



ACTIVITÉS DE RECHERCHE EN BIO-INFORMATIQUE GROUPE FLORIMOND-DESPREZ

09/11/2023





- Présentation du groupe FD
- Activités de recherche

Sommaire





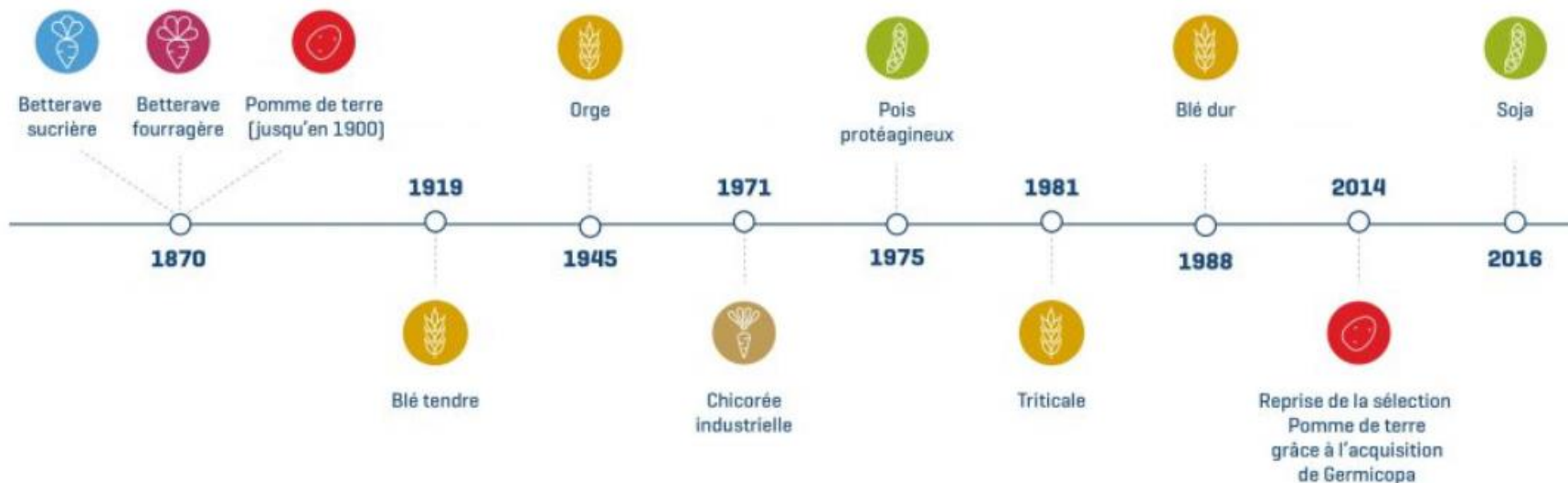
GROUPE FLORIMOND-DESPREZ





Un groupe familial et indépendant, spécialisé dans la production de semences et l'innovation variétale

- Création en 1830
- Localisation à Cappelle-en-Pévèle
- Création variétale pour 10 espèces de grandes cultures





ACTIVITÉS DE RECHERCHE

2





SELECTION VARIETALE

Création de nouvelles variétés

Betterave

Pois
protéagineux

Soja

Blé dur

Triticale

Orge

Blé tendre

Chicorée

Betterave
fourragère

Pomme de
terre



SELECTION VARIETALE

Création de nouvelles variétés

Betterave

Pois
protéagineux

Soja

Blé dur

Triticale

Orge

Blé tendre

Chicorée

Betterave
fourragère

Pomme de
terre

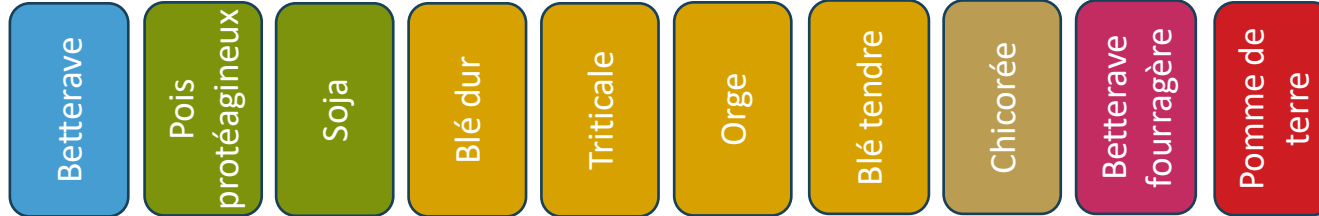
SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique et Statistique

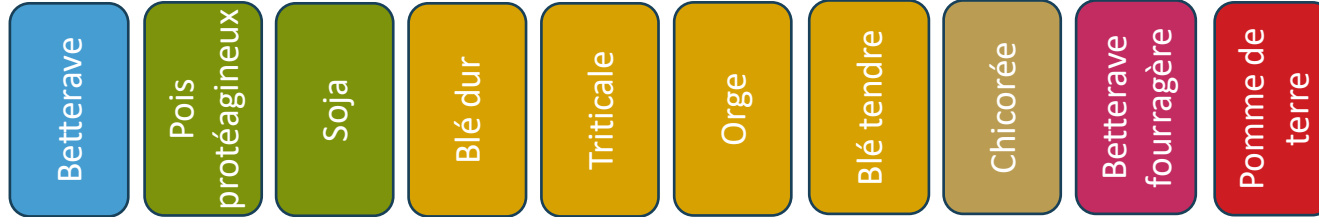
Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique et Statistique

Différentes équipes & Différents espèces

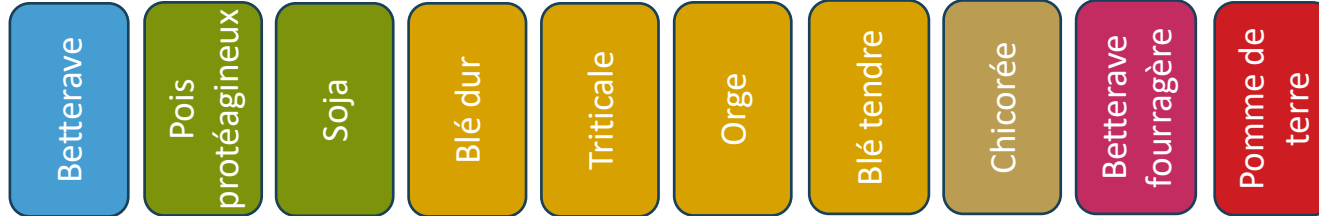
- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

Développement de marqueurs moléculaires



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique et Statistique

Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

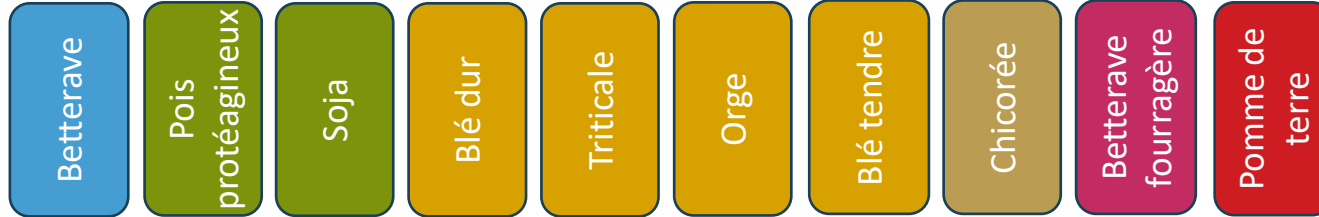
Développement de marqueurs moléculaires

Développement de méthodes d'analyses



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique et Statistique

Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

Développement de marqueurs moléculaires

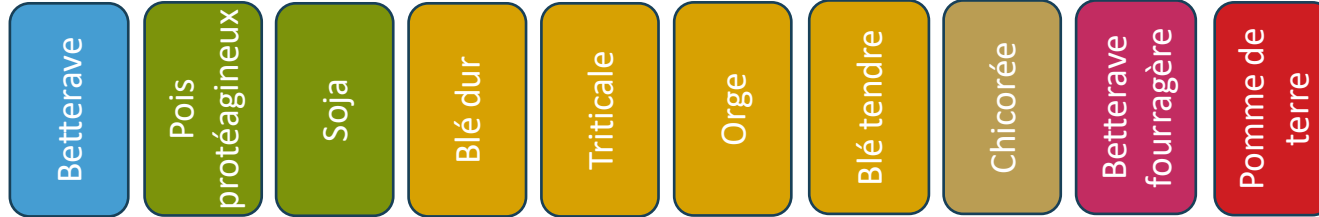
Développement de méthodes d'analyses

Développement d'outils de stockage et d'exploitation des données



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique et Statistique

Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

Développement de marqueurs moléculaires

Développement de méthodes d'analyses

Développement d'outils de stockage et d'exploitation des données



Cas de sélection variétale: en théorie

- Situation: une lignée de blé a été observée en champ comme résistante à une maladie
- Problème: son rendement n'est pas suffisamment bon



Resistance +++
Rendement -



Cas de sélection variétale: en théorie

- Situation: une lignée de blé a été observée en champ comme résistante à une maladie
- Problème: son rendement n'est pas suffisamment bon
- Solution: Introgresser le gène de résistance de la lignée observée dans une lignée élite ayant un meilleur rendement





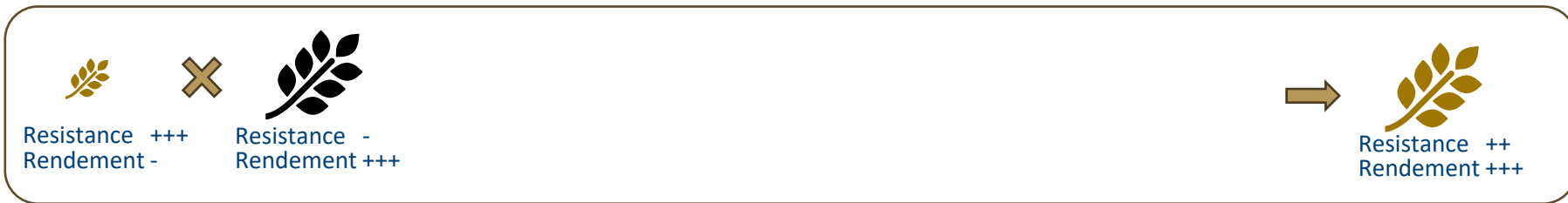
Cas de sélection variétale: en théorie

- Situation: une lignée de blé a été observée en champ comme résistante à une maladie
- Problème: son rendement n'est pas suffisamment bon
- Solution: Introgresser le gène de résistance de la lignée observée dans une lignée élite ayant un meilleur rendement.





Cas de sélection variétale: en pratique



- Processus qui nécessite:
 - Du temps: plusieurs générations sont nécessaires



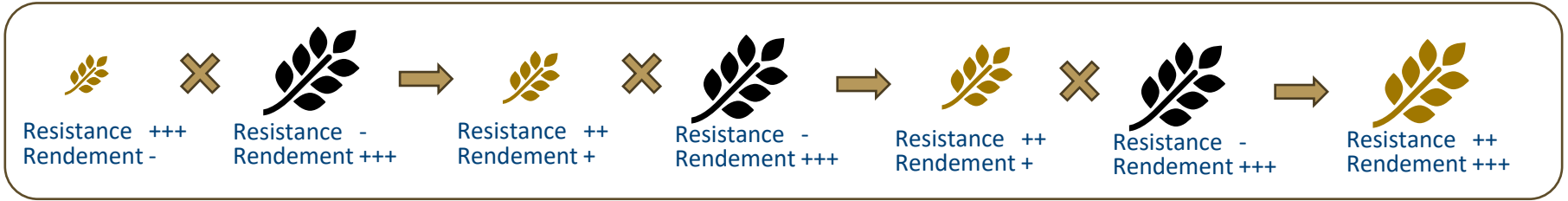
Cas de sélection variétale: en pratique



- Processus qui nécessite:
 - Du temps: plusieurs générations sont nécessaires



Cas de sélection variétale: en pratique



- Processus qui nécessite:
 - Du temps: plusieurs générations sont nécessaires



Cas de sélection variétale: en pratique



- Processus qui nécessite:

- Du temps: plusieurs générations sont nécessaires
- Des ressources:
 - plusieurs descendants par croisements,
 - observation et phénotypage des plantes pour choisir les meilleurs candidats à chaque génération



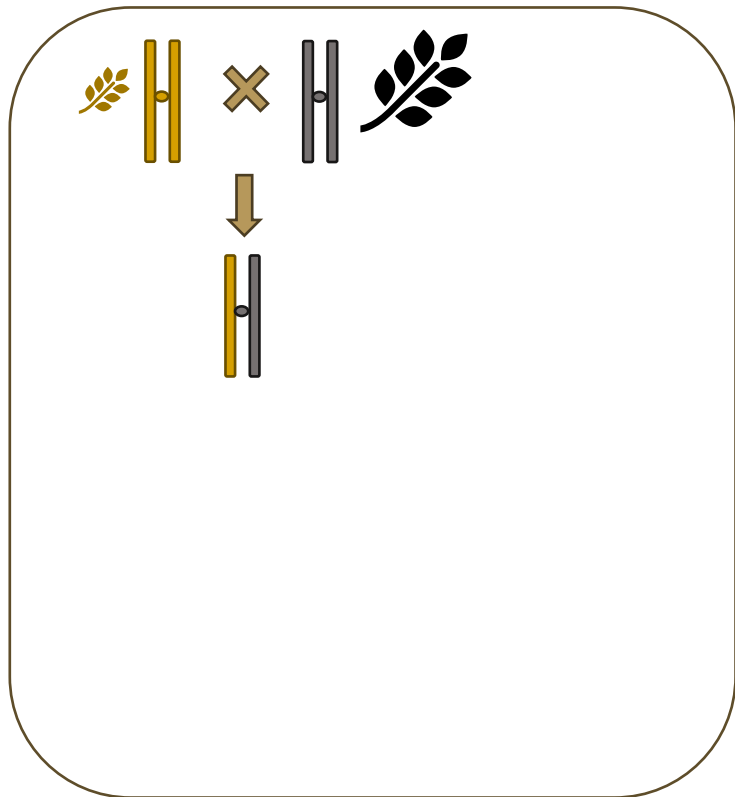
Cas de sélection variétale: en pratique



- Processus qui nécessite:
 - Du temps: plusieurs générations sont nécessaires
 - Des ressources:
 - plusieurs descendants par croisements,
 - observation et phénotypage des plantes pour choisir les meilleurs candidats à chaque génération
- Aide à la sélection via marqueurs moléculaires:
 - Sélection des individus par marqueurs génétiques
 - Diminution du nombre de croisements nécessaire pour purifier une lignée lors d'une introgression
 - Utilisation de marqueurs génétiques sur des plantules au lieu d'attendre les phénotypes



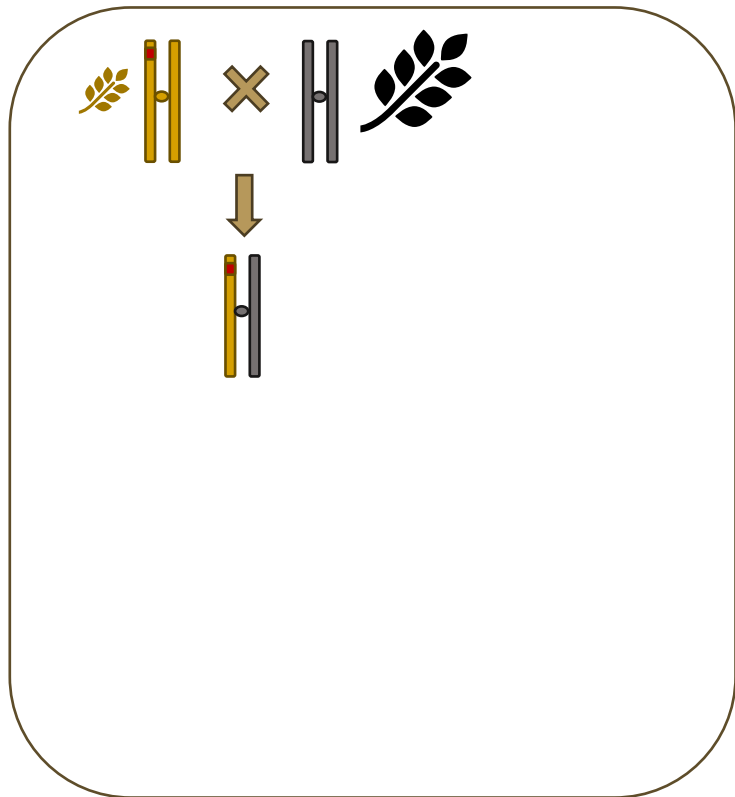
Cas de sélection variétale: en pratique



- 2 types de marqueurs nécessaires
 - Marqueurs associés au fond génétique de la lignée élite vs la lignée résistante



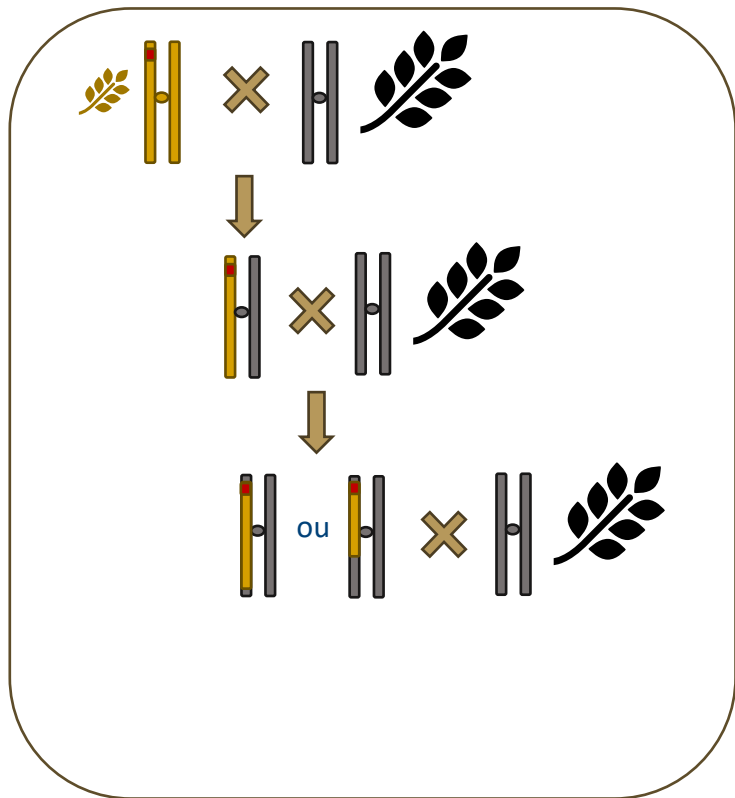
Cas de sélection variétale: en pratique



- 2 types de marqueurs nécessaires
 - Marqueurs associés au fond génétique de la lignée élite vs la lignée résistante
 - Marqueurs associés au gène de résistance



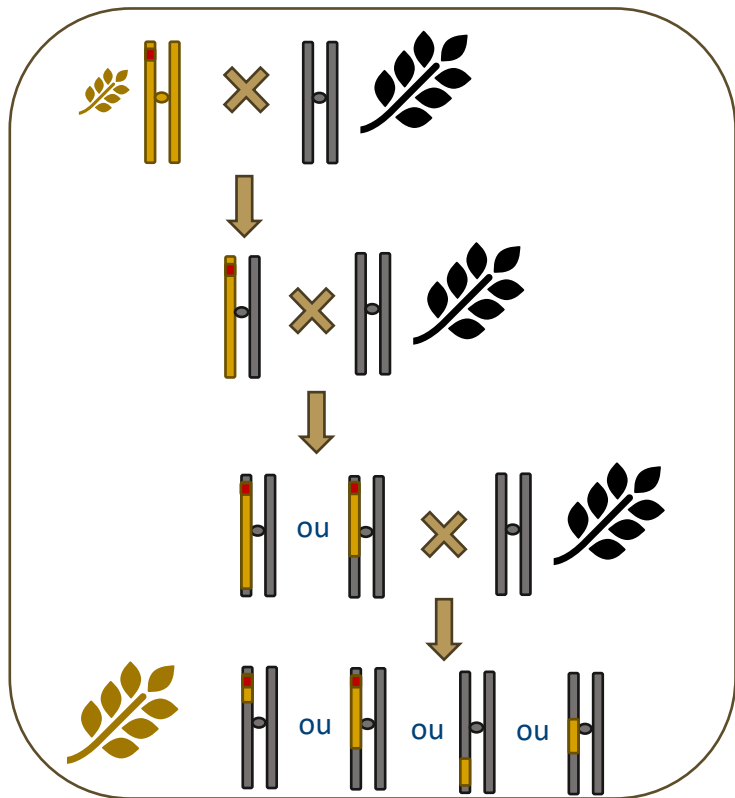
Cas de sélection variétale: en pratique



- 2 types de marqueurs nécessaires
 - Marqueurs associés au fond génétique de la lignée élite vs la lignée résistante
 - Marqueurs associés au gène de résistance
- => sélectionner rapidement des individus à chaque génération portant la résistance en ayant le plus court fragment du parent résistant



Cas de sélection variétale: en pratique



- 2 types de marqueurs nécessaires
 - Marqueurs associés au fond génétique de la lignée élite vs la lignée résistante
 - Marqueurs associés au gène de résistance
- => sélectionner rapidement des individus à chaque génération portant la résistance en ayant le plus court fragment du parent résistant



Sélection d'individus par marqueurs moléculaires

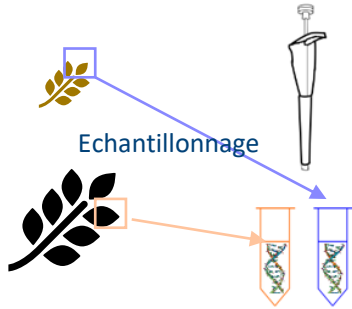
- En routine, 2 millions d'échantillons sont génotypés chaque année
 - Méthode de génotypage: Kaspar
 - Marqueurs utilisés: marqueurs associés à des traits
 - Analyse des données: Par les biologistes en utilisant des outils bioinfo

Type	Coût	Rapidité des résultats	#Marqueurs	Utilisation	Analyses
KASPar	€	+++	10-200	Routine pour les analyses de sélection	Labo



Sélection d'individus par marqueurs moléculaires

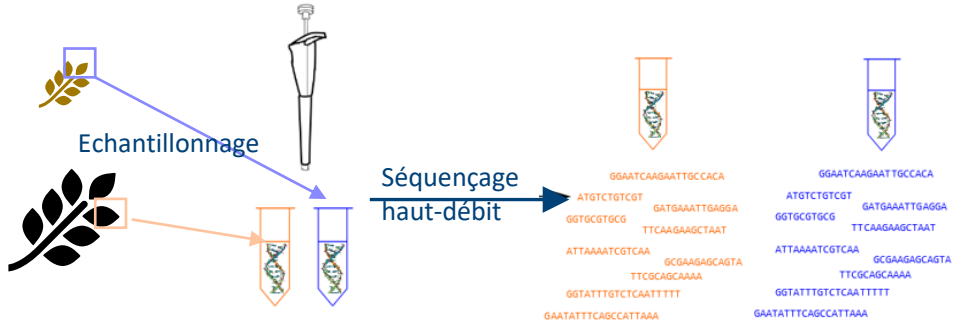
- Comment les marqueurs Kaspar sont-ils développés ?





Sélection d'individus par marqueurs moléculaires

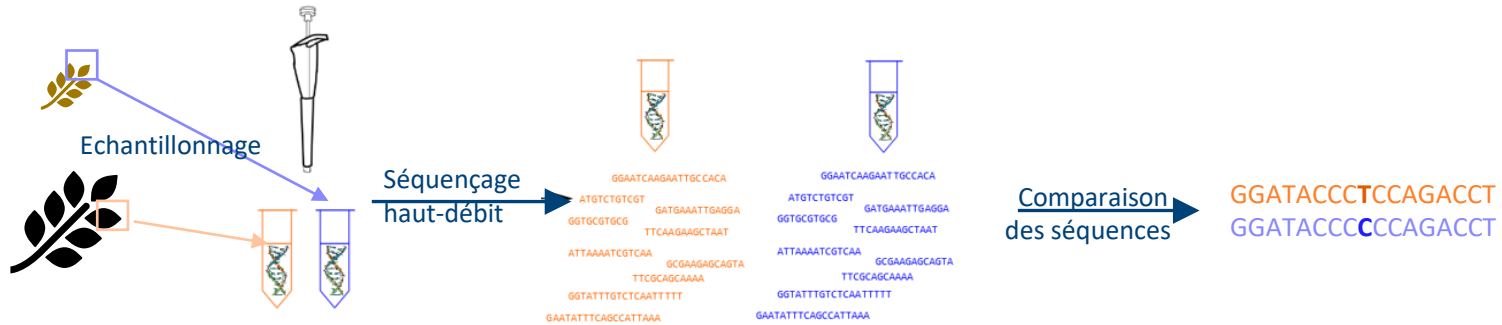
- Comment les marqueurs Kaspar sont-ils développés ?





Sélection d'individus par marqueurs moléculaires

- Comment les marqueurs Kaspar sont-ils développés ?





• Analyses bio-informatiques pour le développement de marqueurs moléculaires

- Assemblage de génomes de référence
- Annotation génique des génomes de référence
- Alignement des données de reséquençage sur les génomes de références
- Analyse des données de reséquençage et application de filtres pour sélectionner des marqueurs spécifiques
- Création de pipeline d'analyse

Type	Coût	Rapidité des résultats	#Marqueurs	Utilisation	Analyses
Reséquençage	€€€	+	Génome complet	Développement de marqueurs	Bioinfo



Utilisation des marqueurs développés par séquençage haut-débit

- Analyses de routine: Kaspar
- Activités de recherche
 - Puce à ADN
 - Recherche de marqueurs associés aux traits d'intérêt (GWAs)
 - Sélection génomique

Type	Coût	Rapidité des résultats	#Marqueurs	Utilisation	Analyses
KASPar	€	+++	10-200	Routine pour les analyses de sélection	Labo
Puce à ADN	€€	++	20K-30K	Sélection génomique, GWAs	Bioinfo
Reséquençage	€€€	+	Génome complet	Développement de marqueurs	Bioinfo



- Innovation

- Projets d'utilisation de nouvelles technologies de séquençage et/ou génotypage
 - Séquençage long-reads
 - Séquençage NGS de régions cibles (baits, Cas9)
 - Génotypage par NGS ciblés
- Projets sur l'exploration de pangénomes



- Innovation

- Projets d'utilisation de nouvelles technologies de séquençage et/ou génotypage
 - Séquençage long-reads
 - Séquençage NGS de régions cibles (baits, Cas9)
 - Génotypage par NGS ciblés
- Projets sur l'exploration de pangénomes

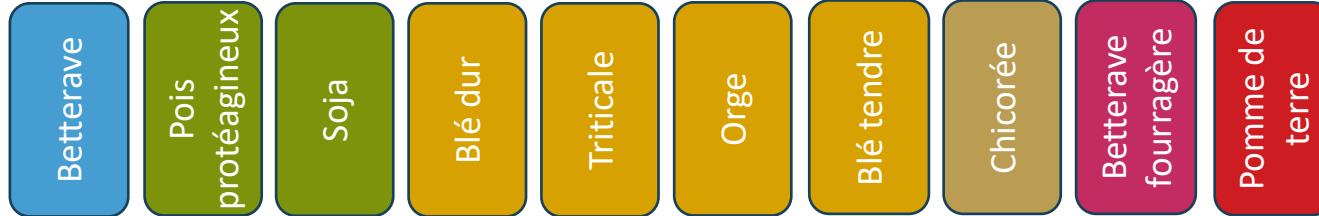
- Facteurs limitants

- Stockage des données (>Tb)
- Temps d'analyse sur de gros jeux de données



SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique, Statistique et Génétique

Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

Développement de marqueurs moléculaires

Développement de méthodes d'analyses

Développement d'outils de stockage et d'exploitation des données



Développement de méthodes d'analyses

- Algorithmes pour les analyses de sélection génomique
 - Modèle prédictif basé sur une population d'entraînement dont le génotype et le phénotype sont connus
 - Prédiction des performances d'un set d'individus, en se basant sur leurs données de génotypage et des données de la population d'entraînement



Développement de méthodes d'analyses

- Algorithmes pour les analyses de sélection génomique
 - Modèle prédictif basé sur une population d'entraînement dont le génotype et le phénotype sont connus
 - Prédiction des performances d'un set d'individus, en se basant sur leurs données de génotypage et des données de la population d'entraînement
- Amélioration du phénotypage
 - Phénotypage est une donnée capitale pour une sélection efficace



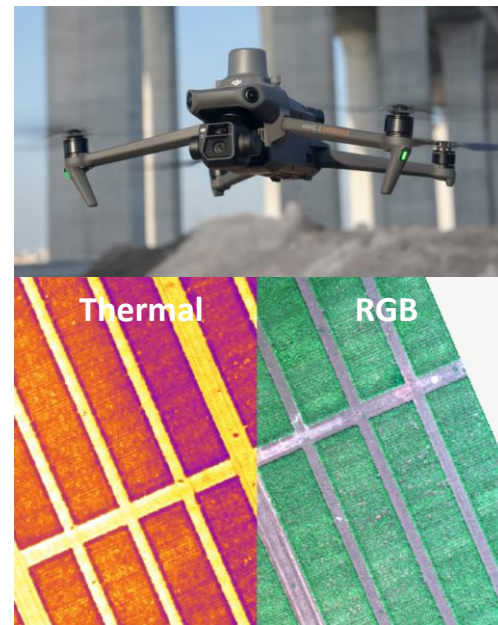
Développement de méthodes d'analyses

- Algorithmes pour les analyses de sélection génomique
 - Modèle prédictif basé sur une population d'entraînement dont le génotype et le phénotype sont connus
 - Prédiction des performances d'un set d'individus, en se basant sur leurs données de génotypage et des données de la population d'entraînement
- Amélioration du phénotypage
 - Phénotypage est une donnée capitale pour une sélection efficace
 - Harmonisation des relevés de données de phénotypage



Développement de méthodes d'analyses

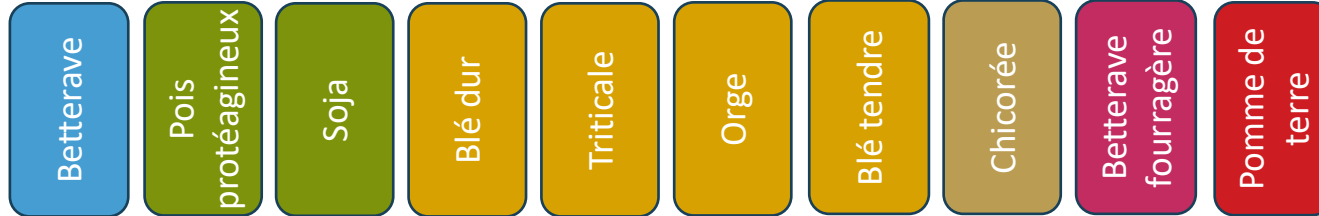
- Algorithmes pour les analyses de sélection génomique
 - Modèle prédictif basé sur une population d'entraînement dont le génotype et le phénotype sont connus
 - Prédiction des performances d'un set d'individus, en se basant sur leurs données de génotypage et des données de la population d'entraînement
- Amélioration du phénotypage
 - Phénotypage est une donnée capitale pour une sélection efficace
 - Harmonisation des relevés de données de phénotypage
 - Développement d'algorithmes pour le phénotypage haut débit via drones
 - Suivi de germination, vigueur, sécheresse, photosynthèse, maladies, ...





SELECTION VARIETALE

Création de nouvelles variétés



SUPPORT DES EQUIPES DE RECHERCHE

Développement d'outils d'aide à la sélection

SUPPORT DES EQUIPES DE RECHERCHE: Bioinformatique, Statistique et Génétique

Différentes équipes & Différents espèces

- Objectifs communs
- Technologies et outils communs
- Harmoniser les outils et les données

Développement de marqueurs moléculaires

Développement de méthodes d'analyses

Développement d'outils de stockage et d'exploitation des données



Développement d'outils de stockage et d'exploration des données

- Données de reséquençage
 - Projet de stockage et de développement de pipeline d'analyse dans Azure



Développement d'outils de stockage et d'exploration des données

- Données de reséquençage
 - Projet de stockage et de développement de pipeline d'analyse dans Azure
- Données de génotypage
 - Développement, pour les sélectionneurs, d'outils de visualisation de larges jeux de données génomiques



Développement d'outils de stockage et d'exploration des données

- Données de reséquençage
 - Projet de stockage et de développement de pipeline d'analyse dans Azure
- Données de génotypage
 - Développement, pour les sélectionneurs, d'outils de visualisation de larges jeux de données génomiques
- Données de phénotypage
 - Faciliter l'accès aux données
 - Accessible sur tablette pour un accès par les sélectionneurs et leurs équipes pour entrer les données
 - Accessible sur PC pour un accès par tous pour exploiter les données



Développement d'outils de stockage et d'exploration des données

- Données de reséquençage
 - Projet de stockage et de développement de pipeline d'analyse dans Azure
- Données de génotypage
 - Développement, pour les sélectionneurs, d'outils de visualisation de larges jeux de données génomiques
- Données de phénotypage
 - Faciliter l'accès aux données
 - Accessible sur tablette pour un accès par les sélectionneurs et leurs équipes pour entrer les données
 - Accessible sur PC pour un accès par tous pour exploiter les données
- Ensemble des données:
 - Développement et gestion de bases de données



Développement d'outils de stockage et d'exploration des données

FloriBase 18WWBAL-001 Expérimentation Batt. 100% CJ

Bande	Passes	Var	Rep	Dénomination	Tem	ObsBT
Bloc	Gauche			REP	NVAR	DT
1	2	4	1	Klein Serpiente	4	1 n.
1	3	1	1	Aca 356	1	1 n.
1	4	13	1	FD17WW0635	4	n.
1	5	8	1	Alhambra	8	2 n.
1	6	2	1	Aca 360	2	n.
1	7	14	1	FD17WW0640	7	n.

DT 22/11/2018 36/12/2018

Desc App Epuration Résumé Valider

ARIST	AspPr	AspHt	GLBt	DisjBT	Cyclo	EpiBT 1	EPIBT 2	Divers 1	Divers 2
nb	TT	TC	V	b/nb	CC	FERT	BS	Sain	Rang
ar	T	C	A	Disj Pr	CC - Cl	CPct	EpiS	PdF	-> V
b	1/2 T	AC	GL	Disj Pr+	Cl	LAX	Erx	EdF	Verse+
b/nb	1/2 P	AH	GL	Disj Ht	Cl - CL	Crb	BStr	JBF	
B+	P	H	GL+	Disj Ht+	Cl	Velu	PM	CrT	
B=	TP	TH	EVT GL	Disj Gl	LG	Supp	Non velu	TPHy	
B-	DH P	DH H	Disj GL	Disj Velu		Grillé		Stressé	

Généalogie = FD02112-3 x Baguette 10 = (Cm1-44 x Frelon) x Baguette 10
 CSe=03439-SHD-63, CFam=FD08116

FloriBase EG / Ellen Goudemand

30/11/2018 11:19 Largeur 1000px Hauteur 900px

Individus Phéno/Génotypage Administration

utilisateurs connectés

Analyse de l'essai 18WWCAP-101

Expérimentation	18WWCAP-101	2018 / Blé tendre / Cappelle
Dispositif	Essai Bloc 24 x 2	Rep. TR : 12 Surf/Parc = 7.50
Description	ESSAI VARIETES	
Objectif		

Rendement 'traité'

Rendement 'non traité'

Répétitions = 12 (soit 2 rép.)
 48 notes théoriques
 48 notes en base
 OK:48 MQ:0 Err:0

Analyse modèle fixe en Bloc (Pst1542/AMFB1):
 ETR = 3.49772
 (ETR correct si <= 5)
 Test d'homosédastéité (Bartlett) validé (0.2164, 1)
 Test de normalité (Shapiro-Wilk) validé (0.3175)
 Pas de résidus suspects (Grubbs) (0.06678)
 Nombre de résidus mis à jour : 48 (notes)

Analyse modèle fixe en Bloc (Pst1542/AMFB2):
 Résultat de l'anova:
 Anova Table (Type III tests)
 Responses: notes

Sum Sq Df	F value	Pr(>F)
(Intercept) 33797.7	14530.8578	< 2.2e-16 ***
DEMO	1315.23	4.4735 0.002336 ***
REP	1	0.0906 0.766057
Residuals	281.23	

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.'
 Nombre de moyennes ajustées générées : 24
 nombre de notes de Tukey générées : 24

FloriBase Dev 18WWCAP-SDS-01 Expérimentation Batt. ? JMA

Annuler

Agitation

Serie n°1 5.00 Farine + eau

Préparation Farine

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17	18WWCAP-101 Npp. 3 Ccm. 17

2 (9-16)

3617792 3617793 3617766 3617767 3617762 3617763 3617774 3617775

MQ Eff Valider

Expérimentation 18WWCAP-101 Affichage : [Minimum (Var)] Paramètres Les notes

Année / Espace / Site	2018 / Blé tendre / Cappelle	
Dispositif	Essai Bloc, 24 x 2	
Caractère	Rendement / Rdt T+NT RES_T	
Notation	27/07/2018	

Répétition 2	5	18	24	12	6	11	17	4	10	23	22	15	3	16	13	19	1	7	8	21	
Répétition 1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21

-7.01 -6.08 -5.14 -4.21 -3.27 -2.34 -1.40 -0.47 0.47 1.40 2.34 3.27 4.21 5.14 6.08 7.01

Analyse de l'essai 18WWCAP-101 / Rendement

Ordre	N°Var	Dénomination	Rdt.T	Class.	%T	Tukey
1	10	Pilon	118.35	1	110.74	A
2	17	FD15WW0065	116.23	2	108.76	AB
3	22	FD16WW0096	114.26	3	106.91	ABC
12	9	Complice	113.27	4	105.99	ABCD
3	13	Rubisko (T3)	113.25	5	105.97	ABCD
18	17	FD15WW0097	110.55	6	103.44	ABCDE
9	10	Complice	110.49	7	103.38	ABCDE
8	10	Bienfait	109.43	8	102.39	ABCDE
13	10	Concrète	108.53	9	101.55	ABCDE
19	10	FD15WW0147	108.30	10	101.34	ABCDE
7	10	Collector	108.09	11	101.14	ABCDE
6	10	Calumet	107.21	12	100.32	ABCDE
1	10	Celulle (T1)	107.15	13	100.26	ABCDE
16	10	FD15WW0063	106.66	14	99.80	ABCDE
14	10	Unik	106.59	15	99.74	ABCDE
24	10	Argument	105.41	16	98.63	ABCDE
2	10	Fructidor (T2)	104.11	17	97.42	ABCDE
15	10	Pilier	103.29	18	96.65	BCDE



MERCI DE VOTRE ATTENTION

QUESTIONS ?