

# Rapport de stage

Baptiste ROUGER

L2 BCST

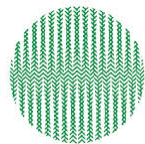
Analyse de la variation de la taille des plantes  
lors d'une expérience de sélection divergente  
pour la date de floraison sur le maïs

21 mai 2015



**Le Moulon**  
**GÉNÉTIQUE**  
**QUANTITATIVE**  
**ET ÉVOLUTION**

UMR de Génétique Quantitative et Evolution - Le Moulon  
Ferme du Moulon  
91190 Gif-sur-Yvette



Institut National de la Recherche Agronomique

**INRA**



UNIVERSITÉ  
**PARIS**  
**SUD**

Directeur de recherche : Olivier MARTIN

Maître de stage : Christine DILLMANN

Enseignant référent : Judith LEGRAND

# Table des matières

Glossaire	0.1 Introduction	3
0.1.1 Le maïs, une espèce modèle d'intérêt agronomique . . . . .		3
Généralités sur le maïs . . . . .		3
Histoire du maïs . . . . .		3
0.1.2 Présentation de l'expérience de sélection divergente . . . . .		4
Généralités . . . . .		4
L'expérience de sélection divergente . . . . .		4
0.1.3 Objectifs du stage . . . . .		4
<b>0.2 Matériel et méthodes</b>		<b>5</b>
0.2.1 Le matériel végétal . . . . .		5
0.2.2 Méthodes . . . . .		5
Mise en forme des données . . . . .		6
Correction des données . . . . .		6
0.2.3 Résultats . . . . .		7
Réponse de la hauteur des plantes à la sélection pour la date de floraison .		7
Corrélation entre hauteur et date de floraison . . . . .		8
<b>.1 Codes sources</b>		<b>10</b>

# Table des figures

1	Comparaison entre téosinte et maïs . . . . .	3
2	Cycle de vie du maïs en culture . . . . .	4
3	Réponse des deux lignées pures de maïs à la sélection pour la date de floraison	5
4	Principe de l'expérience de sélection divergente . . . . .	6
5	Généalogie des plantes sélectionnées chez les deux lignées de maïs F252 et MBS pour l'expérience de sélection divergente en 2012 . . . . .	6
6	Représentation des hauteurs des plantes de la lignée F252 . . . . .	7
7	Représentation des hauteurs des plantes de la lignée MBS . . . . .	7
8	Représentation de la relation entre hauteur et date de floraison . . . . .	8
9	Représentation séparée de la corrélation entre hauteur et date de floraison pour deux familles . . . . .	9
10	Représentation de la corrélation entre hauteur et date de floraison par individus en 2013 et 2014 . . . . .	9

# Glossaire

**Degré-jour** : unité de mesure utilisée pour calculer l'accumulation de chaleur nécessaire à la durée d'un développement biologique comme la croissance d'une plante

**Dérive génétique** : réduction du nombre de génotypes différents dans une population de taille finie du fait de l'échantillonnage aléatoire des gamètes ou des individus

**Monoïque** : se dit d'une plante dont les fleurs mâles et femelles sont distinctes mais situées sur le même pied

**Sélection naturelle** : pression évolutive qui conduit à des changements de la composition génétique d'une population d'une génération à l'autre, en lien avec des différences génétiques des capacités de survie ou de reproduction

## 1 Introduction

### 1.1 Le maïs, une espèce modèle d'intérêt agronomique

#### 1.1.1 Généralités sur le maïs

Le maïs (*Zea mays*) est une plante herbacée annuelle appartenant à la famille des *Poacées* (graminées). Par ailleurs, il s'agit d'une plante monoïque. Lors de sa domestication (cf 0.1.1), le maïs a vu certaines de ses caractéristiques changer pour le voir devenir ce qu'il est aujourd'hui. On peut par exemple citer l'absence de dissémination de ses graines, l'augmentation de leur poids et de leur nombre, la synchronisation des dates de floraison mâles et femelles ainsi qu'une dominance apicale entraînant une modification de l'architecture de la plante.

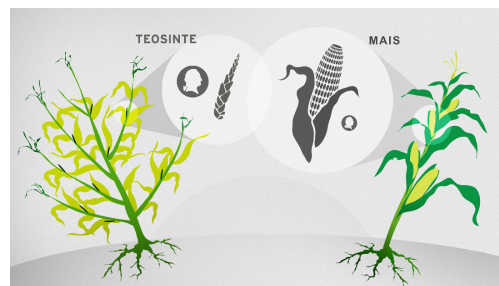


FIGURE 1 – Comparaison entre téosinte et maïs

**Cycle de vie du maïs** Le cycle de vie du maïs (FIGURE 2) commence par la phase de dormance du grain, puis sa germination. Vient ensuite le stade végétatif jeune durant lequel la plante développe ses feuilles juvéniles. Il s'achève par la transition florale qui marque le début du stade végétatif adulte. Celui-ci se termine par la floraison débutant le stade reproducteur. Pour finir, lorsque les graines sont mûres, celles-ci peuvent être récoltées.

#### 1.1.2 Histoire du maïs

Le maïs est une espèce domestiquée depuis environ 9000 ans. Elle est originaire des alentours de Mexico et est issue de la téosinte (*Zea sp.*). Elle s'est rapidement propagée à

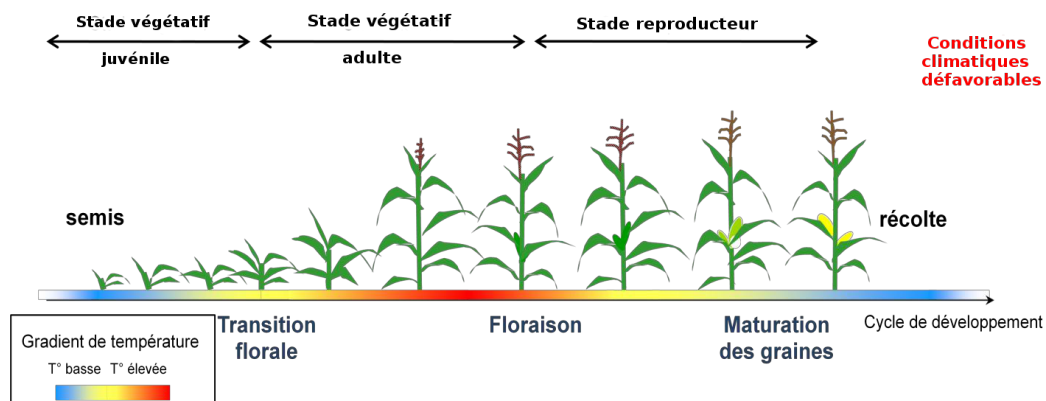


FIGURE 2 – Cycle de vie du maïs en culture

travers le continent américain[1] en quelques milliers d’années. Cette expansion demanda son adaptation à des climats plus tempérés que son environnement naturel tropical, ce qui fut réalisé en sélectionnant les variétés insensibles à la photopériode, relativement précoces et capables d’effectuer leur cycle de croissance avant l’arrivée des conditions climatiques défavorable (FIGURE 2). Elle fut découverte et ramenée pour la première fois en Europe par Christophe Colomb lors de sa découverte de l’Amérique. C’est aujourd’hui une des premières céréales cultivées dans le monde.

## 1.2 Présentation de l’expérience de sélection divergente

### 1.2.1 Généralités

**Des forces évolutives** Les individus d’une population interagissent de façon dynamique avec leur environnement. Ces interactions conduisent à des changements de composition génétique des populations sous l’effet des pressions évolutives. *Développer...*

**L’adaptation** Il s’agit d’un phénomène résultant des forces évolutives, telle que la sélection. Elle témoigne de la plus grande capacité des individus adaptés à survivre dans leur environnement et à s’y reproduire.

### 1.2.2 L’expérience de sélection divergente

L’expérience de sélection divergente a débuté en 1993. Prenant pour origine deux lignées pures de maïs (F252 et MBS847) à très forte homozygotie, les plants les plus précoces et les plus tardifs pour la date de floraison furent sélectionnés pour le semis chaque année. On obtient ainsi deux populations de maïs pour une même lignée pure, marquant une plus ou moins forte réponse à la sélection selon la variété et la direction de la sélection (FIGURE 3).

Par ailleurs, la taille des plantes fut aussi mesurée, pour observer si, comment, et combien un caractère n’intervenant pas dans la sélection artificielle d’un individu pouvait être modifié par celle-ci.

## 1.3 Objectifs du stage

Une analyse fine des paramètres de taille pour chaque population de maïs doit être faite pour déterminer dans quelle mesure les plantes voient leur hauteur varier par la

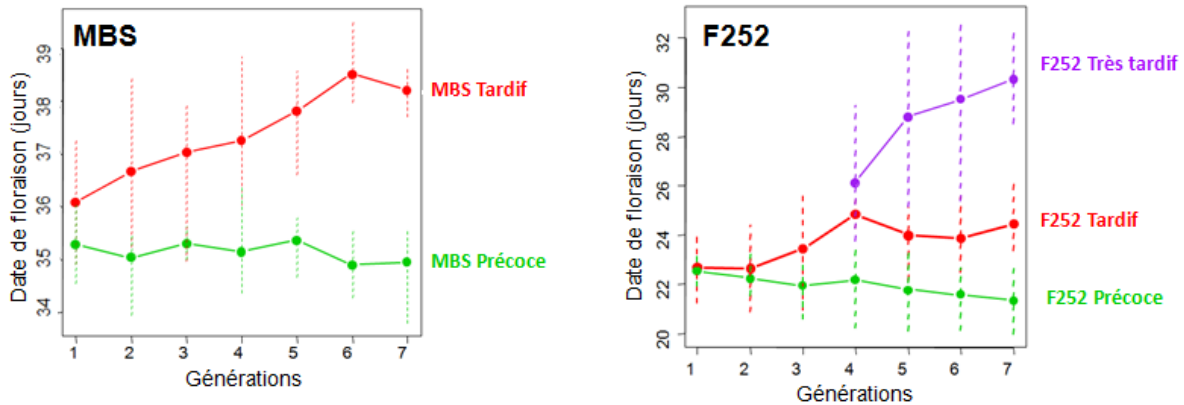


FIGURE 3 – Réponse des deux lignées pures de maïs à la sélection pour la date de floraison

sélection faite sur la date de floraison et, le cas échéant, déterminer quelle pourrait être la source de ces variations (mutations, dérive génétique...).

Après nous être intéressés au matériel utilisé lors des vingt années d'expérimentation, nous nous intéresserons aux méthodes de traitement des données. Nous analyserons ensuite les résultats et les interpréterons. Enfin, nous conclurons ce sujet.

## 2 Matériel et méthodes

### 2.1 Le matériel végétal

L'expérience de sélection divergente (FIGURE 4), débutée en 1993 et toujours en cours, compte 19 générations de maïs. Les lignées de départ utilisées sont deux lignées pures à forte homozygotie : les maïs F252 et MBS847 (notée MBS). On obtient pour chaque lignée deux populations, une tardive et une précoce, sélectionnées pour la date de floraison<sup>1</sup> notée en degrés-jours.

A chaque génération, on réalise pour chaque lignée une autofécondation<sup>2</sup> des dix individus les plus précoces et les plus tardifs. Pour pouvoir retracer le pedigree de chaque individu semé et sélectionné, on réalise des cartes généalogiques (FIGURE 5) des lignées de maïs. Les individus sélectionnés à la dernière génération sont indiqués en couleur, ainsi que tous leurs descendants.

### 2.2 Méthodes

Pour réaliser les analyses nécessaires à la confirmation de nos hypothèses, il faut commencer par formater toutes les données au même format, c'est à dire que chaque plante mesurée sera référencée avec de nombreuses caractéristiques la concernant : l'année, la génération, la lignée (F252 ou MBS), le parent (c'est-à-dire le « grand-parent » de la plante mesurée), le progéniteur (le parent de la plante), la population à laquelle elle appartient

1. Définie lorsque les fleurs mâles et femelles apparaissent

2. Pour réaliser l'autofécondation, on couvre les fleurs des individus de sachets pour recueillir le pollen et protéger les soies. Après vingt-quatre heures, le sachet contenant le pollen est déposé sur la fleur femelle pour réaliser l'autofécondation.

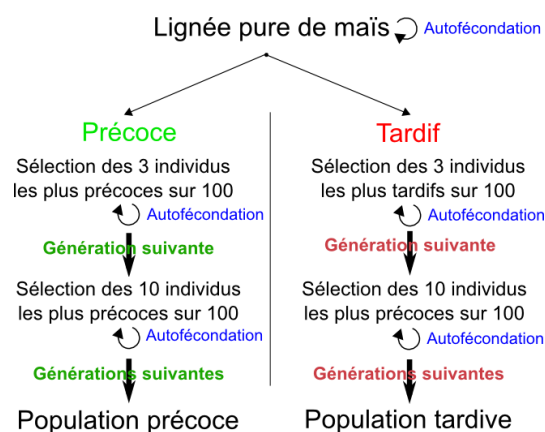


FIGURE 4 – Principe de l'expérience de sélection divergente

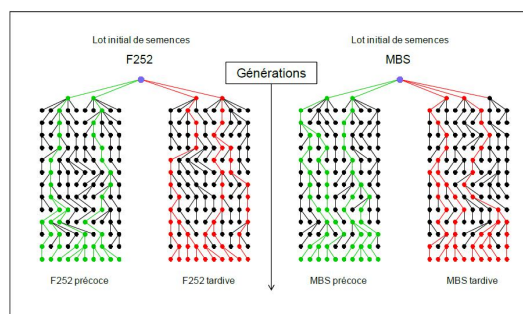


FIGURE 5 – Généalogie des plantes sélectionnées chez les deux lignées de maïs F252 et MBS pour l'expérience de sélection divergente en 2012

(tardive, précoce ou témoin), la famille qui permet de retracer la généalogie, le bloc, la ligne et le numéro de la plante qui permettent de retrouver la position de la plante dans le champs, et bien évidemment la hauteur de la plante.

### 2.2.1 Mise en forme des données

L'ensemble des données que l'on doit rassembler se trouve dans des fichiers de formats différents. Il faut donc les mettre au même format pour commencer à traiter les données. On choisi le format **CSV** pour des raisons de commodité. Ces fichiers sont ensuite traités par le script<sup>3</sup> *tri.R*. Celui-ci retourne alors un fichier de données formatées par année. Puis le script *aggregate.R* permet de rassembler toutes ces données dans un même fichier au format **CSV**.

### 2.2.2 Correction des données

L'impact de l'année et de l'emplacement dans le champ (appelé « effet bloc ») sur la hauteur des plantes a ensuite été constaté . En effet, il apparaît clair que l'année va jouer un rôle important sur la hauteur de la plante, en fonction de la quantité de précipitations, de la température cumulée, etc. De la même façon, les propriétés du terrain peuvent varier localement, selon par exemple la composition du sol, ce qui va là encore influencer sur la hauteur des plantes. Un script<sup>4</sup> a donc été réalisé pour estimer et corriger les données récoltées.

Ce script réalise une analyse multifactorielle (effet de l'année et « effet bloc ») de la variance de la hauteur. On utilise ensuite les résultats de cette analyse pour corriger dans un premier temps l'« effet bloc », et dans un second temps l'effet année.

3. Tous les scripts sont réalisés dans le langage interprété par le logiciel d'analyses statistiques **R**. cf ANNEXES .1

4. Il s'agit du script *DataCorrectv2.R*

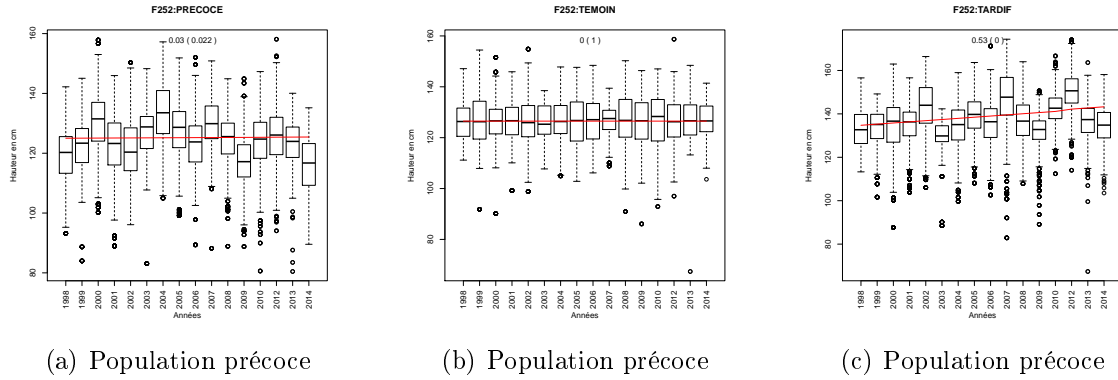


FIGURE 6 – Représentation des hauteurs des plantes de la lignée F252

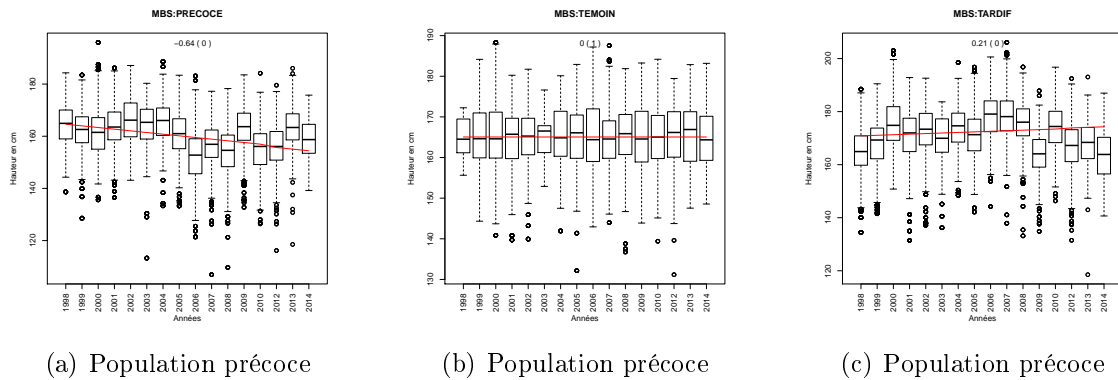


FIGURE 7 – Représentation des hauteurs des plantes de la lignée MBS

## 2.3 Résultats

### 2.3.1 Réponse de la hauteur des plantes à la sélection pour la date de floraison

Grâce aux deux graphiques suivants ( FIGURE 6(c) et FIGURE 7(c)), on peut conclure que les données peuvent être corrigées, puisque les témoins ont tous deux une droite de régression dont le coefficient directeur est de 0.

Par ailleurs, on constate que la variation de taille, bien que significative, est relativement faible au fil des années ( $< 1$  cm/an) à partir de 1998. Cela semble indiquer qu'une réponse à la sélection très rapide a eu lieu, laissant ensuite la place à une réponse beaucoup plus faible, voire insignifiante.

On observe également<sup>5</sup> des tendances à la hausse pour les populations tardives (0.53 cm/an et 0.21 cm/an respectivement pour F252 et MBS) et une tendance à la baisse pour les populations précoces de MBS (-0.64 cm/an). Les populations précoce de F252 semblent cependant augmenter de taille, bien que très faiblement (0.03 cm/an).

---

5. FIGURE 6 et FIGURE 7

### 2.3.2 Corrélation entre hauteur et date de floraison

**Corrélation par famille et par année** Il a également été testée si il y avait bien une corrélation entre la hauteur des plantes et la date de floraison (FIGURE 8). Pour cela, on représente pour chaque famille et chaque année la hauteur moyenne des plantes en fonction de leur date moyenne de floraison.

On peut voir que les points semblent globalement être répartis autour d'une droite, ce qui confirme la corrélation entre hauteur et date de floraison.

On peut également observer une scission entre les populations précoces et tardives des lignées (respectivement dans les couleurs bleu/vert et rouge en FIGURE 8).

On remarque cependant que, malgré que l'ensemble des points confirme la corrélation, certaines familles ne semblent pas suivre cette dernière. Par exemple, la famille F-39 (FIGURE 9) ne montre plus de corrélation, tandis que pour M-49 (FIGURE 9), elle apparaît conservée.

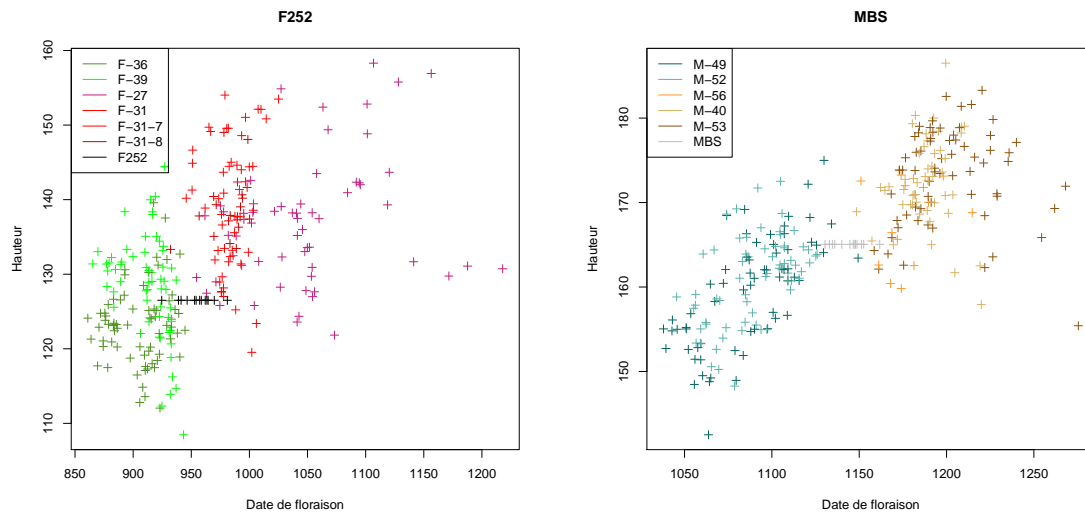


FIGURE 8 – Représentation de la relation entre hauteur et date de floraison  
Chaque point d'une couleur représente la moyenne de la hauteur des progéniteurs d'une famille pour une année (couleurs dans la figure) en fonction de la date moyenne de floraison de ces mêmes plantes. A gauche, la lignée F252 et à droite la lignée MBS.

**Corrélation par individu pour 2013 et 2014** On peut voir que ça fait des choses bizarres...

## Références

- [1] Zerjal, Rousselet, Mhiri, Combes, Madur, Grandbastien, Charcosset, and Tenailon. Maize genetic diversity and association mapping using transposable element insertion polymorphisms. 2011.



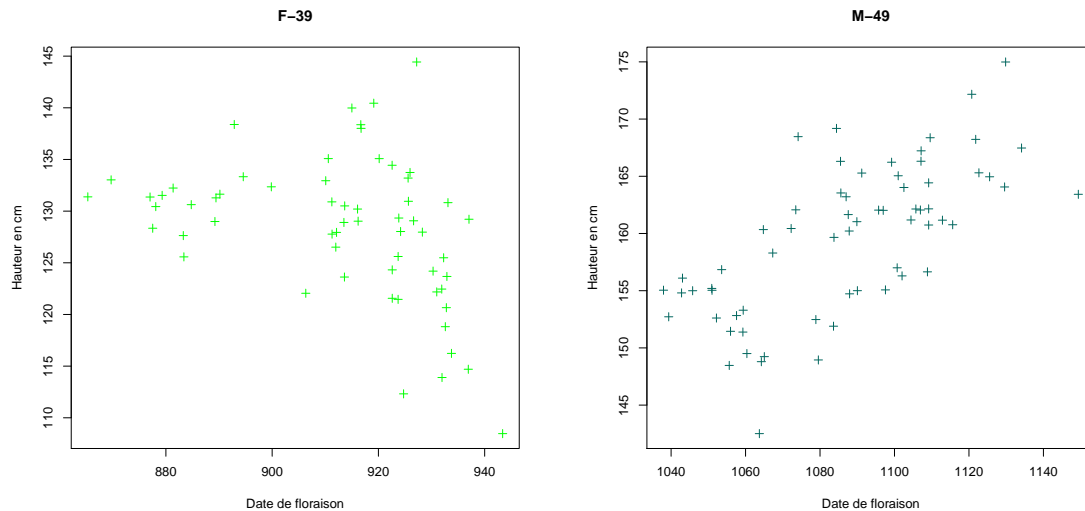


FIGURE 9 – Représentation séparée de la corrélation entre hauteur et date de floraison pour deux familles

Chaque point représente la moyenne de la hauteur des progéniteurs de la famille (F-39 à gauche et M-49 à droite) pour une année en fonction de la date moyenne de floraison de ces mêmes plantes.

## A Codes sources

Les codes sources utilisés pendant ce stage sont disponibles sur Github à l'adresse suivante : <https://github.com/rougerbaptiste/stageL2>.

Voici la fonction de chacun de ces codes :

**aggregate.R** : Ce script a pour fonction de rassembler les données contenues dans les fichiers de hauteurs de chaque année pour les rassembler dans un unique fichier.

**Datacorrectv2.R** : Ce script corrige les données des années et affiche un graphe (*Référence*) résumant les données.

**RprFam.R** : Ce script a pour fonction de représenter les données de hauteurs obtenues pour chaque famille (hauteur absolue et hauteur relative au témoin) et de créer un fichier .pdf contenant les deux graphes.

**RprPop.R** : Ce script a pour fonction de représenter les données de hauteurs obtenues pour chaque population (hauteur absolue et hauteur relative au témoin) et de créer un fichier .pdf contenant les deux graphes.

**tri.R** : Ce script a pour fonction de récupérer toutes les données brutes de hauteurs, de les mettre sous une forme standard et de créer un fichier contenant les données de hauteur par année.

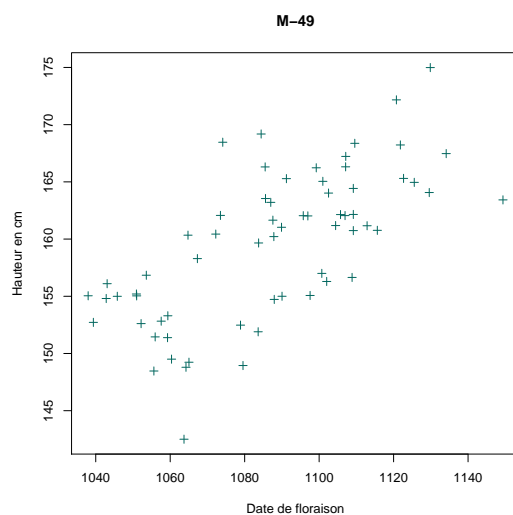


FIGURE 10 – Représentation de la corrélation entre hauteur et date de floraison par individus en 2013 et 2014