



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

## Stratégie de conservation : la génétique seul critère déterminant ?

Réflexions autour de la constitution de  
collections-noyau de rosiers

Thème de la session : Gestion des ressources  
génétiques et pre-breeding

Alix PERNET, Jordan MARIE-MAGDELAINE,  
Cristiana OGHINA-PAVIE, Agnès GRAPIN

IRHS (INRA, AO, UA) et CERHIO (UA, CNRS)

## LA CONSERVATION DES RESSOURCES GÉNÉTIQUES ET LE CONCEPT DE « COLLECTION NOYAU »

La nécessité de conserver et caractériser les RG dans une contexte contraignant

### Les collections noyau

Principaux objectifs :

- Faciliter l'utilisation et l'échange des accessions
- Permettre de transposer des informations

Deux critères essentiels :

- Une **taille réduite**
- Un niveau de redondance le plus faible possible

Sélection d'accessions clés via **l'amande** (« kernel core »)

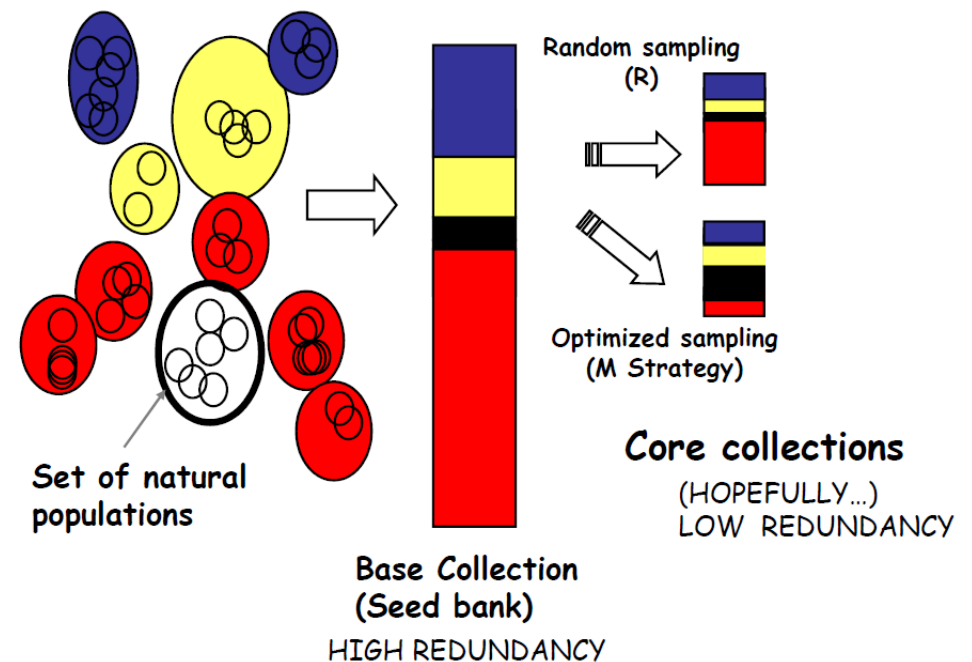




Illustration comparant deux stratégies de constitution de collections noyau (Schoen et Brown, 1993)

## LA PROBLÉMATIQUE : QUELS CRITÈRES POUR DÉTERMINER UNE COLLECTION-NOYAU DE RESSOURCES GÉNÉTIQUES EX SITU ?

### Données disponibles

	Généticiens	Collectionneurs
Profil génétique « neutre »	x	
Profil génétique « sélectionné »	en cours	
Caractérisation phénotypique	x	x
Valeur historique		x
Valeur patrimoniale		x
Valeur commerciale		x
Données éco-géographiques		

 Collections-noyau « classiques »    
  A prendre en compte    
  Résultat différent ?

## L'EXEMPLE DU ROSIER



Rose Madame Falcot.

Plante pérenne

Reproduction sexuée  
préférentiellement allogame

Multiplication végétative des  
rosiers cultivés

Un genre complexe : environ  
150 espèces réparties en 10  
sections dans 1 des 4 sous-genres

Niveaux de ploïdie variés

Hybridations interspécifiques

De très nombreuses variétés,  
réparties en groupes horticoles

**Conservation  
des variétés**

Roseraie  
publique  
ou privée

Parfois en  
réseau :  
FFR

# LES TYPES DE DONNÉES ACQUISES, OBSERVÉES OU CALCULÉES

## Les données collectées par les roseraies et les historiens

### Les données « Passeport »

Variable ou ensemble de variables

Nom de l'obteneur

Période d'obtention

Origine géographique

Donnée  
observée

Donnée  
calculée

### Les données historiques

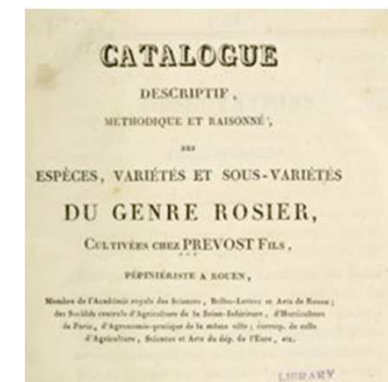
Variable ou ensemble de variables

Groupe horticole

Nombre de références

Valeur patrimoniale

Réputation



## NOTRE ESTIMATION DE LA VALEUR PATRIMONIALE

Variable	Poids attribués aux modalités	Conditions remplies par l'accession
Nom de la variété	1	Ne contient aucun des termes prédéfinis
	2	Contient l'un des termes prédéfinis
	3	Contient plus d'un terme parmi ceux prédéfinis
Période d'obtention	1	N'a pas été obtenue au cours du XVIII ou XIXème siècle
	2	A été obtenue au cours du XVIII ou XIXème siècle
Nombre de références bibliographiques	1	Est décrite par une seule référence ou aucune
	2	Est décrite par 2 à 9 références
	3	Est décrite par 10 références ou plus
	X-1	Est un rosier récent
Réputation	1	N'est pas réputée
	2	Est décrite par une remarque sur sa réputation ou a été primée une fois
	3	Cumule deux remarques ou plus

Notations utilisées pour la conception de la variable « valeur patrimoniale » ; la valeur « X-1 » indique un malus de 1 attribué au score global d'un rosier récent

# LES TYPES DE DONNÉES ACQUISES, OBSERVÉES OU CALCULÉES

## Les données collectées par les roseraies et les généticiens

### Les données phénotypiques observées

Variable ou ensemble de variables
Lustre de la face supérieure des feuilles
Aiguillons
Parfum
Hauteur
Duplicature



*Belle Hélène*



*Du Mortier*



Donnée observée



Donnée calculée

### Les données phénotypiques calculées

Variable ou ensemble de variables
Nombre de fleurs par volume
Forme détaillée des fleurs
Remontance
Couleurs détaillées des fleurs



*Château de Namur*

### Les données génétiques

Variable ou ensemble de variables
Données SSR

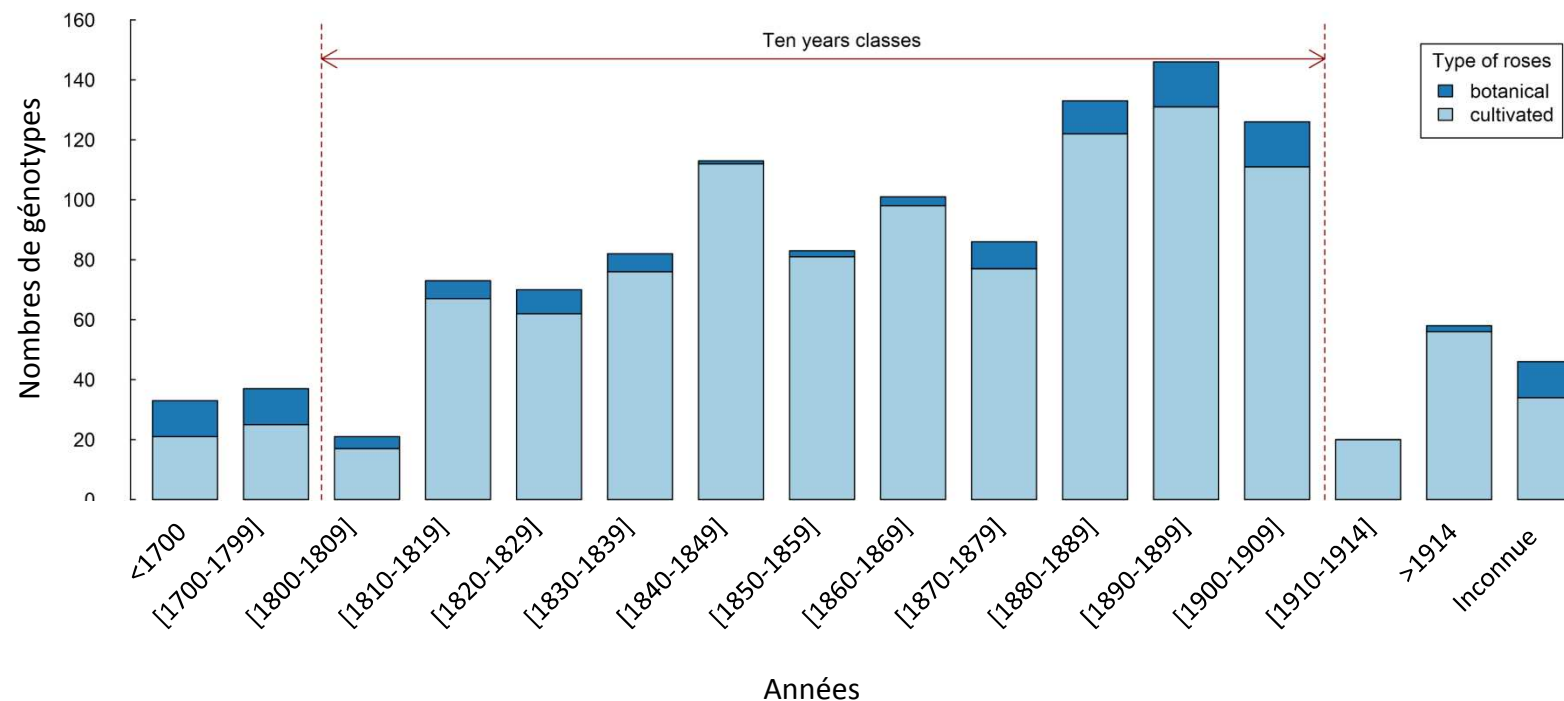


*Mermaid*

## LE PROJET FLORHIGE : CAS D'ÉTUDE

### Le matériel végétal

- 1534 accessions dont 1228 analysés par SSR (avec moins de 30% de données manquantes)
  - 991 roses des 18 et 19<sup>e</sup> siècles (1700-1914)
  - 18 roses anciennes (< 1700)
  - 118 roses botaniques
  - 53 roses modernes (>1914)



Répartition des années d'obtention pour les 1110 variétés de rosiers  
et années d'introduction pour les 118 rosiers botaniques en 15 périodes (Liorzou et al 2016)



## LE PROJET FLORHIGE : STRUCTURATION DE LA DIVERSITÉ NEUTRE



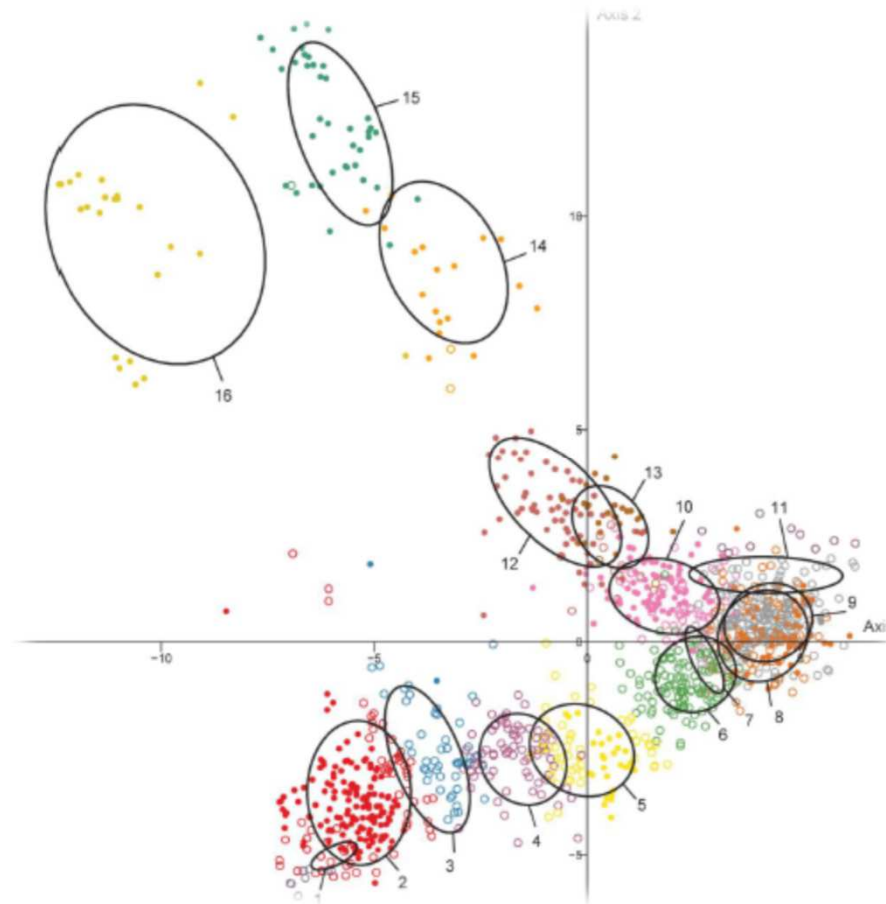
Duchesse de Portland  
av. 1775 Inconnu



Le Rosier Evêque  
1815 Descemet



Du Mortier  
1843 Parmentier



Représentation de la dispersion spatiale des groupes  
génétiques de 1228 rosiers (Liorzou et *al.*, 2016)



'Persian Yellow'  
1837 Willock



'Maréchal Niel'  
1864 Pradel



'Rose à parfum de l'Hay'  
1901 Gravereaux

## LES STRATÉGIES ET MÉTHODES DE CONSTITUTION DES COLLECTIONS NOYAU

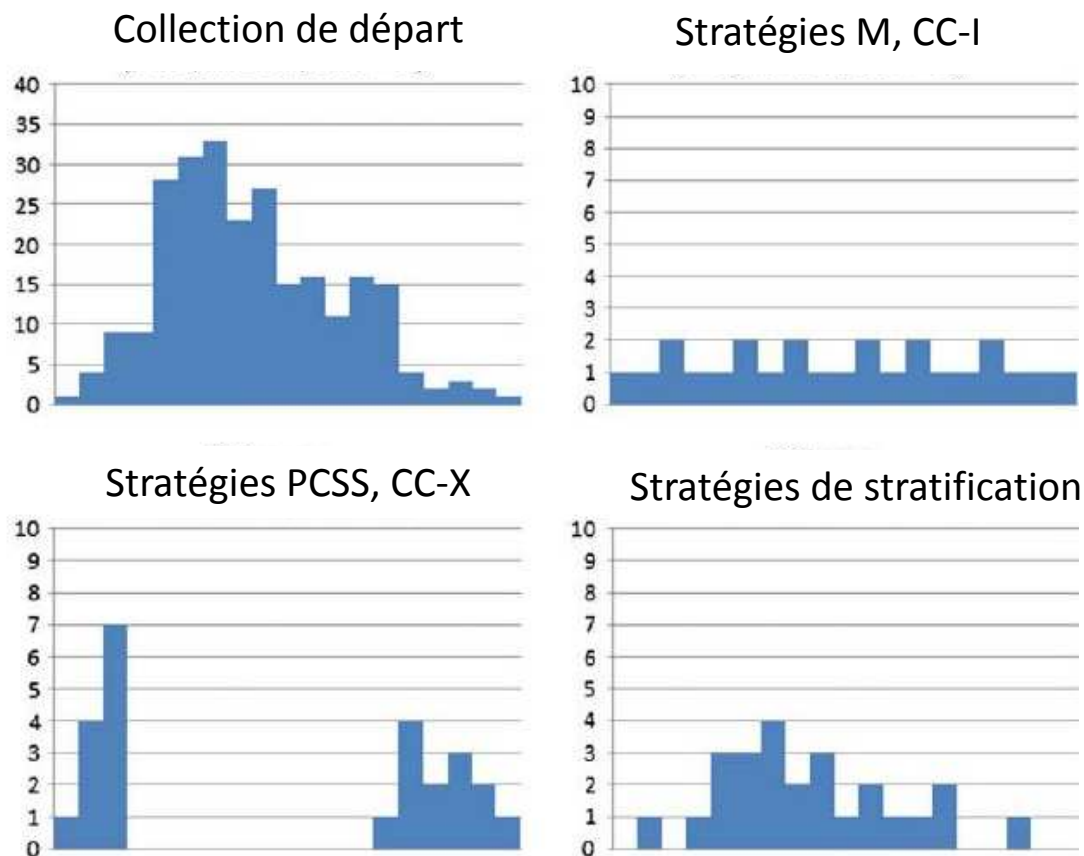
Selon l'objectif ou l'utilisateur final :

**Maximiser la diversité totale**

- Généticiens
- Curateurs de banques de gènes

Représenter la **structuration** de la diversité

- Sélectionneurs



Distribution multimodale d'un trait phénotypique selon les stratégies employées (d'après Odong et al, 2013)

**Taille des collections noyau :** en fonction de la diversité capturée par maximisation

→ Environ 10 % soit 151 entrées

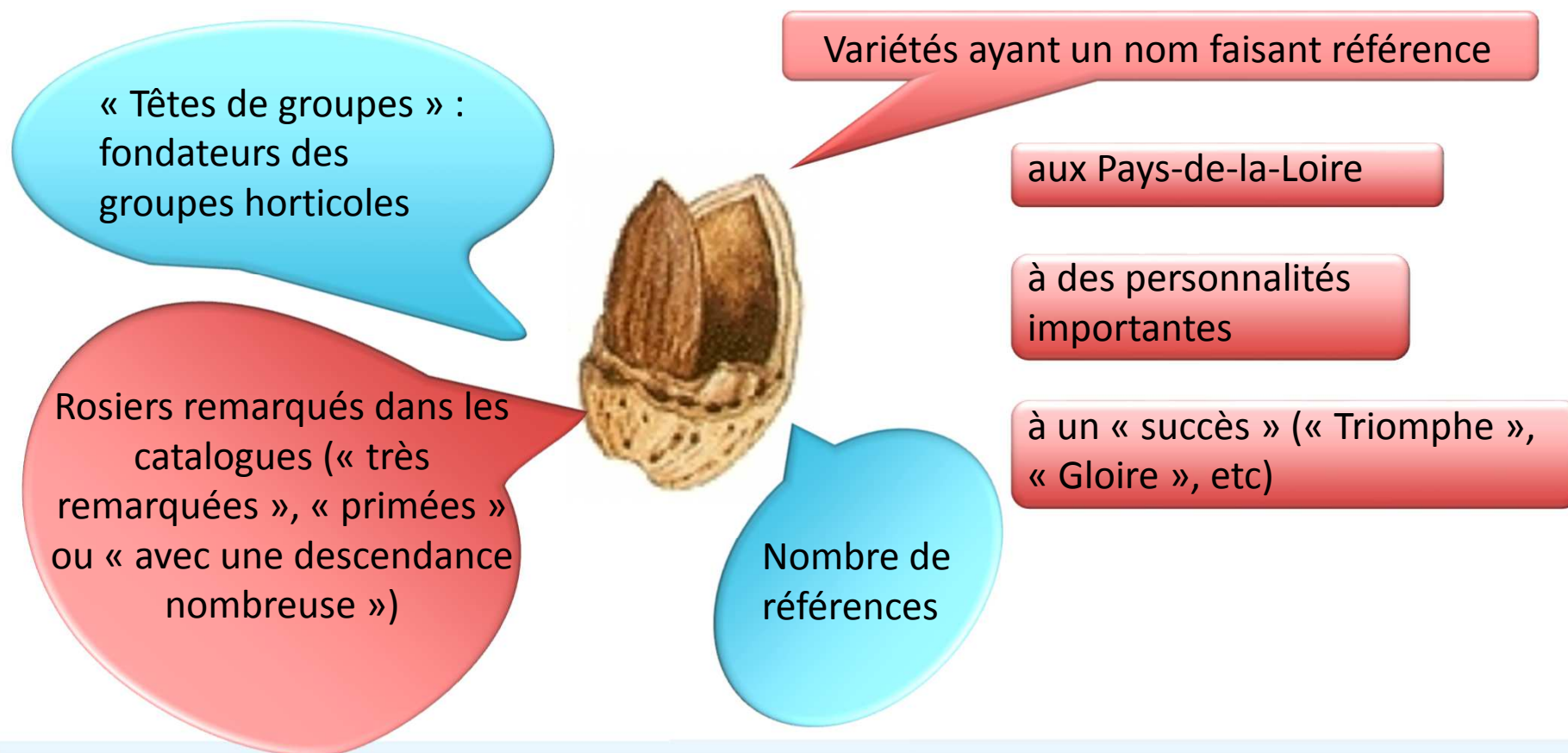
→ Aussi 20 % soit 249 entrées à titre de comparaison

## L'AMANDE : DÉFINITION ET MÉTHODE DE CONSTITUTION

**Définition** : Ensemble d'accessions toujours incluses dans la collection-noyau finale

**Taille** : 26 accessions (15 à 20 % de la taille de la collection-noyau)

**Choix** : selon des données historiques et patrimoniales



## LES CRITÈRES D'ÉVALUATION DES COLLECTIONS NOYAU CONSTITUÉES

### Comparaison de la collection noyau avec la collection entière

#### Indices de diversité (richesse allélique)

- indice de diversité de Nei (Nei)
- indice de Shannon (SH)
- indice de couverture de la variable (CV)

#### Indice de distance génétique (diversité)

- distance modifiée de Rogers (MR)

→ Estimés avec les données SSR

### Comparaison des collections noyau entre elles

Indice de similarité sur le modèle de Sneath et Sokal (simple matching)

$$\frac{a + d}{a + b + c + d}$$

## LES LOGICIELS ET LEURS CARACTÉRISTIQUES

Programme	Mstrat	PowerCore	Corehunter 2.0	R
Référence	Gouesnard et al 2001	Kim et al 2007	Thachuck et al 2009	Studnicki et Debski 2012
Stratégie	De maximisation	De maximisation	De maximisation	De stratification
Vitesse de calcul	Lente	Rapide	Rapide	Rapide
Facilité d'utilisation	Très faible (formats imposés)	Grande	Grande	Moyenne (langage à connaître)
Ergonomie	Faible	Bonne	Moyenne	Selon le logiciel utilisé pour programmer et/ou utiliser les scripts
Ligne de commande	Possible	Impossible	Possible	Possible
Amande	Possible	Possible	Impossible	Possible
Indices	Nei, SH	CV	7 différents (MR, CE, SH, HE, NE, PN, CV)	Aucun
Algorithmes	Un seul	Heuristique ou non-heuristique	13 différents (exh, lr, mixrep, mstrat, rand, etc.)	Selon les packages
Remarques	Evalue la taille optimale et compare les stratégies M et aléatoire	Evalue la taille optimale, mais empêche le choix d'une taille spécifique	Permet l'évaluation ultérieure de collections noyau	Permet de créer ses propres algorithmes, de découvrir de nouveaux packages

## MODALITÉS DE CONSTITUTION DES COLLECTIONS NOYAU : TAILLE DES COLLECTIONS NOYAU OBTENUES EN FONCTION DES DIFFÉRENTS FACTEURS

Stratégie	Logiciel	Sélection d'une amande	Variables						
			SSR	Multicritères	Nested	Sélection « PPP »	Passeport	Phénotypiques	Patrimoniales
Maximisation	Mstrat	Sans	151, 249	151, 249	151		151	151	151
		Avec	151	151	151	151			
	CoreHunter	Sans	151, 249	151, 249	151	151			
	PowerCore	Sans	244	244	244	141	141	141	141
		Avec	244	244	244				
Stratification	R	Sans				151	151	151	151

# LE PROJET FLORHIGE : CARACTÉRISATION DES DONNÉES ET POUVOIR DISCRIMINANT

## Les données collectées par les roseraies et les historiens

### Les données « Passeport »

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Nom de l'obteneur	298	1189	25,1
Période d'obtention	16	1185	1,4
Origine géographique	10	1023	1

Donnée observée

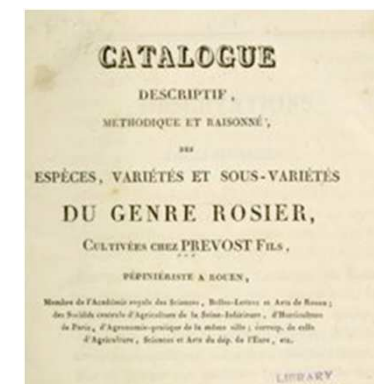


Donnée calculée



### Les données historiques

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Groupe horticole	16	1023	1,6
Nombre de références	20	1412	1,4
Valeur patrimoniale	36	1171	3,1
Réputation	3	412	0,70





# LE PROJET FLORHIGE : CARACTÉRISATION DES DONNÉES ET POUVOIR DISCRIMINANT

## Les données collectées par les roseraies et les historiens

### Les données « Passeport »

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)	SélectionPPP
Nom de l'obteneur	298	1189	25,1	x
Période d'obtention	16	1185	1,4	x
Origine géographique	10	1023	1	x

Donnée observée

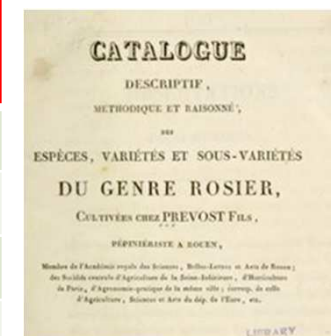


Donnée calculée



### Les données historiques

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)	SélectionPPP
Groupe horticole	16	1023	1,6	x
Nombre de références	20	1412	1,4	
Valeur patrimoniale	36	1171	3,1	x
Réputation	3	412	0,7	





# LE PROJET FLORHIGE : CARACTÉRISATION DES DONNÉES ET POUVOIR DISCRIMINANT

## Les données collectées par les roseraies et les généticiens

### Les données phénotypiques observées

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Lustre de la face supérieure des feuilles	4	34	11,8
Aiguillons	4	107	3,7
Parfum	2	62	3,2
Hauteur	9	1412	0,6
Duplication	6	1079	0,6

### Les données phénotypiques calculées

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Nombre de fleurs par volume	4	42	9,5
Forme détaillée des fleurs	6	96	6,3
Remontance	6	329	1,8
Couleurs détaillées des fleurs	14	1339	1,0

### Les données génétiques

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Données SSR	1526	1534	99,5



*Belle  
Hélène*



*Du  
Mortier*



Donnée  
observée



Donnée  
calculée



*Château  
de Namur*



*Mermaid*

# LE PROJET FLORHIGE : CARACTÉRISATION DES DONNÉES ET POUVOIR DISCRIMINANT

## Les données collectées par les roseraies et les généticiens

### Les données phénotypiques observées

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Lustre de la face supérieure des feuilles	4	34	11,8
Aiguillons	4	107	3,7
Parfum	2	62	3,2
Hauteur	9	1412	0,6
Duplication	6	1079	0,6

#### SélectionPPP

x



*Belle Hélène*



*Du Mortier*



Donnée observée

### Les données phénotypiques calculées

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Nombre de fleurs par volume	4	42	9,5
Forme détaillée des fleurs	6	96	6,3
Remontance	6	329	1,8
Couleurs détaillées des fleurs	14	1339	1,0

#### SélectionPPP

x



Donnée calculée



*Château de Namur*

### Les données génétiques

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Données SSR	1526	1534	99,5



*Mermaid*

## CARACTÉRISATION DES ENSEMBLES DE DONNÉES

Variable ou ensemble de variables	Nombre de classes	Nombre d'accessions	Pouvoir discriminant (%)
Données SSR	1526	1534	99,5
Ensemble des données, hors SSR	1288	1511	85,2
Sélection « PPP »	955	1164	82
Données passeport	917	1189	77,1
Données phénotypiques	942	1339	70,4
Données patrimoniales	893	1412	63,2



## COMPARAISON DES COLLECTIONS NOYAU CONSTITUÉES SELON LE TYPE DE DONNÉES

Données			SSR		Multi critères		Nested		Sélection PPP		Passeport		Phénotypiques		Patrimoniales	
Logiciel	Indice	Collection	T151	H	T151	H	T151	H	T151	H	T151	H	T151	H	T151	H
Mstrat		Nei	73.96		75.62											
		MR	0.16		0.16		0.14				0.15		0.15		0.15	
		SH	5.73		5.67		5.67				5.54		5.56		5.54	
		CV	0.71		0.72		0.66				0.58		0.61		0.56	
CoreHunter		MR	0.21		0.34		0.19		0.17							
		SH	5.95		4.76		5.77		5.70							
		CV	0.73		0.66		0.68		0.64							
PowerCore		MR		0.16		0.16		0.16		0.11		0.13		0.13		0.11
		SH		5.99		5.96		5.58		4.76		4.74		4.73		4.70
		CV		0.94		0.89		0.68		0.19		0.10		0.13		0.13
R		MR							0.14		0.16		0.15		0.14	
		SH							5.52		5.49		5.52		5.49	
		CV							0.61		0.57		0.59		0.57	

## LA COMPARAISON DES NIVEAUX DE DIVERSITÉ RETENUS PAR NOS COLLECTIONS

- Taille idéale entre 10 et 16%
- Capacité des stratégies à capturer la diversité initiale :
  - Stratégie M >> Stratégie Aléatoire
  - Stratégies de maximisation et de stratification **relativement similaires**
- Influence du type de données sur la qualité :
  - **Couverture allélique très variable**
  - Données SSR suffisantes pour la capture d'un maximum d'allèles
  - Données PPP peuvent donner de bons résultats (sauf pour CV) – à voir en fonction des données phénotypiques et patrimoniales
  - Multicritères pas utile et Nested encore moins



Dispersion spatiale des entrées (en orange) d'une collection noyau constituée via la stratégie M

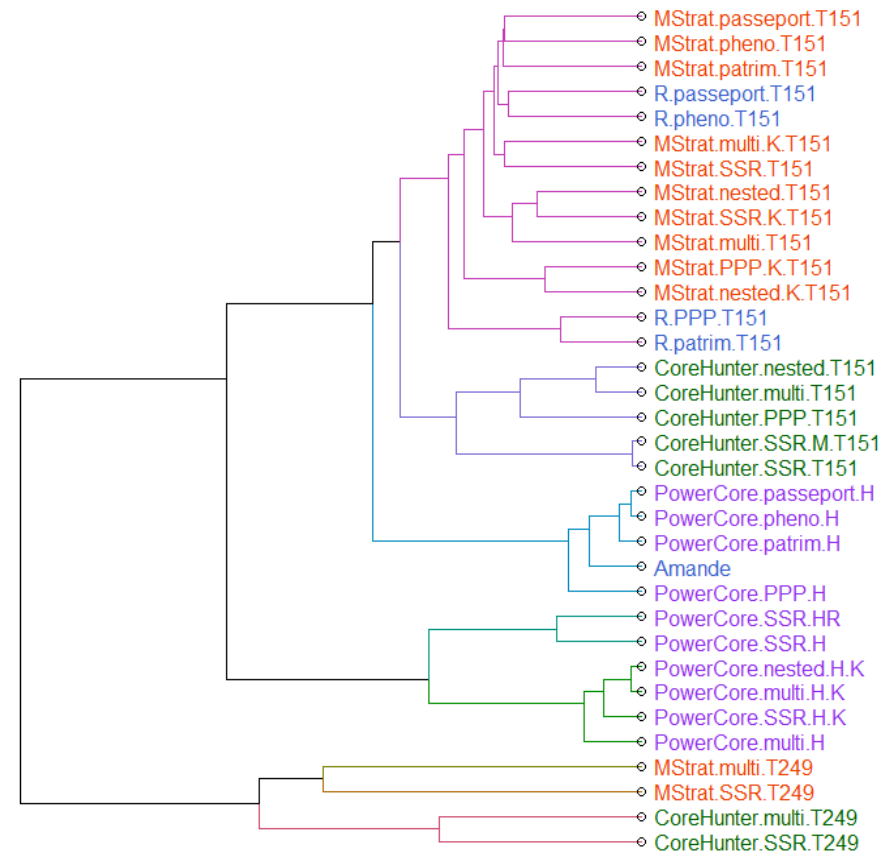
## L'ÉVALUATION DE LA SIMILARITÉ DES COLLECTIONS NOYAU CONSTITUÉES

Collections noyau  
globalement similaires :

- Taux de similarité entre  
62% et 99,5%  
(moyenne de 81%)

Évaluation de l'effet des  
facteurs étudiés :

- Effet **très marqué** de  
la taille et du  
programme
- Effet modéré des types  
de données utilisées
- Effet **très faible** de  
l'amande constituée



Dendrogramme représentant les collections noyau  
constituées en fonction de leur niveau de similarité

## DISCUSSION - CONCLUSION

**Estimation de la valeur patrimoniale** : beaucoup de travail de collecte des données par les historiens et un peu de subjectivité

→ analyser plus finement l'impact de chaque variable sur la valeur patrimoniale et la constitution des collections noyau ?

**Données phénotypiques** : quantité de travail en fonction des différents caractères

→ là aussi, analyser plus finement

**La génétique : seul critère déterminant ?**

C'est possible MAIS

risque d'omettre des accessions mutantes innovantes

des accessions patrimonialement importantes

→ Utiliser une amande

**OU** constituer des collections à partir de données phénotypiques et patrimoniales (ce qu'ont les roseraies en main)

**Génétique OU ensemble PPP : compromis coût / temps / objectifs**



## PERSPECTIVES

### Perspectives méthodologiques

Sur les valeurs patrimoniale et phénotypique : composition (quel apport du champ « Commentaires », intérêt respectif ou cumulé)

Sur les indices d'évaluation en fonction du type de collection à constituer

Sur l'optimisation des SSR

### Perspectives pour la conservation

Ensemble des RG rosiers importantes à conserver

- du point de vue patrimonial → Analyse patrimoniale à faire
- du point de vue de la création variétale → Analyse phénotypique, et génétique pour limiter les apparentements → intérêt d'une collection noyau

Sécurisation de la conservation

- Réseau dynamique de conservation
- Cryoconservation sur une collection noyau

## UMR IRHS

Chastellier Annie

Clotault Jérémy

Dupuis Fabrice

Foucher Fabrice

Gaillard Sylvain

Li Shubin

Liorzou Mathilde

Malécot Valéry

Michel Gilles

Thouroude Tatiana

Tous les autres phénotypeurs

## CERHIO

Antoine Annie

Briée Céline

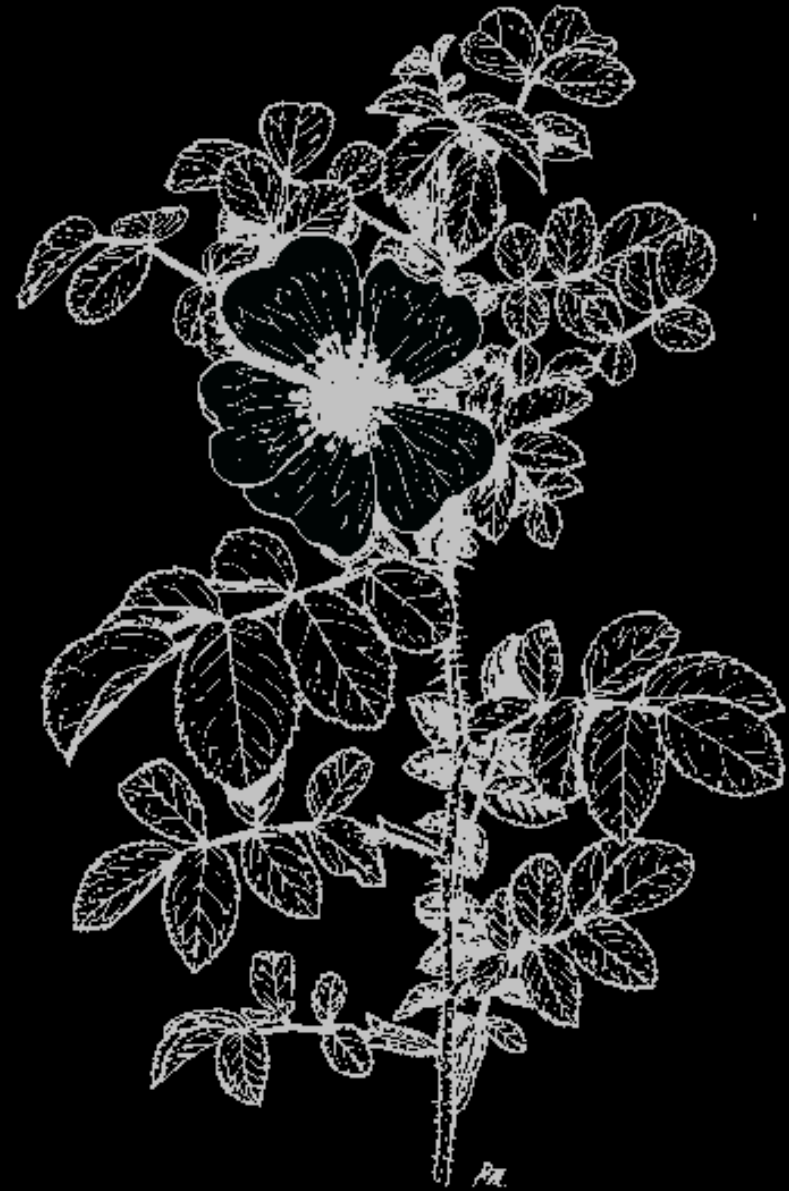
Ferrand Nathalie

## Plateforme Gent Yane

Poncet Charles

Philippe Géraldine

Chevaudonna Karine



*Rosa gallica* L. Dominique Mansion

**Les roseraies partenaires : Loubert, Nantes, Val de Marne**

*Merci pour votre attention*

