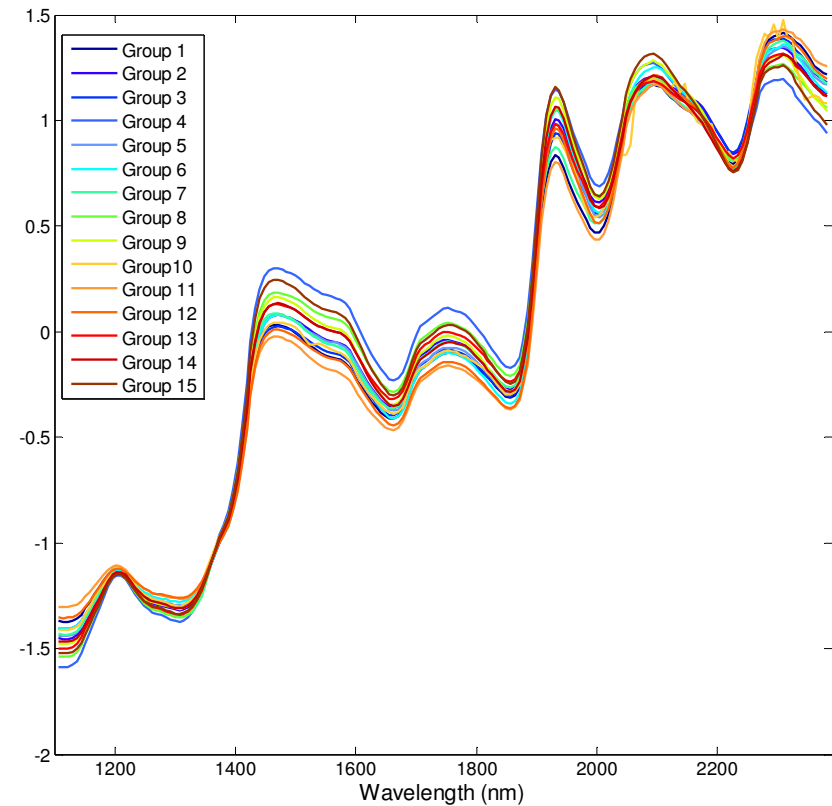
République  
Tchèque



# Résultats K-means Pois (1) - 15 clusters

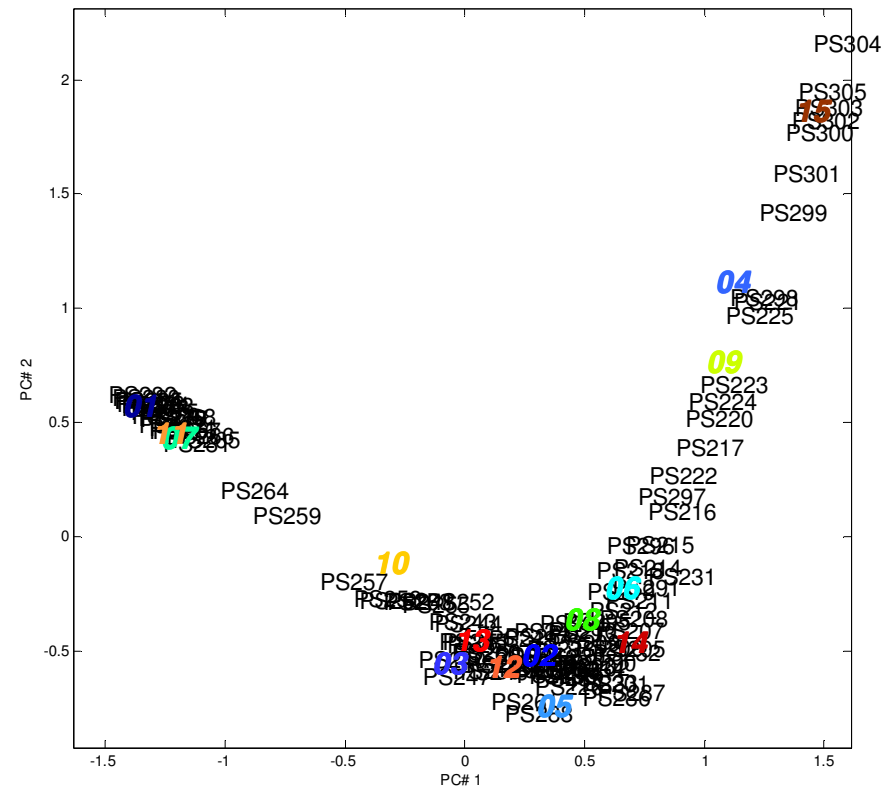
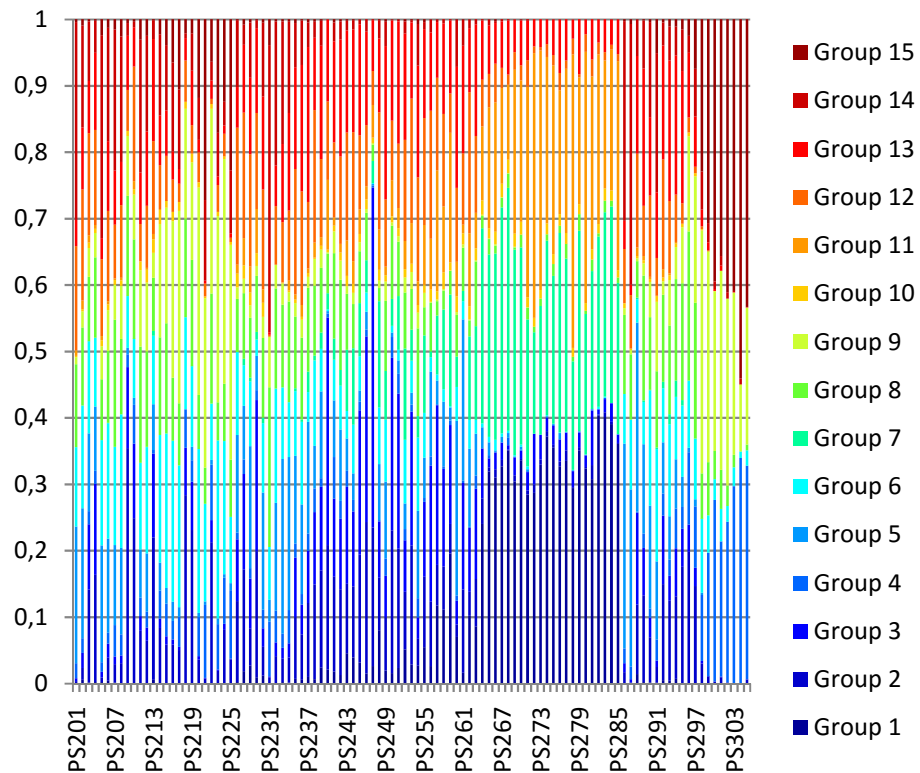


*Images en fausses couleurs*



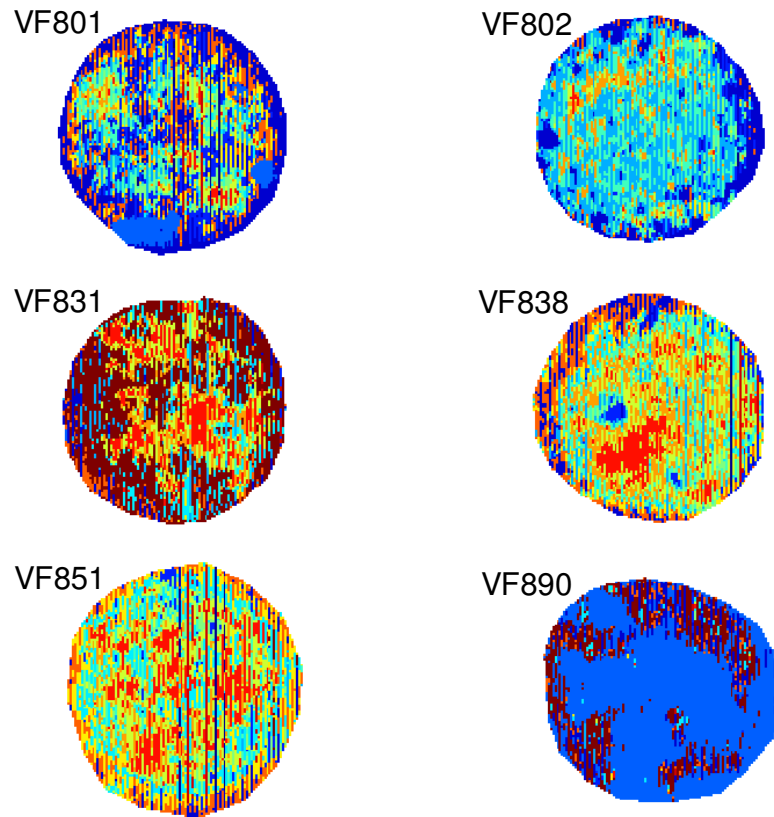
*Barycentres des groupes*

# Résultats K-means Pois (2) - 15 clusters

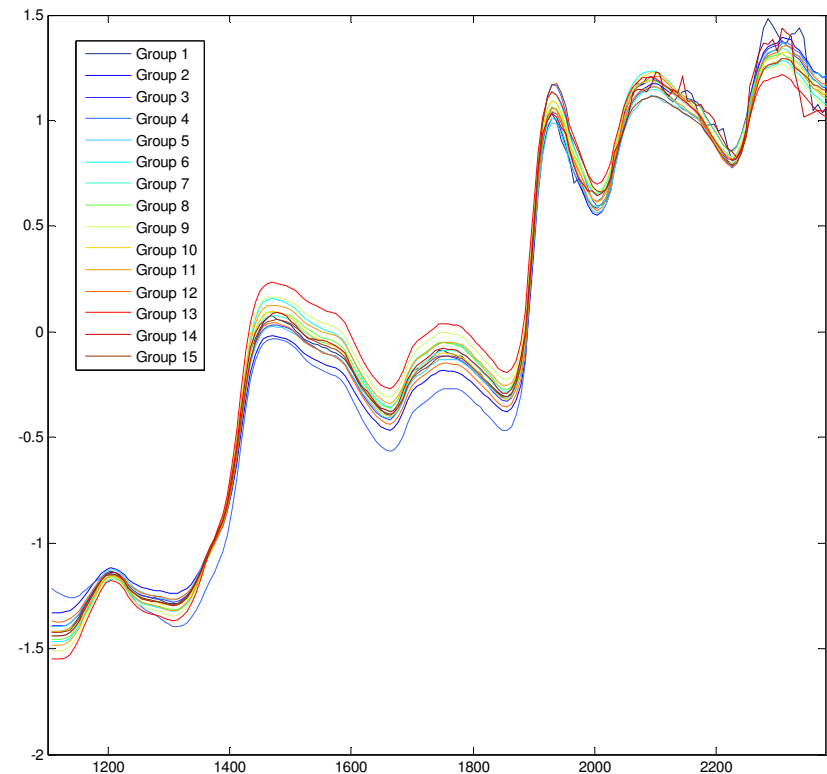




# Résultats K-means féveroles (1) - 15 clusters

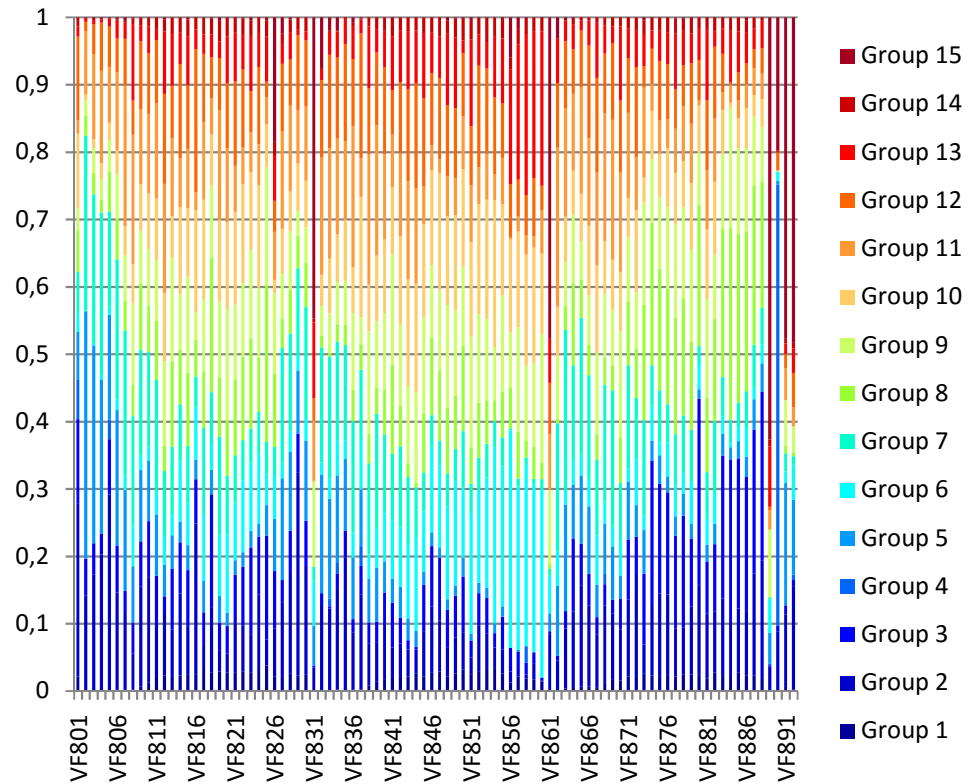


*Images en fausses couleurs*

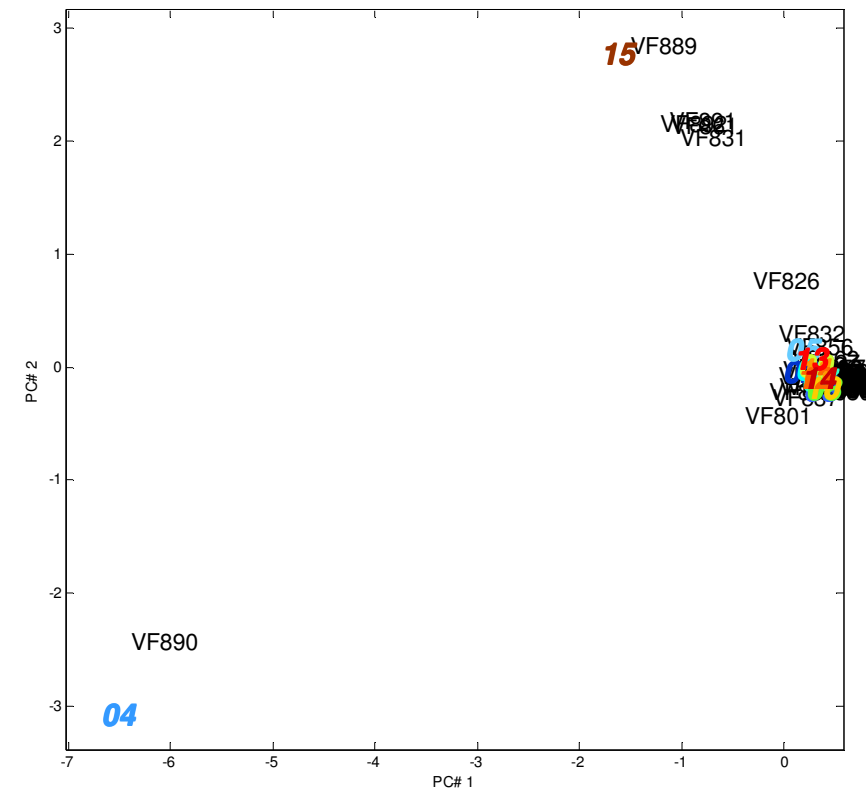


*Barycentres des groupes*

# Résultats K-means féveroles (2) - 15 clusters

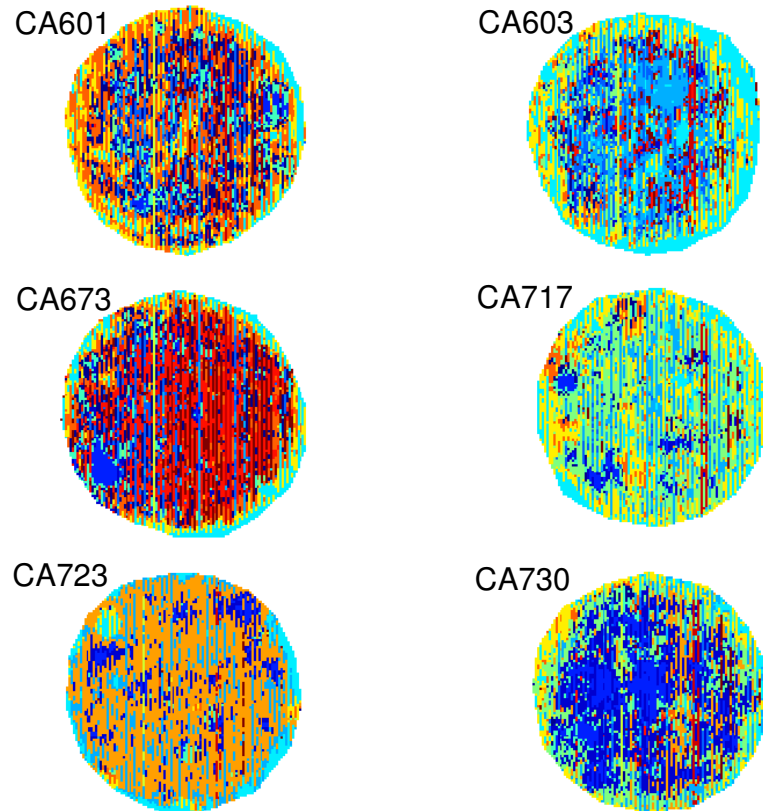


*Proportions des groupes pour chaque image*

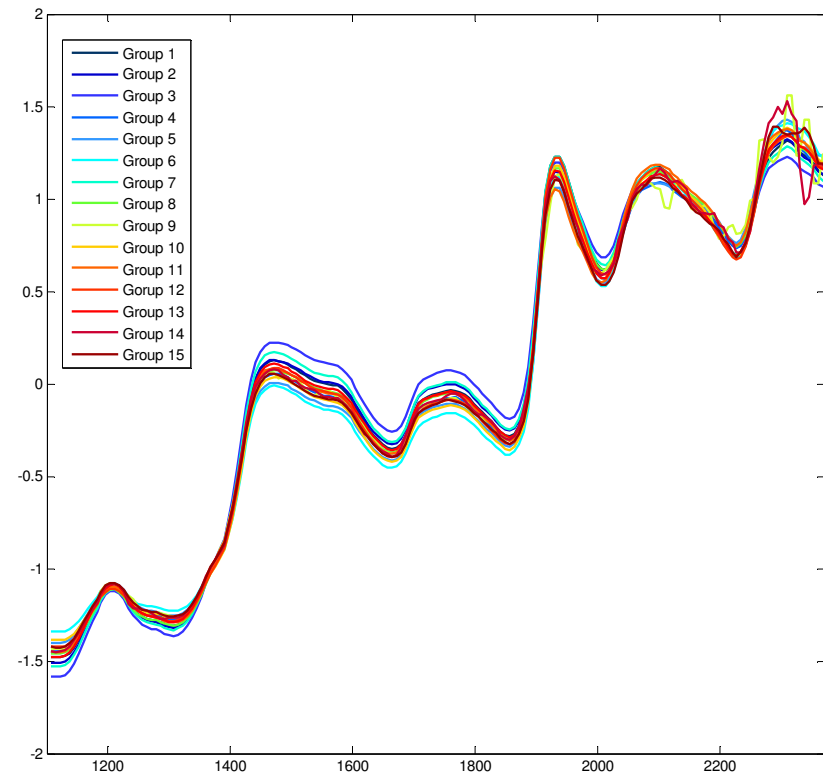


*Carte factorielle de l'analyse des correspondances*

# Résultats K-means Pois Chiche (1) - 15 clusters

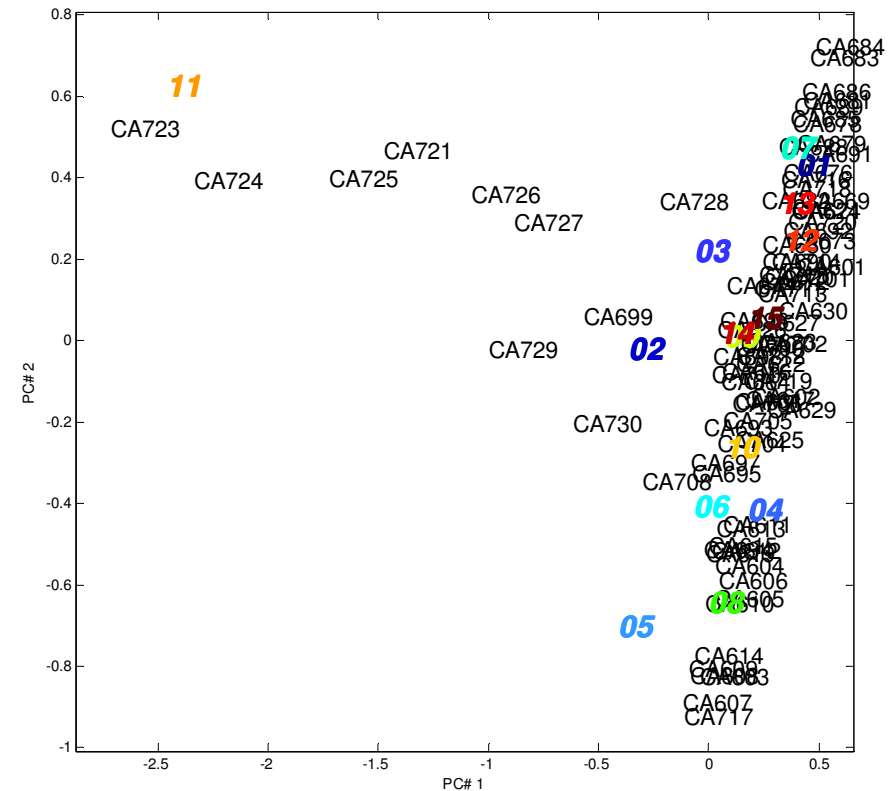
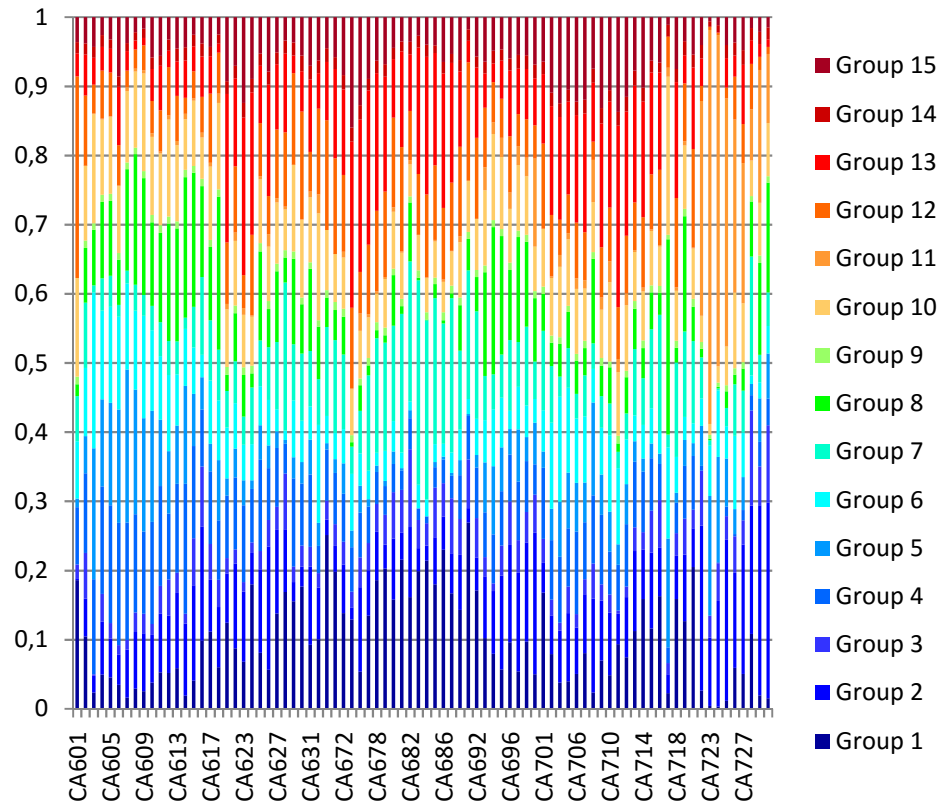


*Images en fausses couleurs*



*Barycentres des groupes*

# Résultats K-means Pois Chiche (2) - 15 clusters



# Conclusions

Les graines de légumineuses broyées de pois, de féveroles et de pois chiche ont été analysées par imagerie hyperspectrale NIR et par des méthodes chimiométriques non supervisées (ACP, ACP concaténée, k-means).

L'ACP permet de séparer rapidement les échantillons en fonction des espèces ou de l'origine au sein d'une même espèce, en utilisant les teneurs en amidon et en protéine. Cependant, ces résultats doivent être confirmés par les analyses chimiques réalisées par les partenaires du projet LEGATO.

L'ACP concaténée peut être utilisée pour comprendre la distribution des composés dans un échantillon et estimer l'homogénéité de celui-ci.

La méthode des Kmeans permet de grouper des échantillons en fonction de leurs signatures spectrales, liées aux teneurs en constituants. Cette méthode met en évidence l'hétérogénéité des échantillons au sein d'une espèce, liée à leur origine, leur sous-espèces et leurs conditions de culture.

# Perspectives

Appliquer les méthodes statistiques non supervisées aux données chimiques et physiques mesurées par les partenaires de LEGATO.

Utiliser la régression multivariée pour relier les images NIR aux teneur en constituants.

## Merci pour votre attention



data\_frame

