Université Montpellier II – Sciences et Techniques du Languedoc

THÈSE

pour obtenir le grade de
Docteur de l'Université Montpellier II

Discipline	Biologie des populations et écologie
École doctorale	Systèmes Intégrés en Biologie, Agronomie, Géosciences, Hydrosciences, et Environnement (SIBAGHE)

Évolution de la composition génétique du tissu nourricier de la graine

Double fécondation, polysporie, et empreinte parentale

par

Aurélie Cailleau

Soutenue publiquement le
13 décembre 2010

Jury

Frédéric BERGER	Professeur, Temasek Lifesciences, Singapour	Rapporteur
Pierre-Olivier CHEPTOU	Chargé de Recherche, CEFE Montpellier	Directeur de thèse
Bernard GODELLE	Professeur, Université Montpellier II	Examinateur
Thomas LENORMAND	Directeur de Recherche, CEFE Montpellier	Directeur de thèse
Denis ROZE	Chargé de Recherche, Station Biologique de Roscoff	Examinateur
Xavier VEKEMANS	Professeur, Université Lille 1	Rapporteur

Remerciements

Je tiens à remercier l'ensemble des personnes ayant contribué de près ou de loin à la réalisation de ce travail.

Un grand merci à mes directeurs de thèse, Thomas Lenormand et Pierre-Olivier Cheptou, pour m'avoir offert la possibilité d'atteindre l'objectif que je m'étais donné dès le début de mon cursus universitaire : réaliser un travail théorique en biologie de l'évolution. Merci pour la qualité de l'encadrement et la formation que j'ai reçues. Merci à Thomas pour m'avoir accompagnée dans les méandres des séries de Taylor et du QLE, merci à POC pour les discussions conceptuelles toujours riches et stimulantes.

Merci aux membres de mes comités de thèse : Denis Roze, Sylvain Glémin, Isabelle Goldringer, Daniel Grimanelli, Eric Imbert, et Sandrine Maurice, pour leurs contributions si pertinentes. La collaboration avec Daniel Grimanelli et son équipe (en particulier Daphné Autran) a été déterminante dans la réalisation d'une grande partie de ce travail. En particulier, l'expérimentation sur le maïs n'aurait pas été possible sans la contribution de Daniel, son expérience, ses contacts, et ses collaborations. Daniel et Daphné ont également eu la gentillesse de relire attentivement mon manuscrit, et m'ont encadrée dans l'exploration de la biologie de l'empreinte génétique, ce qui constitue la base de tout un chapitre de ma thèse (le chapitre 5).

J'ai eu la chance de bénéficier gracieusement, pour l'expérience avec le maïs, de nombreuses opportunités logisitiques : merci à Marty Sachs du Maize Genetic Cooperation Stock Center pour l'envoi des grains, merci au personnel de Syngenta pour avoir accueilli la manip, et merci au CIRAD pour le prêt de la compteuse à grains.

Merci à tous ceux qui m'ont donné de nombreux coups de main techniques et conseils ingénieux ou ont relu mon manuscrit de thèse : Élodie Blanchet, Christian Colin, Florence Débarre, David Degueldre, Céline Devaux, Julien Dutheil, Olivier Flores, Marie-Angélique Laporte, Patrick Mayet, François Munoz, et Nicolas Rode. Merci en particulier à Élodie qui s'est impliquée trois semaines dans la tâche ingrate d'égrainer, trier, compter et peser les grains de maïs, et à Florence qui m'a formée à de nombreux outils logiciels et a fait œuvre de dictionnaire français-anglais vivant pendant trois ans.

Merci aux membres de mon jury, Fréderic Berger, Bernard Godelle, Denis Roze, Xavier Vekemans, et mes directeurs de thèse, pour avoir accepté d'évaluer ce travail.

Merci aux étudiants et post-docs que j'ai cotoyés au CEFE, pour l'ambiance et la bonne humeur : Florence et Nicolas que j'ai cités, mais également Elodie Flaven, Federico Manna, Jonathan Dubois, Camille Guilbaud, Michel Thomman, Romain Gallet, Emeline Pettex, Catherine Soler, Pedro Vale, Jérôme Fort, et tous les autres.

Merci à mes proches, mes sœurs et leurs familles, mes parents, qui ont fait office de soupape lors de mes excursions dans l'Ouest et qui m'ont manifesté tout leur soutien le jour J. Merci à Saloum, qui est là pour le meilleur et pour le pire.

Merci enfin à toutes les personnes qui m'ont aidée à préparer ma soutenance, que je ne listerai pas pour éviter d'en oublier : c'est une phase critique et ce remerciement n'est pas des moindres.

Merci à tous.

Table des matières

Table des matières **iii**

1 Introduction générale **1**
1.1 Les tissus nourriciers de la graine 2
1.2 Scénarios de l'évolution des tissus nourriciers 14
1.3 Éléments théoriques 25
1.4 Orientation de la thèse 33

2 Développements de la théorie de l'hétérozygotie **35**
2.1 Introduction 35
2.2 Méthodologie 36
2.3 Évolution de la double fécondation et des contributions maternelles 38
2.4 Évolution de l'empreinte parentale 42
2.5 Éffet de l'autofécondation 44
2.6 Évolution du périsperme et de la polysporie 46
2.7 Hétérozygotie et évolution d'autres traits 53
2.8 Discussion générale et perspectives 56

3 Analyses comparatives **59**
3.1 Choix des traits et prédictions 59
3.2 Base de données 62
3.3 Tissus nourriciers et dépression de consanguinité 66
3.4 Tissus nourriciers et effets parentaux 70

4 Conflits mâle-femelle et allocation des ressources chez *Zea mais* **73**
4.1 Introduction 73
4.2 Matériel et méthodes 74
4.3 Résultats 76
4.4 Discussion 83

5 Quelle théorie pour l'évolution de l'empreinte parentale ? **85**
5.1 L'empreinte parentale et le conflit d'allocation des ressources 86
5.2 Les théories alternatives 94
5.3 Discussion 101

6 Conclusion générale **103**
6.1 Évaluation des théories 104
6.2 Discussion et perspectives 112

Bibliographie **115**

A Manuscrit : Ploidy and the Evolution of Endosperm of Flowering Plants **135**

B Manuscrit : Patterns of resource allocation and male-female conflict : a case study in maize **151**

C Manuscrit : Masking effect, inbreeding and the evolution of double fertilization **169**

D Manuscrit : Endosperm genetic make-up and inbreeding depression **179**

E Tables des valeurs sélectives de la partie modèles **197**

Liste des figures **201**

Liste des tables **202**

Liste des encadrés **202**

Chapitre 1

Introduction générale

Lorsque les êtres vivants se reproduisent, ils transmettent des ressources à leurs descendants. Il existe une diversité de modes d'allocation des ressources. Certains organismes allouent les ressources aux gamètes, avant la fécondation. C'est le cas par exemple des Cycas et Ginkgos parmi les végétaux, ou des insectes parmi les animaux (les ressources sont contenues dans l'ovule ou l'oeuf). D'autres organismes allouent des ressources aux embryons, après la fécondation. C'est le cas par exemple des plantes à fleurs (ou angiospermes) parmi les végétaux, ou des mammifères parmi les animaux. C'est pourquoi chez ces organismes le jeune embryon se développe sur le parent femelle jusqu'à un stade avancé du développement.

Chez les angiospermes comme chez les mammifères, il existe des tissus spécialisés dans le transfert des ressources de la mère vers l'embryon : le placenta chez les animaux, et l'**albumen** chez les angiospermes. Ces tissus partagent une caractéristique rare : l'empreinte parentale. L'empreinte parentale est l'expression différentielle des gènes selon qu'ils sont hérités paternellement ou maternellement (Surani 1991).

L'albumen des angiospermes présente également d'autres spécificités. La première tient au fait que l'albumen résulte d'une **double fécondation**. La double fécondation est la fécondation concomitante du gamète femelle (l'oosphère) et d'une cellule voisine : la cellule centrale. C'est le produit de la fécondation de la cellule centrale qui se développe en albumen. Une autre spécificité tient au fait que le nombre de **contributions génétiques** (c'est-à-dire de génomes haploïdes) maternelles qui participent à l'albumen est variable d'une espèce à l'autre.

Par ailleurs, il existe généralement des tissus nourriciers annexes dans la graine angiosperme, tels que le **périsperme** (tissu maternel), ou les **cotylédons** (tissu de l'embryon).

Plusieurs théories ont été proposées pour expliquer les spécificités que l'on a décrites. **La théorie du conflit sur l'allocation des ressources** est celle qui a suscité le plus d'investigations. La formulation intuitive de ce conflit est que la stratégie optimale de la femelle consiste à allouer ses ressources équitablement entre tous les embryons, tandis que l'optimum des embryons consiste à biaiser l'allocation de ces ressources en leur propre faveur. La double fécondation, les changements dans le nombre de contributions maternelles, et l'empreinte parentale sont vues, dans le cadre de ce conflit, comme les étapes d'une course aux armements pour le contrôle de l'allocation des ressources.

La théorie alternative à laquelle nous nous sommes le plus intéressés est **la théorie de l'hétérozygotie**. Cette théorie suggère que la double fécondation et les variations dans le nombre de contributions maternelles augmentent l'hétérozygotie, ce qui permet notamment de masquer les mutations délétères récessives. D'autres théories alternatives

ont été proposées, particulièrement pour l'empreinte parentale. Elles sont nombreuses et sont décrites dans un chapitre spécifique (chapitre 5).

L'objectif de cette thèse est d'identifier et de discuter les pressions de sélection auxquelles sont soumis les tissus nourriciers de la graine des angiospermes. Nous avons ainsi rassemblé, organisé, et confronté des informations relatives à des domaines très variés, tels que la modélisation, la phylogénie, la biologie moléculaire, l'embryologie et la physiologie. Dans cette introduction, nous commençons par décrire l'ensemble des tissus nourriciers existants, leur mode de formation, et leurs particularités génétiques. Nous présentons ensuite les scénarios possibles de l'évolution des tissus nourriciers, en nous appuyant sur leurs phylogénies. Nous terminons enfin par l'exposition détaillée des deux principales théories qui ont été proposées pour expliquer l'évolution des tissus nourriciers chez les plantes à graines, et de leurs prédictions.

1.1 Les tissus nourriciers de la graine

La graine est une structure particulière, car elle est formée par l'emboîtement de plusieurs générations. L'apparition de la graine chez les végétaux est intimement liée à l'évolution des cycles de vie. Nous commençons donc par présenter, en prenant quelques exemples parmi les grands groupes de végétaux, les grandes étapes de l'évolution du cycle de vie des végétaux vasculaires. Nous détaillons ensuite les différents types de tissus nourriciers de la graine angiosperme, et leurs particularités génétiques.

1.1.1 Cycles de vies et modes de transmission des ressources

La plupart des végétaux vasculaires ont un cycle de vie où alternent deux stades multicellulaires : un stade "gamétophytique", et un stade "sporophytique". Le **gamétophyte** se développe à partir d'une spore issue de la méiose d'une cellule de sporophyte, et produit des gamètes mâles et/ou femelles. Le **sporophyte** se développe à partir de la fusion de deux gamètes, ou fécondation (Fig. 1.1). Le sporophyte est généralement diploïde (2n), le gamétophyte est donc généralement haploïde (n). La durée et la forme de ces stades varient grandement entre les différents groupes de végétaux.

Chez les plantes sans graine, comme les fougères, sporophyte et gamétophyte sont tous deux des stades libres [1], bien que le sporophyte se nourrisse sur le gamétophyte au début de son développement. Le gamétophyte est hermaphrodite : il produit à la fois des gamètes mâles et femelles.

Chez les plantes à graine, les gamétophytes sont sexués : certains gamétophytes produisent des gamètes mâles, et d'autres produisent des gamètes femelles. Le gamétophyte mâle est libre, généralement disséminé par le vent ou les insectes : il s'agit du pollen. Le gamétophyte femelle n'est pas libre. Il est contenu dans l'ovule, qui est un organe sporophytique spécialisé dans la reproduction. Par conséquent, le gamétophyte des plantes à graine est très réduit.

Chez les plantes à graine mais sans fleurs, les **gymnospermes**, le gamétophyte femelle est réduit à un tissu de quelques milliers de cellules (Fig. 1.2). Plusieurs archégones, contenant chacune une oosphère, se forment sur le gamétophyte. En général, plusieurs oosphères sont fécondées, mais un seul embryon se développe jusqu'à maturité. Les jeunes

1. C'est-à-dire qu'on peut les trouver indépendamment l'un de l'autre.

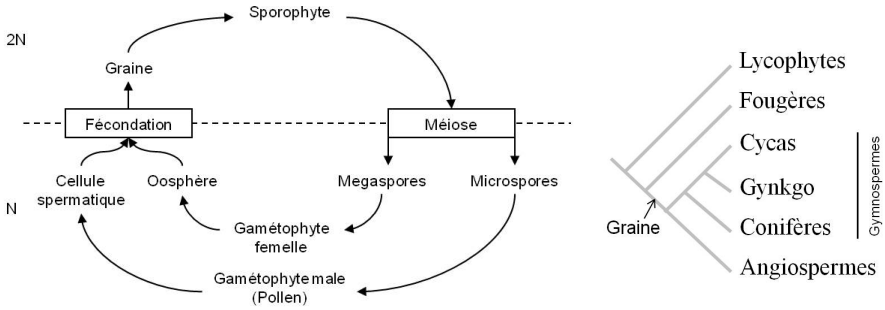

Figure 1.1 – A. Cycles de vie des plantes à graine. Le cycle de vie est formé par l'alternance de sporophytes et de gamétophytes. Les gamétophytes sont généralement différenciés en méga-gamétophyte ou gamétophyte femelle, et en micro-gamétophyte ou gamétophyte mâle : ils produisent respectivement des gamètes femelles ou mâles. B. Phylogénie des végétaux vasculaires. Cette phylogénie montre l'apparition de la graine (Réalisée d'après Donoghue & Doyle 2000).

sporophytes se nourrissent sur le gamétophyte, comme chez les fougères (sur le gamétophyte femelle). Le gamétophyte femelle développe un tissu spécialisé dans la nutrition, l'endosperme. Nous avons donc chez les gymnospermes l'emboîtement de trois générations qui sont i) un sporophyte mature, ii) un gamétophyte femelle, et iii) un embryon de sporophyte. Nous appellerons donc par la suite "sporophyte maternel" un sporophyte porteur de gamétophytes femelles. C'est de cet emboîtement que dérive **la graine**. En effet, la graine provient d'une transformation de l'ovule fécondé. Elle est composée de tissus du sporophyte maternel qui forment des téguments, du gamétophyte transformé en tissus de réserve, et du sporophyte de la génération suivante, c'est-à-dire l'embryon. Lorsque la graine arrive à maturité, elle est libérée et germe. Lors de la germination, le gamétophyte femelle est totalement absorbé par l'embryon et les téguments se décomposent, ce qui met fin à cet emboîtement de générations.

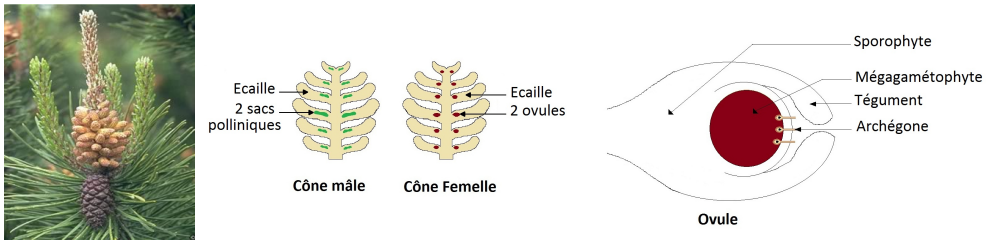

Figure 1.2 – Gamétophytes chez les conifères. Le sporophyte produit des cônes, formés d'écailles (photo). A la base de chaque écaille sont formés des organes reproducteurs. Ces organes sont des ovules pour les cônes femelles, et des sacs polliniques pour les cônes mâles (au centre). C'est à l'intérieur de ces organes que sont formés les gamétophytes. Ainsi, le gamétophyte femelle se développe au sein de l'ovule, protégé par des téguments. Des archégones, qui abritent des oosphères, se développent sur le gamétophyte (à droite). Côté mâle, chaque sac pollinique forme et relâche une importante quantité de pollen.

Chez les plantes à graine et fleurs, **les angiospermes**, les organes reproducteurs sont

tous contenus dans les fleurs, qui sont en général hermaphrodites (Fig. 1.3). Dans la fleur, les étamines portent des anthères, qui produisent le pollen (i.e. le gamétophyte mâle). Un ou plusieurs ovules sont présents. Comme chez les gymnospermes, l'ovule contient le gamétophyte femelle (également appelé "sac embryonnaire"). Cependant, le gamétophyte femelle est encore plus réduit que celui des gymnospermes, il n'est composé que de quatre à treize cellules (selon les espèces). L'une de ces cellules est l'oosphère (le gamète femelle). Des cellules sont accolées à l'oosphère, il s'agit des **synergides**. Lors de la fécondation, elles reçoivent les noyaux des cellules spermatiques et les transmettent aux gamètes. Une cellule au centre est généralement plus volumineuse que les autres cellules, et peut contenir de un à 14 noyaux haploïdes (les noyaux polaires). Il s'agit de la **cellule centrale**. Le développement de cette cellule après la fécondation produira le tissu nourricier de la graine, **l'albumen**. Les **antipodes** sont l'ensemble des autres cellules, leur fonction est mal connue. Lorsque le pollen germe sur une extension de l'ovule, le stigmate, un tube pollinique se forme et amène deux cellules spermatiques jusqu'au gamétophyte femelle ; ainsi se produit la fécondation.

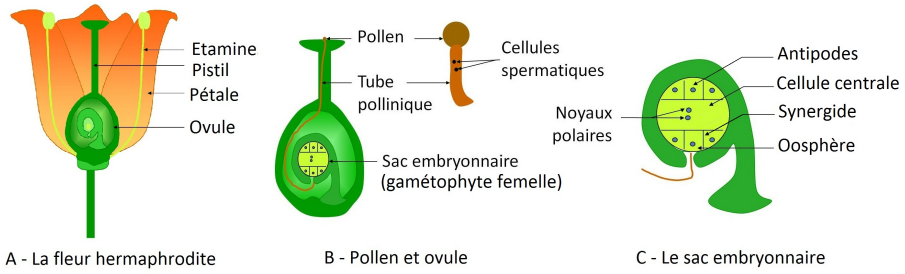

Figure 1.3 – La fleur et les gamétophytes angiospermes. A gauche : la fleur. Au centre : l'ovule (en vert) et le pollen développant un tube pollinique. A droite : le tube pollinique se développe jusqu'au gamétophyte femelle (en vert clair).

1.1.2 L'albumen, un tissu nourricier original

1.1.2.1 Les gamétophytes angiospermes

Pour vraiment comprendre le mode de formation du tissu nourricier, l'albumen, chez les angiospermes, il faut partir de la méiose et la formation du gamétophyte (Fig. 1.4). Lors de la méiose, quatre spores, dont les contenus génétiques diffèrent, sont produites.

En général, le gamétophyte n'est formé qu'à partir d'une seule de ces quatre spores (Maheshwari 1948, 1963) ; on parle de gamétophyte **monosporique**. La spore subit deux ou plus généralement trois mitoses (selon l'espèce). Dans le premier cas, cela conduit à un gamétophyte à quatre noyaux répartis dans quatre cellules, l'oosphère étant entourée de deux synergides et la quatrième cellule étant la "cellule centrale". Il n'y a pas d'antipodes. Il s'agit du type Oenothera. Dans le second cas, cela conduit à un gamétophyte à huit noyaux répartis dans sept cellules. Deux de ces noyaux, appelés **noyaux polaires**, sont dans la cellule centrale. L'oosphère est entourée de deux synergides, tandis que trois antipodes sont alignées de l'autre côté de la cellule centrale. Il s'agit du type Polygonum, qui est présent chez 70% des espèces.

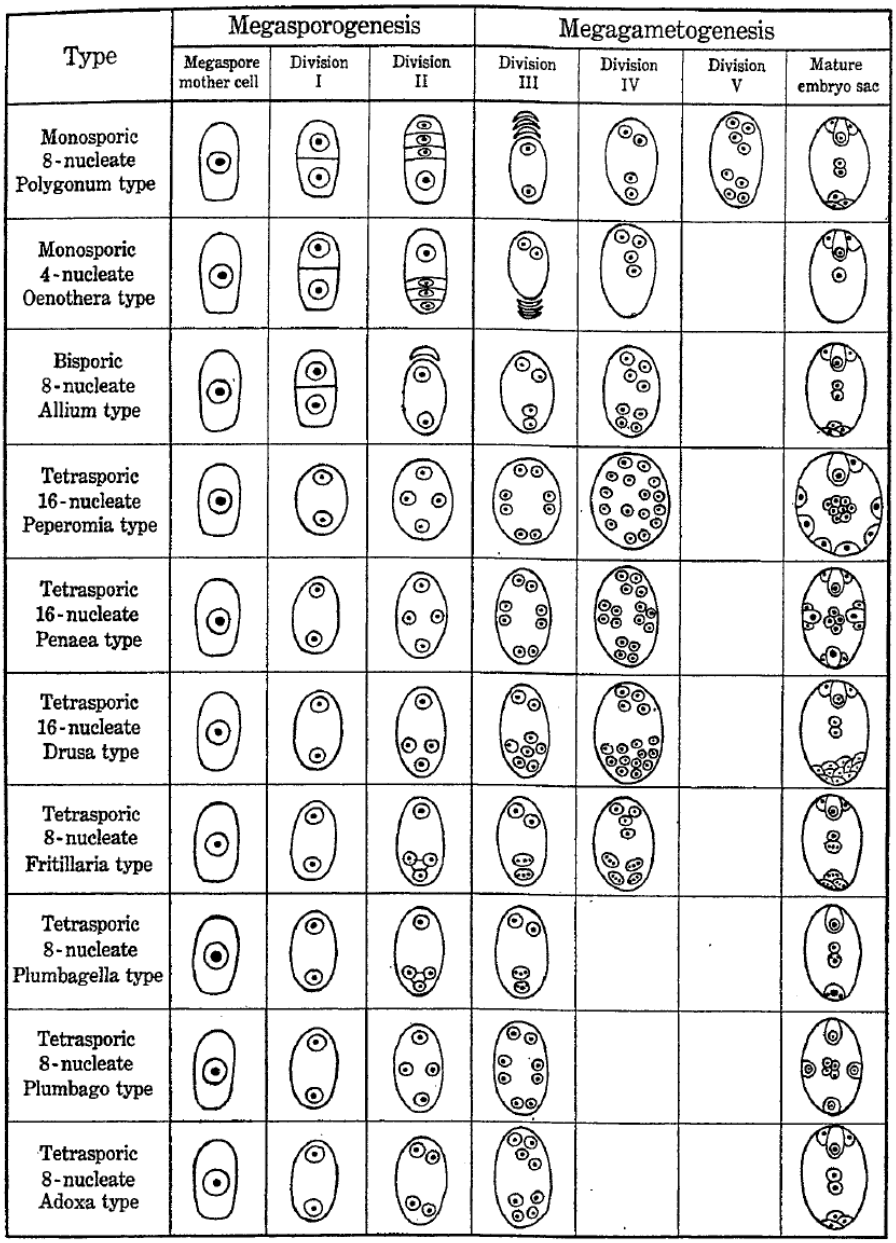

Figure 1.4 – Les différents types de gamétophytes femelles chez les angiospermes d'après Maheshwari (1948). Les colonnes "Megasporogenesis" correspondent aux étapes de la méiose, les colonnes "Megagametogenesis" correspondent aux étapes de la mitose, i.e. de formation du gamétophyte. La dernière colonne correspond au gamétophyte femelle mature. Chaque point représente un génome haploïde. Chaque cercle autour d'un ou plusieurs points représente un noyau. Les cercles de niveau supérieur représentent les contours des cellules.

Parfois, le gamétophyte est formé à partir des deux spores qui se sont séparées lors de la seconde division de méiose (phase équationnelle). On parle alors de gamétophyte **bisporique**. Chacune des spores subit deux mitoses. Les mitoses de la première spore fournissent les noyaux de deux synergides, de l'oosphère, et un noyau polaire. Les mitoses de la seconde spore fournissent les noyaux des trois antipodes, et un autre noyau polaire. Les noyaux polaires sont contenus par la cellule centrale. Il s'agit du type Allium, qui est morphologiquement semblable au type Polygonum.

Enfin, le gamétophyte peut être formé à partir de l'ensemble ou de plusieurs des quatre spores produites par la méiose, qui peuvent subir chacune une ou plusieurs mitoses. On parle alors de gamétophyte **tétrasporique**.

Le nombre de mitoses et/ou de spores initialement présentes, le nombre de cellules du gamétophyte et le nombre de noyaux de la cellule centrale ne sont pas indépendants. D'après Friedman *et al.* (2008), les gamétophytes femelles des angiospermes sont en fait des entités modulaires. Ils sont composés de sous-unités de quatre noyaux (le produit de deux mitoses), et chaque sous-unité fournit un noyau polaire à la cellule centrale. Cela est clairement visible quand on compare les types Oenothera (un seul module), Polygonum (deux modules) et Penaea (quatre modules, voir Fig. 1.4). Un seul des modules fournit les synergides et l'oosphère, tous les noyaux surnuméraires des autres modules forment des antipodes.

1.1.2.2 La double fécondation

Un caractère spécifique des angiospermes en fait un groupe à part, et ajoute un niveau de complexité à ce schéma : la **double fécondation** (Raghavan 2003). Lors de la fécondation (fig. 1.3 C), l'une des cellules spermatiques féconde l'oosphère. L'originalité des angiospermes est qu'en même temps, une seconde cellule spermatique féconde la cellule centrale (Guignard 1899; Nawaschin 1898). Les deux cellules spermatiques étant issues du même gamétophyte mâle [2], elles dérivent d'une même mitose et sont donc strictement identiques d'un point de vue génétique. Le produit de la fécondation de la cellule centrale est une cellule triploïde si le gamétophyte est de type Polygonum. Le tissu triploïde qui se développe à partir de cette cellule constitue le tissu nourricier de la graine, l'albumen.

Comme les cellules spermatiques sont génétiquement identiques, albumen et embryon sont porteurs du même matériel génétique lorsque le sac embryonnaire est monosporique, bien que la contribution génétique maternelle soit doublée pour l'albumen issu d'un sac embryonnaire Polygonum. Dans les cas de polysporie, le gamétophyte n'est pas homogène, mais formé de plusieurs produits de méiose (Fig. 1.4). L'albumen peut alors porter du matériel génétique qui n'est pas porté par l'embryon.

Très peu de familles d'angiospermes peuvent développer une graine sans fécondation de l'albumen. C'est parfois le cas lorsque la reproduction n'est plus sexuée (**apomixie**), et c'est également le cas chez la famille des Podostemaceae (Raghavan 2003, 2006). Inversement, certains taxons de plantes à graine non angiospermes présenteraient une forme de double fécondation (voir encadré 1.1).

2. En général. Chez le maïs, il y a entre 5 et 10% d'hétéro-fécondation (D. Grimanelli, comm. pers.)

Encadré 1.1 – La double fécondation chez les gymnospermes

Une forme de double fécondation à été mise en évidence chez quelques taxons au sein des gymnospermes. La première description explicite de ce phénomène a été fournie pour *Ephedra nevadensis* et *E. trifurca* (Gnetales, Friedman 1990b,a, 1991, 1992). Chez *E. trifurca*, comme chez la plupart de gymnospermes, l'oosphère est incrustée dans l'archégone, qui consiste en général en une grande cellule centrale entourée de petites cellules périphériques. Le noyau de la cellule centrale se divise à maturité en deux noyaux (un noyau "oeuf" et un noyau "canal ventral", Friedman 1991; Raghavan 2003). Au moment de la fécondation, chacun de ces noyaux fusionne, chez *E. nevadensis* et *E. trifurca*, invariablement, avec l'une des deux cellules spermatiques. Chez *E. trifurca*, le second événement de fécondation conduit à la formation d'un zygote qui produit de multiples pré-embryons.
Chez *Gnetum gnemon* (Carmichael & Friedman 1996, 1995), il n'y a pas d'archégone autour de l'oosphère. Le gamétophyte femelle est une grande cellule vacuolée contenant de nombreux noyaux libres. Les deux noyaux des cellules spermatiques fusionnent chacun avec l'un des noyaux libres et indifférenciés. Chacune de ces fusions conduit à la formation de pré-embryons équivalents. Par ailleurs, chez la plupart des gymnospermes, plusieurs archégones sont généralement fécondés pour chaque sac embryonnaire, conduisant à un phénomène de polyembryonie. Mais les différents embryons sont souvent demi-frères, car des grains de pollen différents servent à la fécondation de chaque archégone. Ces formes de fécondation parallèles ne conduisent pas la formation d'un tissu nourricier et les embryons surnuméraires ont tendance à avorter précocement (Raghavan 2003). Ces formes de fécondation ne peuvent donc pas être considérées comme étant fonctionnellement équivalentes à la double fécondation observée chez les angiospermes.

1.1.3 La graine exalbuminée

Si l'albumen est le tissu de stockage des ressources et persiste à maturité chez un grand nombre d'angiospermes, notamment les graminées, il n'est plus qu'un tissu de transfert des ressources vers l'embryon chez d'autres angiospermes. Chez ces espèces, les ressources sont stockées, à maturité, soit par un tissu embryonnaire, les **cotylédons**, soit par un tissu du sporophyte maternel, le **périsperme**.

Les cotylédons sont au nombre de un ou deux chez les angiospermes, et se développent en périphérie du rachis, l'élément central de l'embryon.

Le périsperme est quant à lui dérivé du nucelle, qui est un tissu de l'ovule, c'est-à-dire sporophytique. Le nucelle est formé de une, deux (graine tenuinucellée) ou plusieurs (graine crassinucellée) couches de cellules (Maheshwari 1963). Chez la plupart des espèces, le nucelle est consommé par l'albumen au début de son développement et ne se développe donc pas en périsperme.

1.1.4 L'empreinte parentale

Comme nous l'avons vu, l'albumen présente de nombreuses particularités génétiques (double fécondation, doublement des contributions maternelles, polysporie). Ces particularités induisent des variations dans l'origine et le nombre de chromosomes qui forment ce tissu. Dans cette section, nous décrivons un autre type de particularité, liée à des mécanismes moléculaires qui contrôlent l'expression de ces chromosomes : l'empreinte parentale.

1.1.4.1 Définition

Plusieurs définitions de l'empreinte parentale coexistent dans la littérature. Pour certains auteurs, l'empreinte parentale est simplement l'expression différentielle des allèles d'un même gène selon leurs origines parentales (voir par exemple Kinoshita *et al.* 1999). Cependant, il existe plusieurs mécanismes pouvant être responsables de tels effets (tels que la transmission de transcrits cytoplasmiques). Pour d'autres auteurs, seules les différences qui correspondent à une expression *de novo* peuvent être qualifiées d'empreinte parentale (voir par exemple Grossniklaus *et al.* 1998).

Par ailleurs, les expressions "empreinte paternelle" et "empreinte maternelle" sont parfois utilisées pour décrire la méthylation paternelle ou maternelle, et d'autres fois utilisées pour décrire l'extinction allélique maternelle ou paternelle. Or, s'il est vrai qu'en général, les allèles méthylés ne sont pas exprimés (sont éteints), il existe des exceptions (Sleutels *et al.* 2000).

Enfin, il arrive également que les expressions "empreinte paternelle" et "empreinte maternelle" soient utilisées dans le sens implicite de "contrôle paternel de l'empreinte parentale" et "contrôle maternel de l'empreinte parentale".

Toutes ces utilisations conduisent à un certain nombre d'ambiguïtés. Nous ne parlerons donc pas, ici, d'empreinte paternelle ou maternelle. Nous parlerons d' "empreinte parentale" pour qualifier l'expression différentielle des allèles maternels et paternels lorsqu'elle est liée à des mécanismes épigénétiques. Pour nous, l'empreinte est donc nécessairement la somme d'une activation maternelle et d'une extinction paternelle ou l'inverse, et ne peut qualifier le statut, éteint ou actif, d'un seul allèle. Nous parlerons explicitement

Encadré 1.2 – Les empreintes génétiques dans le monde vivant

L'empreinte génétique (dont l'empreinte parentale est un sous-ensemble) consiste en l'expression différentielle des deux copies d'un même gène dans un individu diploïde. L'empreinte se produit généralement lorsque les ADNs des deux chromosomes sont différentiellement méthylés. En effet, la méthylation est une marque épigénétique qui intervient dans la régulation de l'expression génétique. La méthylation différentielle conduit à l'expression différentielle (ou asymétrique) des deux copies.
L'extinction allélique peut être aléatoire. L'exemple classique est l'inactivation du chromosome X dans les tissus somatiques des mammifères femelles (Lyon 1972), qui peut affecter la copie héritée par le père ou la mère, selon les cellules. Mais on peut également citer l'inactivation aléatoire de nombreux locus sur les autosomes (Pernis *et al.* 1965; Rajewsky 1996; Holländer *et al.* 1998; Bix & Locksley 1998; Chess *et al.* 1994; Rhoades *et al.* 2000; Gimelbrant *et al.* 2005, 2007). Chez l'humain, plus de 300 gènes pourraient être concernés (Gimelbrant *et al.* 2007).
L'extinction peut également ne pas être aléatoire, mais déterminée par l'origine paternelle ou maternelle de l'allèle. C'est par exemple le cas lors de l'**empreinte parentale**. L'empreinte fut définie pour la première fois chez le diptère *Sciara coprophila* où elle est impliquée dans le déterminisme du sexe, et décrite par ailleurs chez les mammifères, la drosophile, et le poisson zèbre. En voici quelques exemples :

1. Inactivation du génome paternel (PGE) chez l'insecte *Sciara coprophila* (Crouse 1960) ;
2. Inactivation paternelle du chromosome X dans les tissus extra-embryonnaires (placenta) des mammifères (Takagi & Sasaki 1975; West *et al.* 1977)
3. Inactivation paternelle du chromosome X dans les tissus somatiques des marsupiaux (Sharman 1971; Cooper *et al.* 1971)
4. Inactivations gènes-spécifiques autosomales chez les mammifères (McGrath & Solter 1984; Efstratiadis 1994) ;
5. Inactivations de transgènes chez le poisson-zèbre (McGowan & Martin 1997) ;
6. Inactivation paternelle de trois gènes d'un mini chromosome X chez la drosophile (Parisi *et al.* 2003).

Ces empreintes remplissent différentes fonctions, telles que le déterminisme du sexe (1) ou une compensation de dosage (hypothèse avancée pour 2, 3). Les inactivations autosomales (4) ont été identifiées pour une centaine de gènes appartenant à sept espèces de mammifères (http ://geneimprint.com/site/genes-by-species). On peut donc raisonnablement penser qu'elles sont répandues dans ce taxon. L'empreinte gène-spécifique autosomale des mammifères est celle qui a suscité le plus d'intérêt dans la littérature. Ces empreintes sont généralement présentes dans le placenta lors du développement embryonnaire. Cependant, certaines ont également été repérées dans des organes comme le cerveau ou le foie. L'explication dominante pour ces empreintes qui sont gènes-spécifiques est l'occurrence d'un conflit mâle-femelle sur l'allocation des ressources (Moore & Haig 1991). Les modalités d'un tel conflit sont détaillées section 1.3.2. Cette explication pourrait d'ailleurs potentiellement être appliquée à l'ensemble des cas 2 à 6 (Hurst 1997).

Gène	Nom complet	Espèce	Allèle + exprimé	Fonction	Fiabilité	Références
FIS2	FERTILIZATION INDEPENDENT SEED2	*A. thaliana*	Maternel	Domaine du groupe Polycomb (dont la fonction est le remodelage de la chromatine)	+	Luo *et al.* 2000; Jullien *et al.* 2006a
FWA (*HDG6*)	FLOWERING WAGENINGEN	*A. thaliana*	Maternel	Facteur de transcription	+	Kinoshita *et al.* 2003
MEA (*FIS1*)	MEDEA	*A. thaliana* , *A. lyrata*	Maternel	Domaine SET du groupe Polycomb	+	Grossniklaus *et al.* 1998; Vielle-Calzada *et al.* 1999; Wang *et al.* 2006; Spillane *et al.* 2007; Kinoshita *et al.* 1999
MPC (*At3g19350*)	MATERNALLY EXPRESSED PolyA Binding protein C-TERMINAL	*A. thaliana*	Maternel	Régulateur de traduction	+	Tiwari *et al.* 2008
PHE1	PHERES1	*A. thaliana*	Paternel	Facteur de transcription	+	Köhler *et al.* 2003, 2005
AtFH5 (FORMIN5)	ARABIDOPSIS FORMIN HOMOLOGUE 5	*A. thaliana*	Maternel	Régulation de l'actine	+	Ingouff *et al.* 2005; Fitzgerald *et al.* 2009
AGL80		*A. thaliana*	Maternel	Facteur de transcription	−	Portereiko *et al.* 2006
FIE	FERTILIZATION INDEPENDENT ENDOSPERM	*A. thaliana*	Maternel	Domaine WD du groupe Polycomb	−	Ohad *et al.* 1996, 1999
SKIP1		*A. thaliana*	Maternel	Régulateur du cycle cellulaire	−	Feil & Berger 2007
HDG9	HOMEODOMAIN GLABROUS9	*A. thaliana*	Maternel	Facteur de transcription	−	Gehring *et al.* 2009; Nakamura *et al.* 2006
HDG8	HOMEODOMAIN GLABROUS8	*A. thaliana*	Maternel	Facteur de transcription	−	Gehring *et al.* 2009; Nakamura *et al.* 2006
HDG3	HOMEODOMAIN GLABROUS3	*A. thaliana*	Paternel	Facteur de transcription	−	Gehring *et al.* 2009; Nakamura *et al.* 2006
ATMYBR2		*A. thaliana*	Maternel	Facteur de transcription	−	Gehring *et al.* 2009
AT5G62110		*A. thaliana*	Paternel	Facteur de transcription potentiel	−	Gehring *et al.* 2009
ZmFIE1 (*FIE1*)	*Zea mais* FIE1	*Zea mais*	Maternel	Domaine du groupe Polycomb	+	Danilevskaya 2003
ZmFIE2 (*FIE2*)	*Zea mais* FIE2	*Zea mais*	Maternel	Domaine du groupe Polycomb	−	Danilevskaya 2003
MEG1	MATERNALLY EXPRESSED GENE1	*Zea mais*	Maternel	Transfert des ressources	−	Gutierrez-Marcos *et al.* 2004
Nrp1	No-apical-meristem related protein1	*Zea mais*	Maternel	Facteur de transcription	−	Guo *et al.* 2003
Mez1	Maize enhancer of zeste1	*Zea mais*	Maternel	Domaine du groupe Polycomb	+	Haun *et al.* 2007
R [ea]	R	*Zea mais*	Maternel	Pigmentation de l'aleurone	+	Brink 1956; Kermicle 1970
Tubulin [ea]		*Zea mais*	Maternel	Protéine de stockage (tubuline)	−	Lund *et al.* 1995
dzr1 [ea]	10-kDa zein regulation1	*Zea mais*	-	Protéine de stockage (zéine)	−	Chaudhuri & Messing 1994
Mee1	Maternally expressed in embryo 1	*Zea mais*	Maternelle	NA	+/−	Jahnke & Scholten 2009
OsFIE1	*Oryza sativa* FIE1	*Oryza sativa*	Maternel (+4 jap)	Domaine du groupe Polycomb	−	Luo *et al.* 2009

Table 1.1 – Les gènes à empreinte parentale dans l'albumen. Les noms redondants (*FIE*, *FIS*) ne sont donnés qu'une fois pour ne pas surcharger le contenu. ea=empreinte allélique. NA= Information non disponible. Les gènes pour lesquels la fiabilité est + sont ceux pour lesquels on a démontré qu'aucune transcription n'avait lieu avant fécondation ou pour lesquels l'activité post-transcriptionnelle a été vérifiée. Les critères pour les autres gènes sont l'occurrence d'effets parent-dépendant associés à l'hypo-méthylation de l'ADN dans la cellule centrale.

d'extinction (maternelle ou paternelle), de méthylation (maternelle ou paternelle), ou de contrôle (maternel ou paternel) dans les cas appropriés.

Des cas d'expressions génétiques monoalléliques sont rencontrés de manière plus générale dans le monde vivant, ils sont décrits Encadré 1.2.

1.1.4.2 L'empreinte parentale chez les angiospermes

Chez les angiospermes, toutes les cibles des empreintes identifiées à ce jour avec certitude sont dans l'albumen. Pourtant, Jahnke & Scholten (2009) suggèrent que le gène *Mee1* du maïs pourrait avoir une empreinte dans l'embryon. L'inventaire de tous les gènes qui ont *ou pourraient* avoir une empreinte dans l'albumen, et leurs caractéristiques, est donné Table 1.1.

Peu de gènes ont, pour le moment, été proposés comme candidats comme cibles d'empreintes, et ceux qui le sont n'appartiennent qu'à quatre espèces : *Zea mais*, *Arabidopsis thaliana*, *A. lyrata* et *Oryza sativa*. L'empreinte n'a été confirmée (les mécanismes qui contrôlent les différences d'expression n'ont été décrits) que pour un petit nombre d'entre eux (fiabilité notée "+" dans la Table 1.1).

Quelques études permettent de soupçonner que d'autres gènes exprimés dans l'albumen seront bientôt ajoutés à la liste actuelle des gènes à empreinte. Les gènes à empreinte chez *A. thaliana* sont généralement moins méthylés dans l'albumen que dans l'embryon, et sont peu ou pas exprimés dans les autres tissus (Gehring *et al.* 2009). Gehring *et al.* (2009), ont identifié 50 gènes qui correspondent à ces critères et méritent d'être étudiés plus attentivement (voir également Henikoff *et al.* 2009). Guo *et al.* (2003) et Stupar *et al.* (2008) identifient une expression différentielle des allèles maternels et paternels pour 3 à 5% des gènes de maïs (ce qui correspondrait à plusieurs centaines de gènes, Gutierrez-Marcos 2009). Reste à démontrer si ces différences d'expression correspondent ou non à des empreintes parentales.

1.1.5 La génétique du tissu nourricier de la graine

Il est admis que les deux copies de chaque gène d'un individu diploïde sont exprimées simultanément au même niveau (Gimelbrant *et al.* 2007).

L'ensemble des particularités génétiques que nous avons décrites dans ce chapitre d'introduction modifient le ratio d'expression des gènes dans les tissus nourriciers. Par conséquent, l'expression des différentes contributions (maternelles et paternelles) est parfois asymétrique. Par ailleurs, chacun des tissus nourriciers que nous avons décris partage plus ou moins de gènes en commun avec l'embryon. Cela aura d'importantes conséquences à la fois pour la théorie du conflit sur l'allocation des ressources, et pour la théorie de l'hétérozygotie.

C'est pourquoi par la suite, nous utiliserons toujours la même nomenclature pour désigner les contributions génétiques exprimées dans le tissu nourricier. Les gènes d'origine maternelle et portés par l'embryon sont notés m. Les gènes issus du sporophyte maternel qui ne sont pas portés par l'embryon sont notés m'. La contribution génétique d'origine paternelle (héritée par le pollen) est notée p. Cette nomenclature permet d'exprimer simplement la composition génétique d'une entité donnée.

L'embryon est toujours diploïde, formé d'une contribution génétique maternelle m et d'une contribution génétique paternelle p. L'albumen peut être génétiquement iden-

tique à l'embryon ($1m : 1p$) si le gamétophyte femelle est de type Oenothera. Il peut être triploïde ($2m : 1p$) si le gamétophyte femelle est de type Polygonum. Il peut être formé de contributions maternelles différentes de celles portées par l'embryon, m', si le gamétophyte femelle est bi ou tétrasporique (Table 1.2). Si le ratio d'expression des contributions génétiques est affecté par l'empreinte parentale à un locus donné, ce locus à une expression $1m : 0p$, $0m : 1p$ ou intermédiaire.

Tissu	Composition génétique	Code synthétique
Tissu nourricier gymnosperme	$1m_{11} : 0p$	$1m : 0p$
Périsperme 2N	$1m_1 1m_2 : 0p$	$1m1m' : 0p^*$
Monosporique 2N (Oenothera, Nuphar, Cotylédons)	$1m_{11} : 1p$	$1m : 1p$
Monosporique 3N (Polygonum, Amborella)	$2m_{11} : 1p$	$2m : 1p$
Bisporique 3N (Allium, Endymion, Drusa)	$1m_{11}1m_{12} : 1p$	$2m : 1p$ pour les gènes non échangés pendant la recombinaison, $1m1m' : 1p$ pour les autres
Tétrasporique 3N (Adoxa, Drusa)	$1m_{11}1m_{21} : 1p$ ou $1m_{11}1m_{22} : 1p$	$1m1m' : 1p^*$ / $1m1m' : 1p$ ou $2m : 1p$ selon les gènes
Tétrasporique 5N (Penaea, Plumbago, Fritillaria, Plumbagella)	$1m_{11}1m_{12}1m_{21}1m_{22} : 1p$	$2m2m' : 1p^*$
Tétrasporique 9N (Peperomia)	$2m_{11}2m_{12}2m_{21}2m_{22} : 1p$	$4m4m' : 1p^*$
Tétrasporique 15N (Peperomia)	$2m_{11}4m_{12}4m_{21}4m_{22} : 1p$	$6m8m' : 1p^*$

Table 1.2 – Constitution génétique des différents tissus nourriciers. Soit m_1 et m_2 une paire de chromosomes du sporophyte maternel. Soit m_{11} et m_{12} les chromatides de m_1, et m_{21}, et m_{22} les chromatides de m_2. Au moment de l'appariement, m_{11} s'apparie avec m_{21}, et m_{12} avec m_{22}. Après recombinaison, les chromatides de chacune des paires sont identiques à la recombinaison près. A la fin de la méiose, quatre spores sont formées, chacune contient l'une des contributions m_{11}, m_{12}, m_{21}, ou m_{22} (Fig. 1.5). Soit $m = m_{11}$ la contribution qui est transmise à l'embryon. $m' = m_{21}$ est donc la contribution qui n'est pas transmise à l'embryon. m_{12} et m_{22} contiennent à la fois des gènes identiques par descendance (m) et des gènes différents (m') de ceux portés par l'embryon (respectivement à des taux r et $1 - r$ si r est le taux de recombinaison). L'ensemble des gènes portés par un même chromosome dans l'embryon (m) ne sont pas nécessairement portés par un seul et même chromosome dans le tissu nourricier (cas notés d'une astérisque).

La plupart des particularités citées ici conduisent à une asymétrie d'expression des gènes maternels et paternels, et génèrent des **effets parentaux** : deux hétérozygotes réciproques (c'est à dire deux individus qui ont hérité deux allèles donnés identiques, mais l'un à reçu une copie, par exemple a par la mère, et l'autre, par exemple A, par le père, quand l'autre à reçu a par le père et A par la mère) n'ont pas le même phénotype.

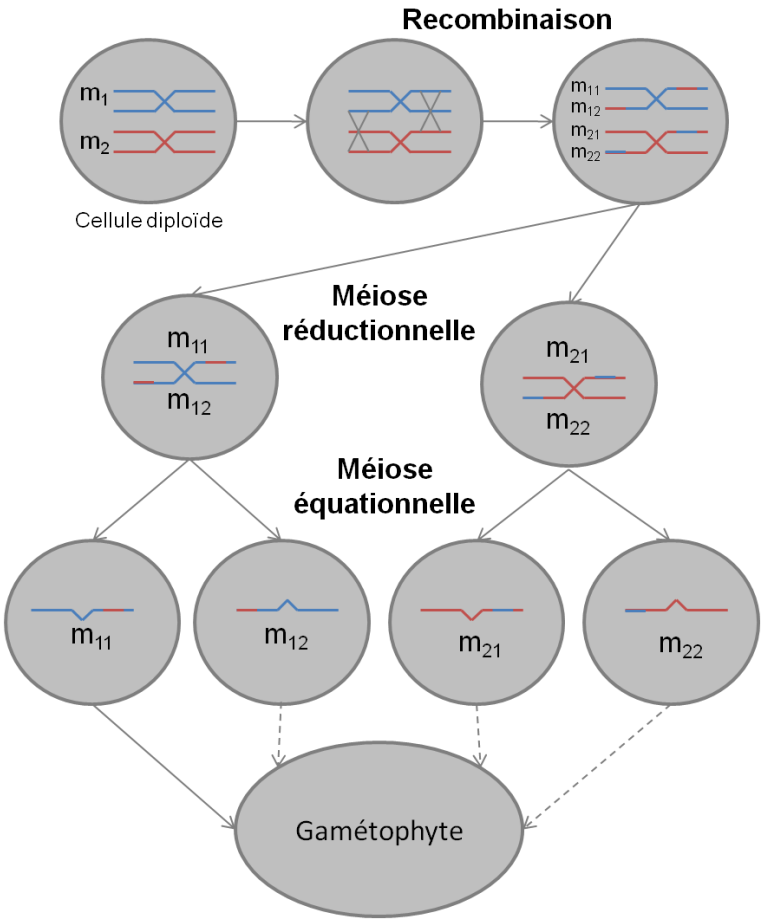

Figure 1.5 – La méiose. La méiose est la division qui permet le passage d'une phase diploïde à une phase haploïde. Elle consiste en la formation de quatre produits de méiose contenant chacun n chromatides, à partir d'une cellule contenant $2n$ chromosomes. La méiose est précédée de la duplication du matériel génétique (réplication de l'ADN), après laquelle chaque chromosome est formé de deux chromatides identiques. Les chromosomes de la cellule diploïde s'apparient, et donnent lieu à la recombinaison et conduisent à la formation de chromatides différentes pour un même chromosome. Puis deux divisions ont lieu successivement. La division réductionnelle conduit à la séparation des paires de chromosomes. La division équationnelle consiste ensuite en la séparation des deux chromatides de chaque chromosome. Chez les végétaux, les produits de méiose participent à la formation d'un ou plusieurs gamétophytes.

1.2 Scénarios de l'évolution des tissus nourriciers

1.2.1 Origine de la double fécondation et des contributions maternelles

1.2.1.1 L'albumen : homologue de l'embryon ou du gamétophyte ?

Deux hypothèses ont été historiquement avancées concernant les origines de l'albumen. Chacune suppose que l'état ancestral du tissu nourricier de la graine angiosperme est gamétophytique, comme chez les gymnospermes. La première hypothèse suggère que l'albumen est homologue [3] à un embryon surnuméraire ayant acquis une fonction de nutrition. Dans cette hypothèse, les changements dans les contributions maternelles ne se produisent qu'après l'acquisition de cette fonction de nutrition (Nawaschin 1898 dans Friedman 2001b, Sargant 1900). La seconde hypothèse suggère que l'albumen est homologue du gamétophyte, et n'a donc pas eu à acquérir de fonction nutritive (Strasburger 1900; Coulter 1911). Cette hypothèse ne présuppose pas de l'ordre d'apparition de la double fécondation et des changements dans les contributions maternelles. Elle suppose cependant qu'au moment où la double fécondation évolue, le moment de la fécondation est déjà précoce (c'est la cellule mère du tissu nourricier qui doit être fécondée).

L'étude du développement précoce de l'albumen chez les angiospermes basales révèle un certain nombre de similitudes avec le développement de l'embryon. En effet, le développement précoce de l'albumen comme de l'embryon implique une première division inégale en un domaine apical et un autre basal, c'est-à-dire une organisation bipolaire. Les domaines ont des développements nettement indépendants (Floyd & Friedman 2000; Friedman 2001a; Boisnard-Lorig *et al.* 2001). Cette indépendance serait également caractéristique du développement des embryons. Cependant, l'albumen est une structure déterminée, dont le développement est limité par le développement de la graine, tandis que l'embryon est une structure indéterminée et initie la phase d'embryogenèse qui va perdurer après que la graine a atteint sa maturité (Floyd & Friedman 2000). Des études de développement de l'albumen *in vitro*, libéré des contraintes de l'ovule, ont également mis en évidence des analogies avec la structure bipolaire de l'embryon (Kranz *et al.* 1998). Enfin, l'occurrence d'événements de double fécondation produisant des embryons surnuméraires chez les gymnospermes (Willson & Burley 1983) va plutôt en faveur l'hypothèse de l'homologie à l'embryon surnuméraire.

Friedman (2001b) décrit aussi un certain nombre d'analogies intéressantes entre les développements du gamétophyte femelle chez les gymnospermes et de l'albumen chez les angiospermes. Il suggère que l'une des phases de croissance qui précède la fécondation du gamétophyte chez les gymnospermes pourrait être homologue de la phase qui produit l'albumen, et qui a lieu après la fécondation chez les angiospermes (cf. Fig. 1.6). D'autre part Brown *et al.* (1999) décrivent, chez *Arabidopsis thaliana*, un patron (par alvéolation) qu'ils estiment analogue à celui observé lors du développement du gamétophyte femelle chez les gymnospermes. L'albumen d'*Arabidopsis thaliana* est susceptible de se développer sans fécondation, à partir de la fusion des deux noyaux maternels : c'est le cas chez certains mutants apomictiques. Ceci suggère plutôt une homologie de l'albumen à un ancien tissu nourricier issu du gamétophyte. Le fait qu'il soit possible de développer un albumen sans fécondation dans certains cas (*Arabidopsis*, Spillane *et al.* 2007) supporte plutôt l'hypothèse de l'homologie au gamétophyte.

Plus récemment, certains auteurs suggèrent que l'albumen ne serait homologue ni à un embryon, ni au gamétophyte, mais serait apparu *de novo* (voir par exemple Williams

3. Deux structures homologues sont des structures de deux espèces différentes, qui ont chacune évolué à partir d'une même structure présente chez l'ancêtre commun à ces deux espèces.

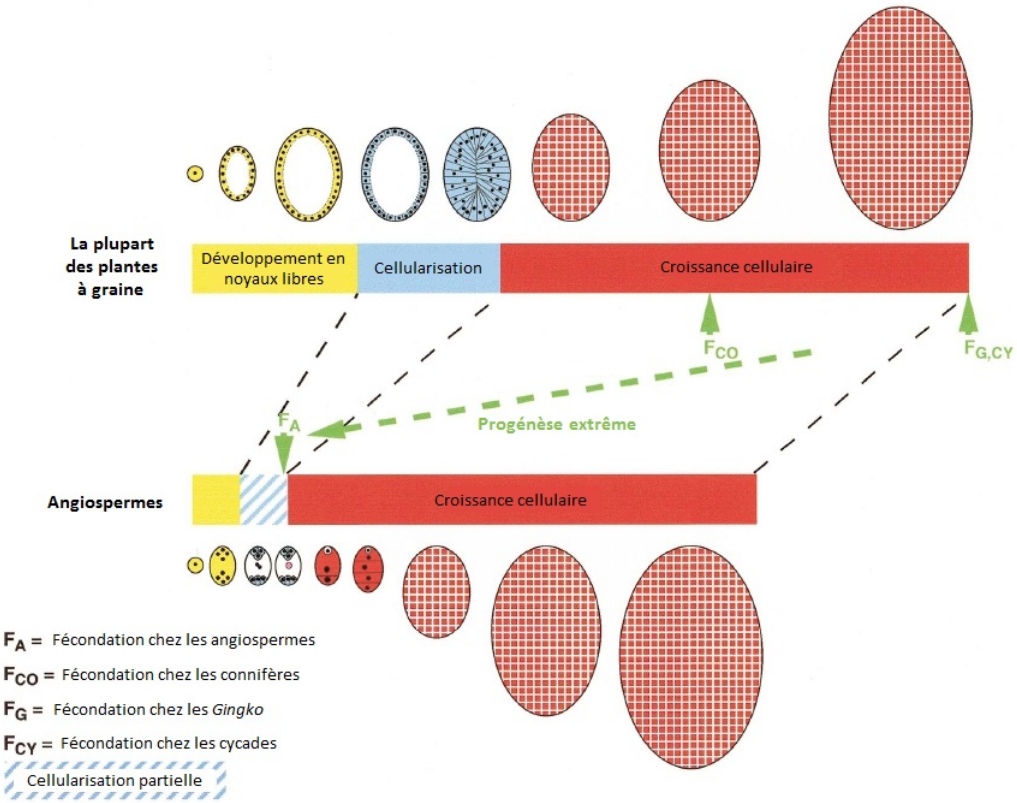

Figure 1.6 – Développements comparés du gamétophyte femelle chez les plantes à graine. Le développement du gamétophyte de la plupart des plantes à graine non-angiospermes (en haut) est caractérisé par trois phases : une phase de développement en noyaux libres, la cellularisation et une phase de croissance cellulaire. Chez les conifères, la fécondation a lieu au milieu de la phase de croissance cellulaire, alors que la plupart des ressources de la graine sont déjà accumulées. Chez les cycadales, la fécondation a lieu à la fin de la phase de croissance cellulaire, alors que l'ensemble des ressources de la graine sont déjà accumulées. Le développement du gamétophyte des angiospermes (ici Polygonum, en bas) est réduit à quelques divisions nucléaires (ici trois) suivi d'une cellularisation. La fécondation a lieu à ce stade, alors que très peu de ressources sont accumulées dans la graine. C'est ensuite que l'albumen accumule les ressources de la graine. Le développement de l'albumen suit les mêmes étapes que celui du gamétophyte (syncitium, cellularisation, phase de croissance cellulaire). Modifié d'après Friedman (2001b).

& Friedman 2004). D'après Friedman & Williams (2003), les premières angiospermes auraient été porteuses de gamétophytes femelles à quatre cellules, dépourvus d'antipodes et de second noyau polaire (type Oenothera), comme chez les Nymphéales et Austrobaileyales. Les gamétophytes du type Nymphéa (identiques à Oenothera) seraient issus d'une réduction importante du gamétophyte ancestral supposé analogue à celui des autres plantes à graine, passant par la perte des archégones et la réduction du nombre d'œufs à un seul. Une duplication du module de base à quatre cellules aurait ensuite généré l'expression d'un second module à l'extrémité chalazale chez l'ancêtre commun aux angiospermes non basales, et les antipodes constitueraient un duplicata stérilisé de l'œuf et ses synergides. Si hypothèse de modularité était juste, le sac embryonnaire ne serait en effet ni strictement homologue au gamétophyte des gymnospermes, ni homologue à un embryon surnuméraire.

1.2.1.2 Evolution de la génétique de l'albumen

Les deux hypothèses concernant les tissus homologues à l'albumen supposent que l'état ancestral du tissu nourricier est celui trouvé chez les gymnospermes et les plantes qui n'ont pas de graine : un gamétophyte $1m : 0p$. Il est également supposé, de manière implicite, que l'état polysporique est dérivé (ce qui parait raisonnable, car plus de 70% des espèces angiospermes sont monosporique et les espèces polysporiques sont dispersées dans la phylogénie suggèrent plusieurs évolutions indépendantes de cette caractéristique, voir Fig. 1.9). De ce fait, les deux hypothèses concernant les tissus homologues à l'albumen peuvent correspondre à deux scénarios de l'évolution de la génétique de l'albumen : $1m : 0p \rightarrow 1m : 1p \rightarrow 2m : 1p$ et $1m : 0p \rightarrow 2m : 0p \rightarrow 2m : 1p$ (Fig. 1.7).

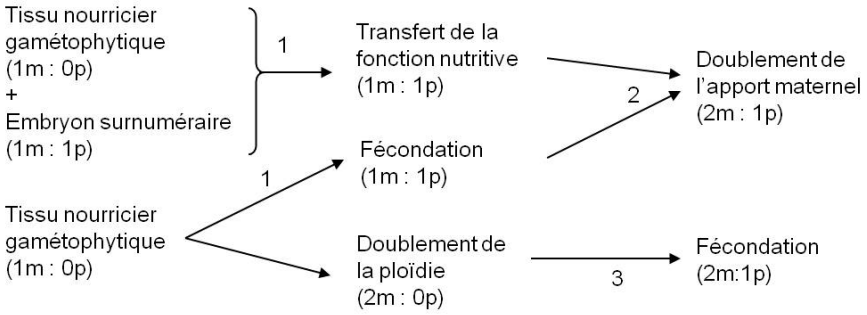

Figure 1.7 – Hypothèses concernant l'évolution de la génétique de l'albumen. 1 : transition $1m : 0p \rightarrow 1m : 1p$, 2 : transition $1m : 1p \rightarrow 2m : 1p$ et 3 : transition $2m : 0p \rightarrow 2m : 1p$. D'après la littérature, soit l'albumen est issu d'un embryon surnuméraire, tel que ceux trouvés chez les gymnospermes, qui a acquis une fonction nutritive et remplace progressivement l'albumen, soit l'albumen est issu de la fécondation du gamétophyte ancestral, haploïde ou déjà di- ou polyploïde. Dans les deux cas, le tissu nourricier ancestral est $1m : 0p$. Si l'albumen est l'homologue à un embryon surnuméraire, il est $1m : 1p$ dès le départ mais passe par l'acquisition de la fonction nutritive. Le doublement des contributions maternelles, donc l'état $2m : 1p$, vient ensuite. Si l'albumen est homologue au gamétophyte, deux scénarios génétiques sont possibles : ou bien la double fécondation précède le doublement des contributions maternelles, ou bien c'est l'inverse.

D'après Friedman & Ryerson (2009), il aussi parcimonieux de supposer que l'état

$1m : 1p$ a précédé l'état $2m : 1p$ ou de supposer l'inverse. D'après nous, en ajoutant l'information donnée par les groupes externes aux angiospermes, c'est-à-dire gymnospermes et fougères, l'étude de la phylogénie permet de supposer que l'évolution a suivi le scénario $1m : 0p \rightarrow 1m : 1p \rightarrow$ 2m :1p (voir Figure 1.8).

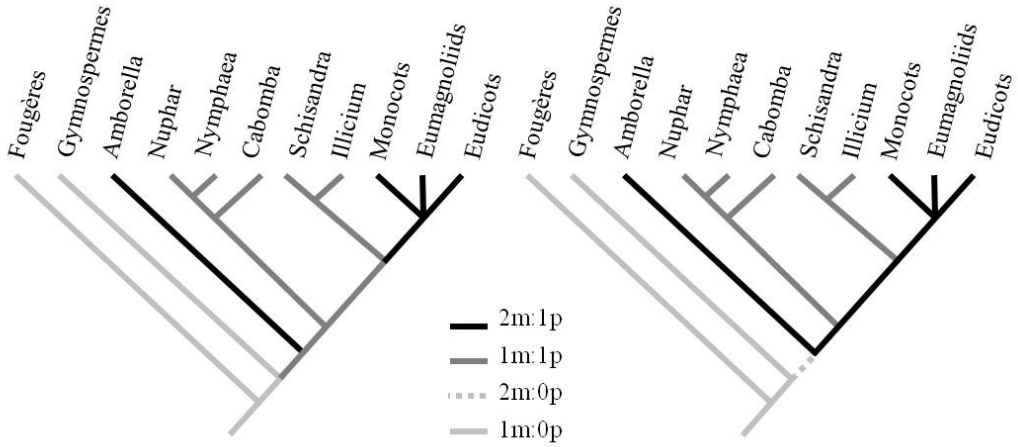

Figure 1.8 – Phylogénie de l'évolution des contributions parentales dans l'albumen. Modifié à partir de Friedman & Ryerson (2009) : nous avons ajouté les groupes externes aux angiospermes, c'est-à-dire gymnospermes et fougères. Dans le scénario de gauche, il y a une transition $1m : 0p \rightarrow 1m : 1p$ et deux transitions $1m : 1p \rightarrow 2m : 1p$, c'est-à-dire trois pas. Dans le scénario de droite, il y a une transition $1m : 0p \rightarrow 2m : 0p$, une transition $2m : 0p \rightarrow 2m : 1p$ et deux transitions $2m : 1p \rightarrow 1m : 1p$, c'est-à-dire quatre pas. De plus, le dernier scénario requiert un état $2m : 0p$ qui n'a été observé dans aucun taxon. Ainsi, le scénario de gauche est le plus parcimonieux. Résultat publié, voir Annexe A.

1.2.2 Une évolution récente de l'empreinte parentale

L'empreinte n'a été étudiée que chez quatre espèces, dont deux brassicacées (*A. thaliana*, *A. lyrata*) et deux graminées (*Z. mais*, et *O. sativa*). Chez de nombreuses espèces, les hybrides de plantes ayant des niveaux de ploïdie différents avortent précocement (Lin 1984; Cooper 1951). L'avortement n'a pas lieu si le ratio des contributions parentales est respecté dans l'albumen (par exemple si l'albumen et 4 : 2 au lieu de 2 : 1). C'est pourquoi ces avortements sont interprétés comme un effet de l'empreinte parentale (Vinkenoog *et al.* 2003). Si cette interprétation est juste, le fait que ce phénomène soit répandu chez de très nombreuses espèces suggère que l'empreinte est également assez largement répandue.

On connait la phylogénie d'un certain nombre des gènes cibles de l'empreinte, mais la phylogénie de l'empreinte parentale à proprement parler (c'est-à-dire les noeuds à partir desquels les empreintes de ces gènes cibles sont mises en place) est peu connue chez les plantes.

On peut donc se demander à quels moments ont évolué les empreintes que l'on connait.

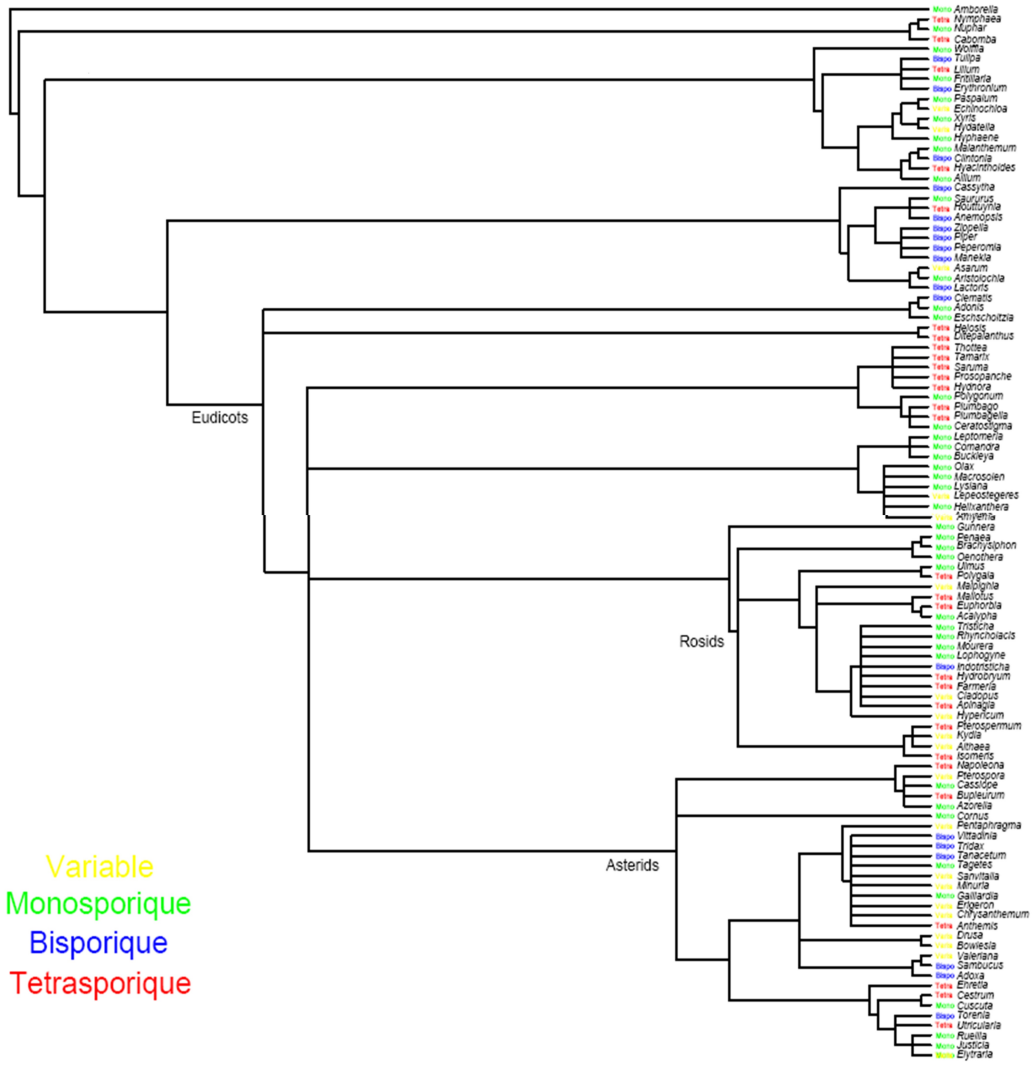

Figure 1.9 – Phylogénie de la polysporie réalisée à partir des données de Maheshwari (1963), et de Phylocom [a] (Webb *et al.* 2008). Dans une représentation exhaustive, 70% des espèces présenteraient un état monosporique.

a. Outil permettant d'extraire la phylogénie de taxons d'intérêt de la phylogénie de référence pour les angiospermes, l'APG.

Ci-après, nous décrivons plusieurs éléments qui indiquent une évolution récente des empreintes connues, remontant au mieux au niveau de la famille, à savoir les graminées pour les empreintes de *FIE2* et *Mez1*, et les brassicacées pour l'empreinte de *MEA*.

Pour rappel, la liste de l'ensemble des gènes à empreinte auxquels nous faisons référence est donnée Table 1.1.

1.2.2.1 Gènes *Fertilization Independent Endosperm* (*FIE*)

Les *FIE* sont homologues à *Embryonic ectoderm development*(*Eed*), qui est requis pour le maintien de l'extinction de certains allèles de gènes à empreinte chez la souris.

En ce qui concerne les *FIE*, nous disposons d'une information pour cinq gènes. Il s'agit de *FIE* chez *A. thaliana*, *FIE1* et *FIE2* chez *Z. mais*, et *FIE1* et *FIE2* chez *O. sativa*. *FIE1* est à empreinte chez *Z. mais* et probablement chez *O. sativa*. Le gène *FIE* d'*A. thaliana* et les gènes *FIE2* de *Z. mais* et *O. sativa* ne sont pas à empreinte. Cela suggère que cette empreinte a évolué chez un ancêtre commun de *Z. mais* et *O. sativa*, probablement une graminée.

1.2.2.2 Gènes *Enhancer of Zeste*

Les gènes *MEDEA* (*MEA*) et *Maize enhancer of zeste1*(*Mez1*) sont des *Enhancer of Zeste* (*E(z)*, Spillane *et al.* 2007). Des gènes de ce type sont connus chez la plupart des animaux (Haun *et al.* 2007). *MEA* est l'homologue au gène *Enhancer of Zeste* de la souris (Feil & Berger 2007).

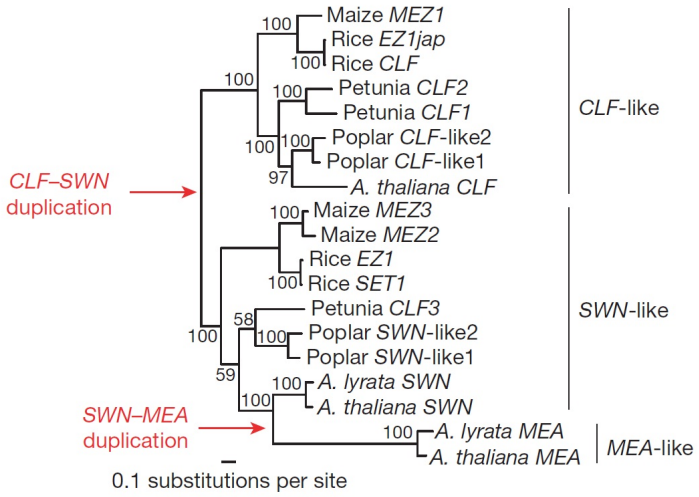

Figure 1.10 – Phylogénie des E(z)-like genes. On voit deux évènements de duplication : le premier évènement est à l'origine de CLF et SWN, et le second événement est la duplication de SWN, la nouvelle copie étant MEA. Tiré de Spillane *et al.* (2007).

Spillane *et al.* (2007) ont pu établir une phylogénie des gènes *E(z)-like* (fig. 1.10). *MEA* est le produit de la duplication de *SWINGER* (*SWN*) chez l'ancêtre commun d'*A. thaliana* et *A. lyrata*. *SWN* est lui même le produit d'une duplication de *Curly Leaf*

(*CLF*) chez un ancêtre beaucoup plus ancien, partagé par les eudicotylédones et les graminées.

Parmi les gènes de cette phylogénie, ni les gènes CLF-like (auxquels appartient *Mez1*), ni les gènes SWN-like n'ont de rôle dans le développement de la graine. En effet, les doubles mutants connus pour ces gènes produisent des graines normales (Spillane *et al.* 2007). *MEA* est donc différencié de *SWN* et *CLF* par sa fonction.

MEA est également différencié de *SWN* et *CLF* par son expression. *MEA*, *SWN* et *CLF* sont tous les trois exprimés dans le gamétophyte femelle, mais les expressions de *SWN* et *CLF* sont très réduites par rapport à l'expression de *MEA*. Cette différence se maintient pendant le développement précoce de la graine, car *MEA* est détectable pendant la phase de développement syncitiale de l'albumen contrairement à *SWN* et *CLF*.

Spillane *et al.* (2007) notent que *SWN* est sujet à une forte sélection purifiante, contrairement à *MEA*, ce qui suggère que c'est *SWN* qui a retenu la fonction ancestrale, qui n'etait donc probablement pas liée au développement de la graine.

Ces observations suggèrent une néofonctionnalisation de *MEA* et permettent de dater l'acquisition d'un rôle dans le développement de la graine par ce gène à l'époque de son apparition potentielle, c'est-à-dire il a entre 35 et 85 millions d'années. Sachant que *MEA* est à empreinte à la fois chez *A. thaliana* et *A. lyrata*, on peut également supposer que l'acquisition de l'empreinte de *MEA* a eu lieu dans cette période.

Étant données les absences d'empreinte de *SWN*, *CLF*, *Mez2* et *Mez3*, il est probable que l'empreinte de *Mez1* chez le maïs ait évolué de manière indépendante et convergente à celle de *MEA* chez *Arabidopsis* (Haun *et al.* 2007; Jullien & Berger 2009).

1.2.3 Evolution des graines à cotylédons et à périsperme

Les graines à périsperme sont rencontrées à divers endroits de la phylogénie (Figure 1.11) : chez *Hydatella* (une Nymphéacée, Friedman 2008), chez les Piperacées (*Piper*, le poivre), chez les Musacées (*Musa*, le bananier, Chin 1995), et chez certains genres des Amaranthacées (tels que *Beta*, la betterave ou *Spinacia*, l'épinard).

L'évolution vers le stockage des ressources dans les cotylédons semble être une tendance générale des angiospermes. Le ratio embryon : graine est une mesure de la taille de l'embryon relativement à la taille de la graine. Plus ce ratio est élevé, plus les ressources sont stockées dans les cotylédons. Forbis *et al.* (2002) ont établi une phylogénie de ce ratio et montrent qu'il existe une tendance générale allant dans le sens d'une augmentation du ratio embryon : graine, aussi bien au sein des angiospermes, qu'au sein des gymnospermes (Figures 1.11, 1.12 et 1.13). Cela soutient l'idée que la graine exalbuminée est dérivée, et correspond donc à une tendance vers l'évolution de graines dans lesquelles les ressources sont stockées dans l'embryon à maturité.

Les Eurosidées (pois, arachide, lupin, soja, haricot) sont le plus grand groupe présentant un ratio embryon : graine élevé. Ce ratio est en moyenne de 0.76 pour l'ensemble des Eurosidées, et de 0.94 chez les Fagacées (chênes, hêtres et châtaigniers). Beaucoup de familles d'Eurosidées ont un ratio de 1. En dehors des eudicotylédones, *Ceratophyllum*, certaines Alismatales (Hydrocharitacées et Zosteracées, qui sont des plantes aquatiques), les Burmanniacées, et les Orchidales (Les Orchidacées sont un cas particulier, car la nutrition du jeune embryon est assurée par les mycorhizes [4]) présentent des ratios élevés

4. Associations symbiotiques entre les plantes et les champignons.

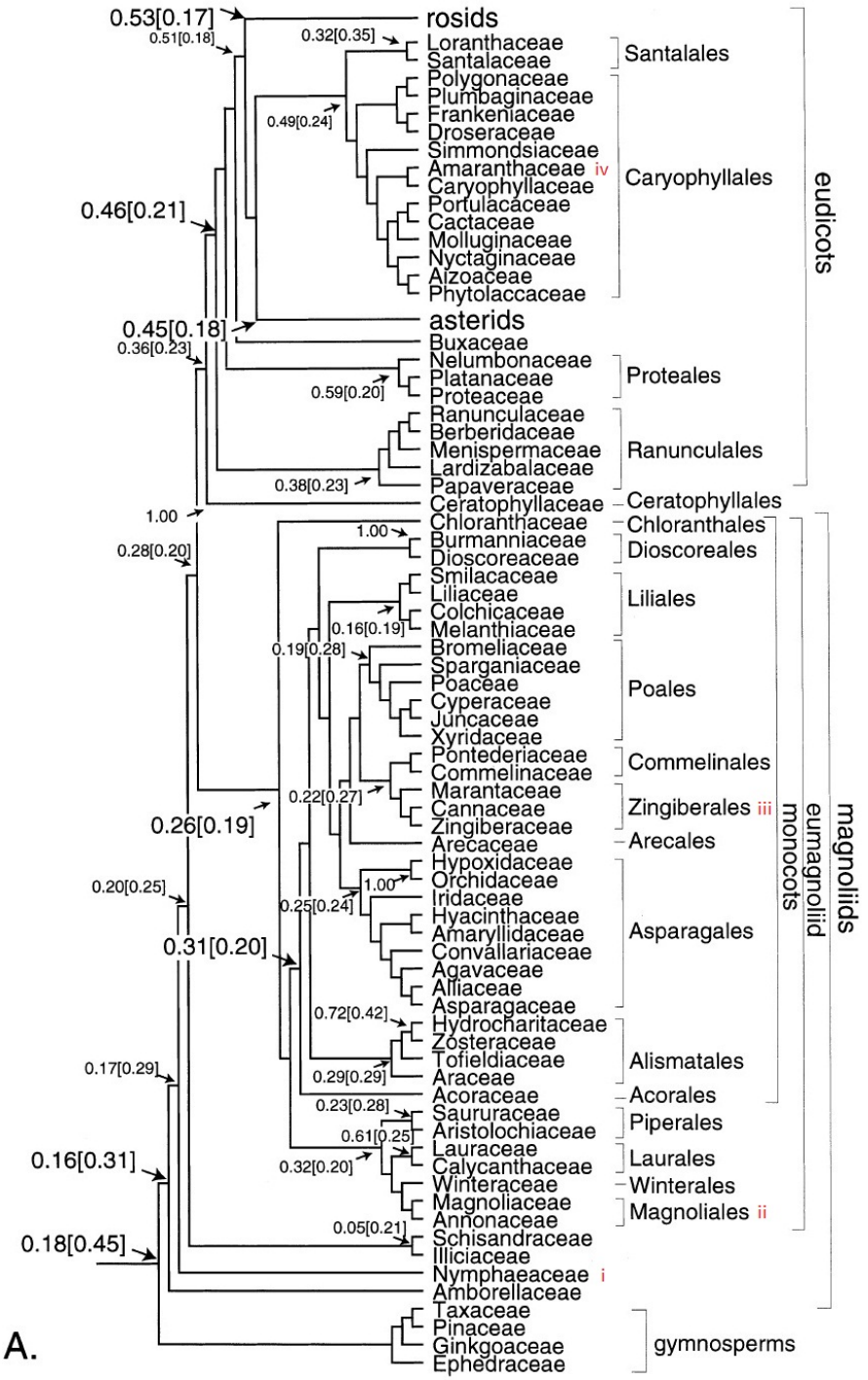

Figure 1.11 – Evolution du ratio Embryon : Graine chez les angiospermes. Le ratio embryon : graine est une mesure de la taille de l'embryon relativement à la taille de la graine. Les chiffres des noeuds représentent l'estimation des valeurs de ratio Embryon : Graine, les chiffres entre crochets représentent l'erreur standard. Les indices correspondent aux positions des groupes dans lesquels on trouve des graines à périsperme : i-*Hydatella* ii-Piperacées iii-Musacées iv-ex-Chénopodiacées. Tiré et modifié de Forbis *et al.* (2002).21

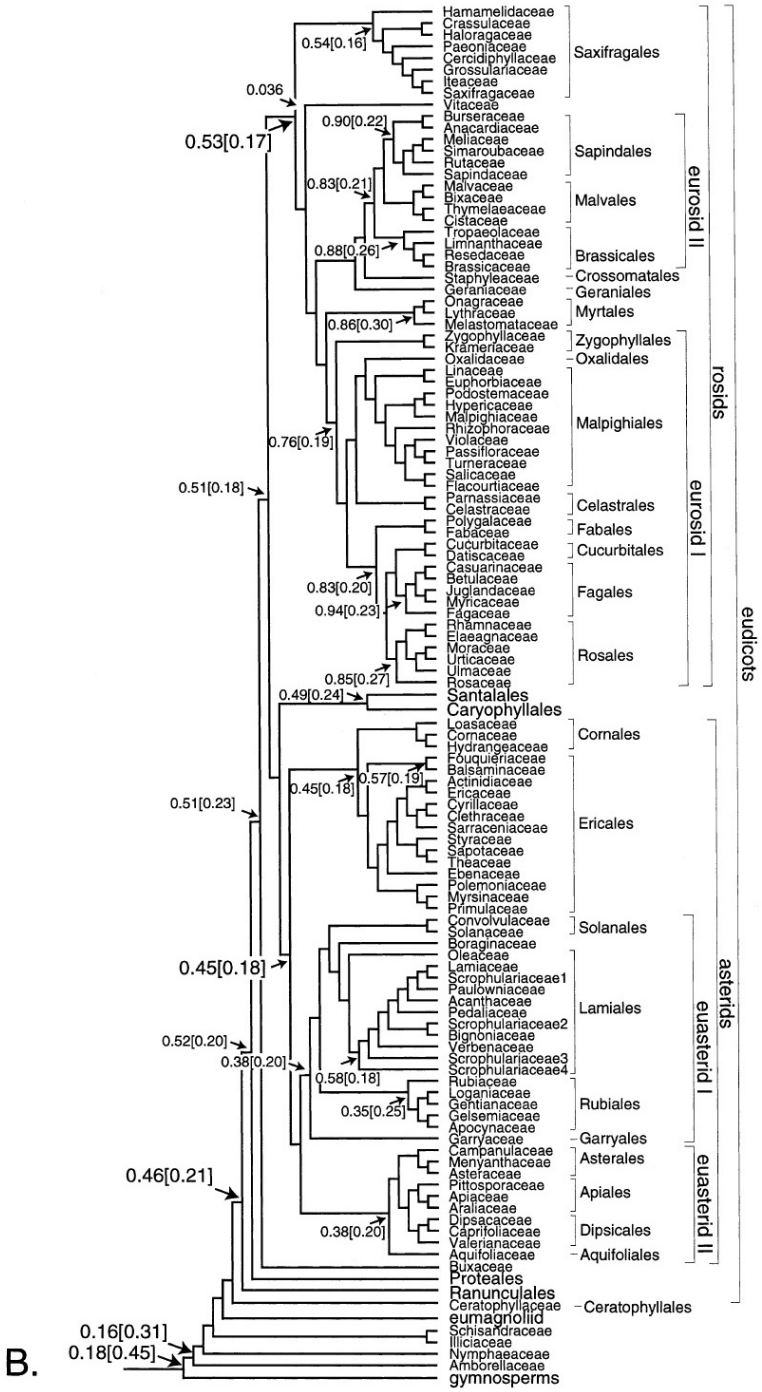

Figure 1.12 – Evolution du ratio Embryon : Graine détaillé pour les eudicotylédones. Suite de la figure 1.11.

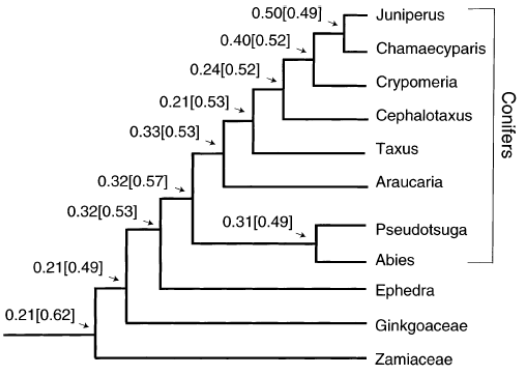

Figure 1.13 – Evolution du ratio Embryon : Graine chez les gymnospermes. Tiré de Forbis *et al.* (2002).

(0.72 en moyenne pour les Alismatales précitées, et 1 pour les autres). La plupart des espèces (dont les gymnospermes) présentent un ratio inférieur à 0.50. Les graminées (ou Poacées), dont les graines sont considérées comme les modèles de la graine albuminées, ont un ratio de 0.19. Dans certains groupes, on trouve une tendance à la diminution du ratio : Vitacées (qui sont des Rosidées), Santalales, et quelques familles parmi les Saxifragales.

1.2.4 Evolution des contributions génétiques des tissus nourriciers de la graine angiosperme

La synthèse de cette section permet de visualiser les étapes possibles de l'évolution de la graine angiosperme (Fig. 1.14). Il est admis que l'ancêtre commun sporophytique des angiospermes et des autres végétaux vasculaires se nourrissait sur un tissu gamétophytique $1m : 0p$. La double fécondation a évolué chez l'ancêtre commun des angiospermes, et le tissu nourricier de ce groupe est devenu $1m : 1p$. Certaines angiospermes branchées à la base de la phylogénie (Nymphéales et Illiciales) présentent encore cet état. L'état $2m : 1p$ a évolué de manière convergente chez *Amborella* et l'ancêtre commun à l'ensemble des autres angiospermes. Par la suite, une diversité de schémas s'est mise en place. Certaines angiospermes ont perdu la seconde contribution maternelle (les Oenothéracées), sont passées vers un gamétophyte polysporique, et/ou ont effectué une transition vers un stockage des ressources dans les cotylédons.

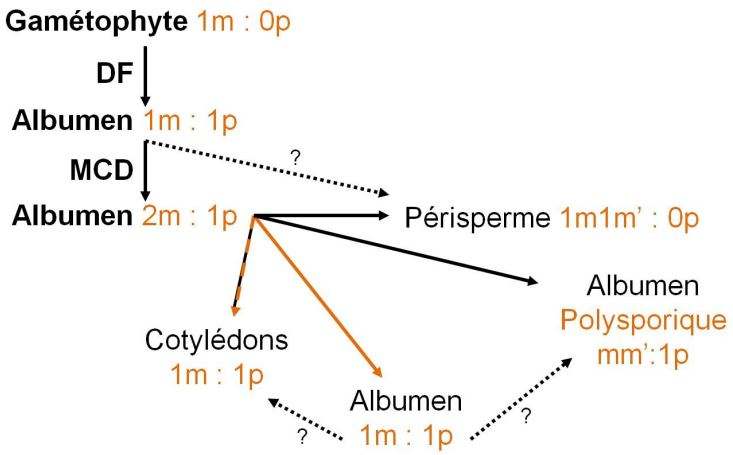

Figure 1.14 – Scénarios possibles de l'évolution des graines angiospermes. DF Double Fécondation, MCD Doublement des Contributions Maternelles.

1.3 Éléments théoriques

1.3.1 Le rôle clef du timing de l'allocation des ressources

La graine angiosperme présente de multiples innovations génétiques et formes de tissus nourriciers, alors que la graine gymnosperme présente quasiment toujours pour tissu nourricier un gamétophyte $1m : 0p$ [5]. Une transition importante doit pouvoir expliquer cette différence entre les gymnospermes et les angiospermes.

La double fécondation et l'empreinte parentale (i.e. l'expression différentielle des gènes d'origine maternelle et paternelle dans l'albumen) ne peuvent avoir lieu que si l'allocation des ressources à la graine a lieu après la fécondation, puisque ces innovations impliquent la contribution paternelle. Chez les angiospermes, contrairement aux gymnospermes, l'allocation des ressources a lieu après la fécondation. **La question à laquelle il faut répondre pour comprendre les différences entre angiospermes et gymnospermes est donc : quelle est la cause de la transition d'une allocation pré-fécondation des ressources à une allocation post-fécondation chez les angiospermes ?**

La première hypothèse est adaptative. Lorsque les ressources sont allouées précocement, une partie d'entre elles est investie inutilement dans des descendants exprimant des mutations délétères (Westoby & Rice 1982). Allouer les ressources tardivement pourrait permettre aux angiospermes de les allouer sélectivement aux embryons présentant les meilleurs génotypes (Lloyd 1980; Willson & Burley 1983; Westoby & Rice 1982; Korbecka *et al.* 2002). Les gymnospermes contournent ce problème par une stratégie alternative. Dans ce taxon, le gamétophyte est porteur de plusieurs archégones. Ainsi, les 3/4 des gamétophytes peuvent être fécondés par plusieurs mâles (Westoby & Rice 1982). En général, un seul des embryons produits se développe jusqu'à maturité, et cet embryon est le plus vigoureux (Westoby & Rice 1982; Sorensen 1982). Chez les gymnospermes, l'allocation sélective des ressources a donc lieu à travers la sélection des embryons après la fécondation (Willson & Burley 1983).

La seconde hypothèse suggère également que l'allocation tardive des ressources permet de mieux les gérer. Accumuler les ressources après la fécondation présente l'avantage de minimiser les pertes si certaines graines ne sont pas fécondées. La stratégie alternative des gymnospermes pourrait être de réabsorber les ressources allouées, mais cette possibilité ne semble pas avoir été explorée dans la littérature.

La dernière hypothèse suggère qu'une pression évolutive liée à la colonisation de nouveaux milieux par l'ancêtre des angiospermes a sélectionné une accélération du développement (ou progénèse, Takhtajan 1976), générant en cascade la réduction du gamétophyte et l'accumulation des ressources après la fécondation (Stebbins 1974, 1976). De nombreux éléments concernant l'anatomie du bois, les feuilles, et les structures embryologiques suggèrent effectivement une progénèse générale chez les angiospermes (Gould 1977).

Deux théories majeures proposent d'expliquer pourquoi l'allocation post-fécondation des ressources chez les angiospermes a entrainé l'évolution de nombreuses particularités génétiques (double fécondation, doublement des contributions maternelles, polysporie, empreinte parentale) dans l'albumen. Il s'agit de la théorie du conflit sur l'allocation des ressources et la théorie de l'hétérosis (ou théorie de l'hétérozygotie), que nous présentons ci-après.

5. Ou presque ; parfois plusieurs gamétophytes sont fusionnés dans la même graine, Willson & Burley 1983, p.111.

1.3.2 Théorie du conflit sur l'allocation des ressources

D'après la théorie du conflit sur l'allocation des ressources, l'allocation post-fécondation des ressources a mis en place les conditions nécessaires à l'expression d'un conflit génétique. Un conflit génétique se produit lorsque la direction de la sélection est différente pour un même trait i) pour des gènes présents à différents locus, ii) selon le tissu dans lequel les gènes qui contrôlent ce trait sont exprimés, ou iii) selon le stade du cycle de vie (embryon, male, femelle) dans lequel les gènes qui contrôlent ce trait sont exprimés (Burt & Trivers 2006).

Chez les plantes à graine, la probabilité pour un embryon d'atteindre un stade où il peut se développer de manière autonome (stade autotrophe) varie en fonction de la quantité de ressources accumulées dans la graine (voir e.g. Leishman *et al.* 2000; Turnbull *et al.* 1999; Moles & Westoby 2004; Moles *et al.* 2005 pour des comparaisons inter-espèce et Biere 1991, *Lychnis flos-cuculis*, Krannitz *et al.* 1991, *A thaliana* pour des comparaisons intra-espèce).

Il n'est possible d'allouer qu'une quantité limitée de ressources à chaque graine. Cette quantité dépend i) de la quantité de ressource allouées par le sporophyte à sa reproduction et ii) de la manière dont cette quantité est partagée entre les embryons. D'après la théorie du conflit sur l'allocation des ressources, il existe un conflit qui porte sur la répartition des ressources allouées aux différentes graines portées par un sporophyte.

Pour comprendre les modalités de ce conflit, nous commençons par discuter la relation entre quantité de ressource accumulée et probabilité pour l'embryon d'atteindre un stade de développement autonome. Nous en déduisons les optimums d'allocation pour les différents "acteurs" du conflit. Nous discutons ensuite des conditions dans lesquelles ces optimums pourraient converger, pour en déduire les conditions dans lesquelles un conflit est effectivement attendu. Nous terminons par les prédictions de cette théorie.

1.3.2.1 Quantité de ressources accumulées par la graine et probabilité pour l'embryon d'atteindre un stade autotrophe

Plusieurs facteurs sont susceptibles d'affecter la relation entre la quantité de ressources accumulées par la graine et la probabilité pour l'embryon d'atteindre un stade autotrophe.

Les ressources accumulées dans la graine participent directement au développement de l'embryon. La probabilité pour l'embryon de survivre jusqu'à germination (moment à partir duquel il commence à faire de la photosynthèse) devrait donc augmenter avec la quantité de ressources accumulées, et ce d'autant plus que les conditions sont stressantes. Plus précisément, on s'attend à ce que i) il y ait un seuil de ressources en deçà duquel la probabilité de survie est nulle, car l'embryon a besoin d'une quantité de ressource minimale pour atteindre un stade de développement autonome ; ii) la probabilité de survie augmente rapidement juste au dessus de ce seuil, car les ressources supplémentaires permettent de survivre à plus d'aléas ; et iii) la probabilité de survie augmente de moins en moins vite au fur et à mesure que la quantité de ressources déjà accumulées est grande, et fini par saturer.

1.3.2.2 Allocation des ressources par le sporophyte

Nous avons vu la manière dont la quantité de ressources accumulées par une graine affecte la probabilité que l'embryon se développe jusqu'à un stade de développement autonome. Cependant, le succès reproducteur d'un sporophyte ne dépend pas que de ce

paramètre. C'est un produit entre le taux de graines survivant jusqu'à maturité et le nombre de graines initialement produites. Or le nombre de graines initialement produites est égal au total des ressources disponibles divisé par la quantité de ressources allouées par graine. Il n'est donc pas possible de maximiser les deux composantes du produit "nombre de graine * quantité de ressources par graine" à la fois. Le modèle de Smith et Fretwell (1974) permet de prédire le mode de répartition des ressources qui maximise ce produit (Fig. 1.15).

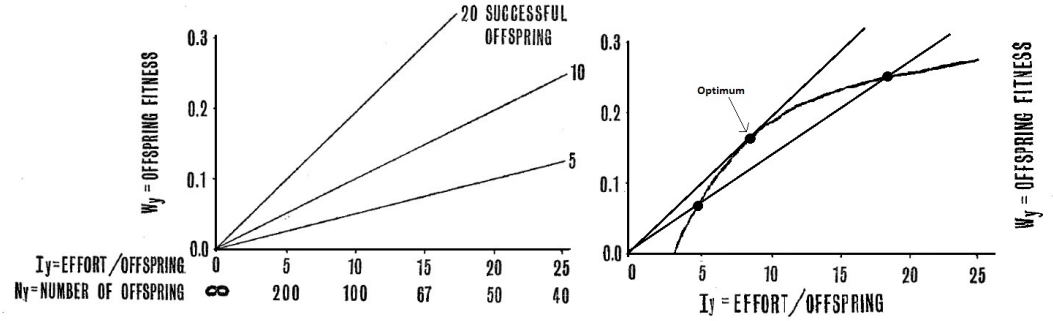

Figure 1.15 – Modèle de Smith et Fretwell. A gauche : si I_Y est la quantité de ressource accumulée par descendant, et N la quantité de ressource disponible, $N_Y = N/I_Y$ est le nombre de graines qu'une plante peut produire. Plus les graines accumulent de ressources, plus la probabilité que l'embryon se développe jusqu'à maturité est grande. Plusieurs hypothèses peuvent être faites sur la manière dont la survie des graines dépend de la quantité de ressources accumulées, et chacune fournit un nombre différent de descendants viables. A droite : la probabilité de survie des graines est supposée dépendre de la quantité de ressources accumulées par une fonction qui sature asymptotiquement, si bien que la probabilité que l'embryon se développe jusqu'à maturité est plafonnée. Seules les fonctions entre la quantité de ressources accumulées et la probabilité que l'embryon se développe jusqu'à maturité qui coupent la courbe saturante peuvent être réalisées par la plante. Le point qui maximise la relation entre quantité de ressources accumulées et probabilité de survie fournit la quantité de ressources par graine qui permet de maximiser le produit entre le nombre de graines initialement produites, et le taux de graines survivant jusqu'à maturité.

Dans le modèle de Smith et Fretwell (1974), les ressources sont allouées équitablement entre tous les embryons. Cependant, il a été démontré qu'un certain nombre de graines sont avortées par le sporophyte maternel, ce qui a plutôt été interprété comme une allocation sélective (non égalitaire) des ressources (voir par exemple Stephenson 1981; Lee & Bazzaz 1986; Ganeshaiah & Uma Shaanker 1992). On pourrait en effet supposer que la probabilité pour un embryon de se développer jusqu'à maturité dépend de la qualité génétique de l'embryon et/ou l'albumen. Dans ce cas, une allocation sélective des ressources aux embryons présentant les meilleures chances de se développer jusqu'à maturité augmenterait plus encore le succès reproducteur du sporophyte (Westoby & Rice 1982; Willson & Burley 1983; Haig 1990). Cette hypothèse selon laquelle le sporophyte maternel pourrait allouer sélectivement les ressources aux embryons ne fait pas l'unanimité. La difficulté est d'imaginer un mécanisme par lequel le sporophyte maternel pourrait avoir l'information de qualité génétique. Cette question ressemble beaucoup à celle posée pour

le choix de partenaire (signaux mensongers, théorie du handicap, etc., Willson & Burley 1983). Laisser les graines être en compétition pourrait être un moyen indirect (pour lequel disposer de l'information de qualité génétique n'est pas nécessaire) pour sélectionner certains embryons (Westoby & Rice 1982). Cependant, la capacité de compétition n'est pas nécessairement corrélée à la qualité génétique. S'il était montré que l'allocation sélective des ressources n'est pas possible, l'avortement de nombreuses graines devrait être interprété (dans le cadre du modèle de Smith et Fretwell) comme un moyen d'ajuster le nombre de graines produites en fonction des ressources disponibles.

1.3.2.3 Les sources du conflit

Dans la littérature, deux visions du conflit sur l'allocation des ressources coexistent. La première vision (associée au modèle de Smith et Fretwell) stipule que l'optimum du sporophyte est égalitaire, alors que l'optimum des embryons contenus dans chaque graine est de biaiser l'allocation des ressources en leur propre faveur. La seconde vision stipule que l'optimum du sporophyte n'est pas égalitaire, mais plutôt d'avorter sélectivement les embryons qui ont une mauvaise qualité génétique, alors que l'optimum des embryons est de ne pas être avortés. Ces visions convergent vers l'idée que l'optimum du sporophyte est d'allouer moins de ressources à chaque graine que ce qui correspondrait à l'optimum des embryons.

Cependant, la quantité de ressources qu'un embryon doit acquérir pour maximiser sa valeur sélective n'est pas illimitée. Même en supposant que la probabilité de survie de l'embryon augmente toujours avec la quantité de ressources acquises, il existe un plafond au delà duquel un biais dans l'allocation des ressources diminue la valeur sélective inclusive [6] de l'embryon.

En effet, les embryons sont apparentés les uns aux autres. Un gène présent dans l'un est donc présent dans un autre avec une probabilité r (cette probabilité mesure l'apparentement entre embryons). Les conditions dans lesquelles le plafond est atteint sont données par la règle de Hamilton (Figure 1.16, Hamilton 1964) : un individu (par exemple un embryon) aura un comportement égoïste plutôt qu'altruiste (ici, une allocation des ressources biaisée plutôt qu'égalitaire) tant que le bénéfice b du comportement altruiste pour l'ensemble des autres individus multiplié par leur apparentement r à l'individu focal est inférieur au coût c (ici un 'manque-à-gagner') pour cet individu focal ($br < c$).

L'apparentement des embryons entre eux diminue lorsque le nombre de pères fécondants augmente. C'est pourquoi certains auteurs défendent l'idée que le conflit femelle-embryon n'existe que parce les embryons sont fécondés par des pères différents. D'après ce point de vue, le conflit parent-descendant est en fait un conflit mâle-femelle (ou conflit sexuel, Queller 1983; Mazer 1987). Cependant, Haig (1986) souligne que les différentes mégaspores (i.e. les produits de méiose à partir desquels se développent les différents gamétophytes femelles) sont aussi en conflit. Il est évident que l'intensité de ce conflit là ne dépend pas du nombre de pères présents. Il semble donc que la vision individu-centrée des conflits manque de clarté. C'est pourquoi nous proposons la formulation gène-centrée qui suit.

6. La valeur sélective inclusive est une généralisation de la valeur sélective individuelle, qui prend en compte la valeur sélective de 'copies' de soi présentes dans d'autres individus.

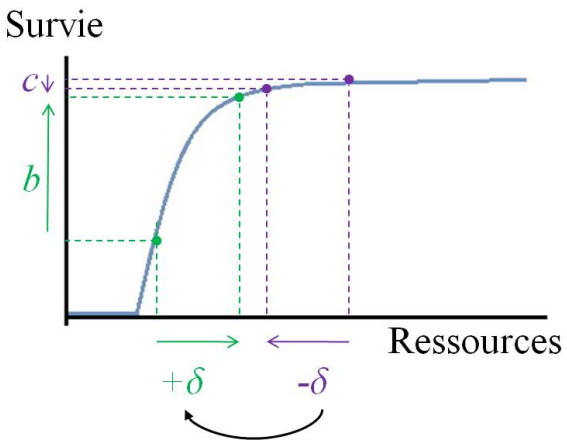

Figure 1.16 – Illustration de la règle de Hamilton. La quantité de ressources δ qui n'est pas acquise par un embryon donné représente un coût c pour cet embryon, en terme de survie. Le transfert de cette même ressource à un autre embryon représente un bénéfice b pour ce dernier. Si les deux embryons sont apparentés à un taux r, et que $br > c$, un comportement altruiste est prédit.

1.3.2.4 Une formulation génétique du conflit

Nous reprenons les notations présentées en section 1.1.5. Soit m la contribution héritée maternellement par un embryon focal, m' la contribution maternelle qui n'est pas héritée par cet embryon focal (mais l'est par 50% des embryons portés par le sporophyte maternel), et p_i les contributions paternelles héritées par les différents embryons, chaque père i apportant une contribution p différente.

Nous proposons que le conflit sur l'allocation des ressources est composé :

1. d'un conflit entre mégaspores, i.e. entre les m et m'. Ces contributions ne sont pas contenues par les mêmes embryons, et sont donc en potentiellement en conflit pour l'acquisition des ressources ;
2. d'un conflit entre pères, i.e. entre les différents p_i. Ces contributions ne sont pas contenues par les mêmes embryons, et sont donc en potentiellement en conflit pour l'acquisition des ressources ;
3. d'un conflit mâle-femelle, i.e. entre m et p. Ces contributions sont contenues par le même embryon, et sont donc l'objet d'un conflit intra-génomique. Il est nécessaire de bien comprendre les deux autres conflits, pour comprendre celui-ci, qui en résulte directement.

Le conflit entre les différents p_i dépend de la probabilité r_p de trouver une copie p_i focale dans les autres embryons (donc du taux de multi-paternité) mais aussi des coûts c pour la copie focale (de recevoir moins de ressources) et des bénéfices b (de recevoir plus de ressources) pour les autres copies. Le conflit a lieu lorsque $br_p < c$ (Hamilton 1964).

En cas de production d'embryons clonaux, r_p est égal à 1. Il n'est jamais avantageux pour p_i de biaiser l'allocation des ressources, puisqu'un tel biais diminue le nombre de copies de lui-même reproduites à travers les autres embryons. En cas de monoandrie [7]

7. Tous les embryons d'une femelle donnée sont fécondés par un seul mâle.

sporophytique (i.e. tous les embryons sont fécondés par des contributions paternelles provenant du même sporophyte), r_p est égal à 1/2, l'optimum de p_i dépend donc de b et c. Enfin, en cas d'allofécondation complètement polygame, r_p est nul, la valeur sélective de p_i est d'autant plus forte que les ressources sont préférentiellement allouées à l'embryon qui en est porteur.

Le conflit entre m et m' dépend des mêmes paramètres que le conflit entre les p_i, à la différence près que la probabilité de trouver une copie de m dans les autres embryons, r_m, ne dépend pas du taux de multi-paternité. Si la reproduction est sexuée, il est fixe, égal à 1/2.

Le conflit entre m et p_i résulte de la différence entre la valeur de r_p (qui est au plus 1/2 en cas de reproduction sexuée), et la valeur de r_m (qui est donc fixe, égale à 1/2). Ce conflit est nul en cas de monoandrie sporophytique, puisque alors $r_p = r_m$, par exemple dans le cas où la plante serait complètement autofécondante (les deux autres conflits subsistent tant que r_p et r_m ne sont pas égaux à 1). Comme ce conflit requiert la présence des contributions p_i pour s'exprimer, il n'a bien sûr d'expression qu'après fécondation.

Le conflit mâle-femelle peut disparaître alors que les deux autres persistent. Contrairement à l'idée suggérée par Queller (1983) et Mazer (1987), le conflit sur l'allocation des ressources n'est donc pas réductible au conflit mâle-femelle.

Ici, nous considérons donc une théorie du conflit sur l'allocation des ressources qui suggère qu'il existe un conflit entre les différentes contributions génétiques en jeu (m, m', p_i) portant sur la quantité de ressources allouées aux différentes graines portées par un sporophyte donné.

1.3.2.5 Prédictions

L'albumen joue une fonction cruciale dans l'acquisition des ressources par la graine. C'est pourquoi Charnov (1979) propose que l'évolution des contributions parentales dans ce tissu est due au conflit sur l'allocation des ressources (voir aussi Westoby & Rice 1982; Willson & Burley 1983; Queller 1983, 1984; Law & Cannings 1984; Bulmer 1986; Haig & Westoby 1989; Friedman 1995; Hardling & Nilsson 2001, 1999). Plus précisément, **la double fécondation, les changements dans les contributions maternelles et l'empreinte parentale sont supposés être des adaptations et contre-adaptations conduisant à des augmentations et diminutions du contrôle des contributions m, m' et p sur l'allocation des ressources, dans le contexte d'une course aux armements générée par les conflits génétiques** (Uma Shaanker & Ganeshaiah 1997).

Un certain nombre de modèles se penchent sur les prédictions qui peuvent être faites pour chacun des traits. On peut distinguer deux catégories.

La première catégorie de modèles s'attache à dériver l'optimum de la taille de graine du tissu nourricier, en fonction des contributions parentales qu'il contient. Ces modèles montrent que cet optimum converge d'autant plus vers l'optimum du sporophyte que l'albumen est apparenté au sporophyte et vers l'optimum de l'embryon que l'albumen est apparenté à l'embryon (Queller 1983, 1984; Willson & Burley 1983; Law & Cannings 1984; Hardling & Nilsson 1999; Bulmer 1986; De Jong *et al.* 2005; Zhang & Jiang 2000).

La seconde catégorie de modèles cherche un avantage direct à l'évolution des différents traits. Étrangement, aucun ne s'intéresse à l'évolution de la double fécondation.

Beaucoup de ces modèles s'intéressent aux conditions d'évolution de l'empreinte parentale (Mochizuki *et al.* 1996; Haig 1997; Spencer *et al.* 1998; Haig 1999; Hurst 1999; Spencer *et al.* 1998, 1999). Ces modèles prédisent que les gènes activateurs de croissance doivent être maternellement éteints, et paternellement activés, et inversement pour les inhibiteurs de croissance. C'est le cas d'*IGF2* et *IGF2R* chez la souris. *IGF2* est un activateur de croissance maternellement éteint, et *IGF2R* est un inhibiteur de *IGF2* qui est paternellement éteint (Willison 1991). La copie maternelle d'*IGF2R* régule donc l'expression de la copie paternelle d'*IGF2*.

Dans le cadre du conflit entre m et m', l'optimum de m est de biaiser l'allocation des ressources en faveur des embryons qui le portent. Ainsi, d'après nous cette prédiction selon laquelle tous les activateurs de croissance devraient être maternellement éteints est fausse. La formulation génétique du conflit introduite précédemment permet de bien comprendre que la prédiction doit être relative : on s'attend à ce que davantage d'activateurs de croissance soient éteints paternellement plutôt que maternellement.

Un grand nombre de ces modèles se sont également focalisés sur l'évolution d'un albumen $1m : 1p$ vers un albumen $2m : 1p$ (Haig & Westoby 1989; Stewart-Cox *et al.* 2004; Hardling & Nilsson 2001). Haig & Westoby (1989) proposent par exemple que le changement d'un ratio $1m : 1p$ vers un ratio $2m : 1p$, en doublant les doses d'inhibiteurs de croissance, accentue l'effet des empreintes en faveur de la contribution maternelle. Ce modèle fait donc l'hypothèse que l'empreinte précède l'évolution du doublement des contributions maternelles.

Nous avons beaucoup traité de l'effet du régime de reproduction sur l'intensité du conflit. Un autre facteur qui pourrait faire varier l'intensité du conflit est le fait que l'espèce soit annuelle ou pérenne. En effet, comme on l'a déjà précisé, le conflit porte à la fois sur la quantité de ressources allouées à la reproduction et sur la distribution de ces ressources entre les embryons. La part du conflit liée à la quantité de ressources allouées à la reproduction devrait être moins importante chez les espèces annuelles que chez les espèces pérennes, puisque les espèces annuelles retiennent *a priori* moins de ressources pour leur survie. Cependant, d'autres facteurs pourraient varier entre espèces annuelles et pérennes, qui pourraient faciliter, dans certaines conditions, l'évolution des traits liés au conflit chez les annuelles (voir Bulmer 1986; Zhang 1998; Zhang & Jiang 1998; Haig & Wilkins 2000). Il n'est donc pas évident de faire de prédiction en ce qui concerne l'effet de l'annualité et de la pérennité.

1.3.2.6 Pour résumer

Les facteurs clefs affectant la force du conflit entre m, m' et les différents p_i pour l'allocation des ressources sont :

1. le régime de reproduction ;
2. le fait que l'espèce soit annuelle ou pérenne (mais il n'est pas possible de faire de prédiction) ;
3. la relation entre quantité de ressources accumulées et augmentation de valeur sélective (qui détermine b et c dans la règle de Hamilton, Fig. 1.16).

Une lacune essentielle des modèles semble être l'absence de modèles spécifiques à l'évolution de la double fécondation.

1.3.3 Théorie de l'hétérosis

L'hétérosis est l'augmentation de valeur sélective qui résulte de l'augmentation d'hétérozygotie. L'hétérosis peut résulter du masquage des mutations délétères lorsqu'elles sont récessives, ou de la superdominance.

1.3.3.1 Prédictions verbales

La plupart des prédictions de la théorie de l'hétérosis sont verbales. D'après Brink & Cooper (1940, mais voir aussi Brink & Cooper 1947; Tiffney 1981; Takhtajan 1991; Donoghue & Scheiner 1992; Stebbins 1976), le passage d'une accumulation des ressources de pré à post fécondation ouvre la possibilité d'accumuler les ressources dans un tissu diploïde, porteur de deux copies de chaque gène, donc potentiellement hétérozygote et donc moins sensible à l'occurrence de mutations délétères. Si l'accumulation des ressources avant fécondation est liée à une pression de sélection entrainant l'accélération du développement, l'hétérosis apporté par la double fécondation pourrait accélérer l'accumulation des ressources dans le tissu nourricier (Donoghue & Scheiner 1992; Stebbins 1976). Les mêmes auteurs suggèrent que l'hétérosis pourrait être encore accru par la polysporie (Brink & Cooper 1940; Willson & Burley 1983).

Friedman *et al.* (2008) émettent des prédictions verbales sur l'hétérozygotie attendue pour les différents types de tissus nourriciers. D'après leurs prédictions, l'hétérozygotie augmente dans la direction : tissus monosporiques → tissus bisporiques → tissus tétrasporiques. Dans ces calculs, l'hétérozygotie ne dépend pas de la fréquence des différentes contributions génétiques, mais seulement de leurs présences/absences. De ce fait, le passage d'un albumen $1m : 1p$ à un albumen $2m : 1p$ ne change pas l'hétérozygotie. Il n'est pas évident qu'il en soit de même pour l'hétérosis.

1.3.3.2 Modèles

Il n'y a pas de modèle qui se penche sur les prédictions qui peuvent être faites sur les conséquences des changements d'hétérozygotie sur l'évolution de la double fécondation et des contributions maternelles (doublement et polysporie).

Par contre, il y a quelques modèles de l'évolution de l'empreinte parentale. Spencer & Williams (1997) et Spencer & Williams (1997) montrent que l'évolution de l'empreinte parentale n'est attendue que pour des locus présentant des mutations délétères dominantes, sauf si le taux de recombinaison est faible. L'occurrence d'empreinte parentale sur des locus présentant des mutations récessives représente un coût, lié à l'exposition accrue de ces allèles délétères. La plupart des mutations délétères sont généralement récessives (Les valeurs de dominance vont généralement de 0.1 à 0.4, voir par exemple Garcia-Dorado & Caballero 2000; Vassilieva *et al.* 2000; Szafraniec *et al.* 2003). Pearce & Spencer (1992) et Spencer & Williams (1997) en concluent que l'empreinte parentale représente un coût du point de vue de l'hétérozygotie, et qu'il est improbable qu'elle ait évolué sans pression de sélection supplémentaire, telle qu'un conflit sur l'allocation des ressources.

La théorie de l'hétérosis a été beaucoup moins explorée que la théorie du conflit en ce qui concerne l'évolution de la constitution génétique des tissus nourriciers. Tout est à faire du point de vue de la modélisation pour établir des prédictions.

1.4 Orientation de la thèse

L'évaluation des pouvoirs explicatifs de la théorie de l'hétérosis et de la théorie du conflit sur l'allocation pour l'évolution des tissus nourriciers de la graine requiert deux étapes : l'étude théorique et la confrontation aux données. Un grand nombre de prédictions théoriques ont été énoncées concernant la théorie du conflit sur l'allocation des ressources, alors que très peu l'ont été, et seulement verbalement, pour la théorie de l'hétérosis. Dans cette thèse, nous développons donc d'abord un ensemble de modèles visant à étudier les prédictions de la théorie de l'hétérosis (chapitre 2).

La théorie du conflit sur l'allocation a été confrontée à un certain nombre de données. Cependant, la plupart des données qui confortent cette théorie concernent l'empreinte parentale chez les mammifères. La théorie de l'hétérosis n'a été confrontée à quasiment aucune donnée. Rassembler les données concernant les tissus nourriciers de la graine chez les plantes est donc une étape essentielle pour tester dans quelle mesure les deux théories s'appliquent aux plantes. Nous exploitons ensuite les données disponibles dans la littérature, que nous avons rassemblées, pour établir des corrélations entre l'état des tissus nourriciers et différents traits tels que le mode de reproduction ou la dépression de consanguinité, et pour tester quelques prédictions des deux théories que nous avons présentées (chapitre 3).

L'une des hypothèses implicites de la théorie du conflit sur l'allocation est que les ressources qui ne sont acquises par un embryon ne sont plus disponibles pour les autres embryons, c'est-à-dire qu'il y a de la compétition. Nous avons mis en place une expérience visant à étudier l'allocation des ressources chez le maïs, ce qui nous permet d'aborder la relation entre le conflit sur l'allocation des ressources et les effets mâles et femelles sur la taille des graines (chapitre 4).

Parmi les spécificités du tissu nourricier, l'empreinte parentale est celle qui a le plus stimulé la formulation de diverses théories. Les données concernant l'empreinte parentale chez les plantes n'ont jamais été rigoureusement confrontées aux prédictions de la théorie du conflit sur l'allocation des ressources. Nous confrontons donc ces données non seulement aux prédictions de la théorie du conflit, mais aussi aux prédictions des théories alternatives. Deux de ces théories se démarquent (chapitre 5).

En conclusion, nous faisons l'inventaire des faits allant en faveur et en défaveur des théories du conflit et de l'hétérozygotie, et énonçons quelques perspectives possibles.

Chapitre 2

Développements de la théorie de l'hétérozygotie

D'après Brink & Cooper (1940) et Willson & Burley (1983), la double fécondation crée de l'hétérosis dans le tissu nourricier, parce qu'elle le rend hétérozygote. L'hétérosis est l'augmentation de valeur sélective qui résulte du masquage des mutations délétères et de la superdominance dans les hétérozygotes. Les changements dans le nombre de contributions maternelles de l'albumen devraient eux aussi modifier l'hétérozygotie et l'hétérosis (Willson & Burley 1983), de même que l'empreinte parentale (Pearce & Spencer 1992; Spencer & Williams 1997).

Comme précisé dans le chapitre d'introduction, peu de modèles mathématiques fournissent des prédictions sur l'évolution de l'expression des contributions mâles et femelles dans le cadre de la théorie de l'hétérozygotie, et ils sont tous focalisés sur l'évolution de l'empreinte parentale (Pearce & Spencer 1992; Spencer & Williams 1997). Dans ce chapitre, nous proposons donc des modèles afin d'établir des prédictions mathématiques pour la théorie de l'hétérozygotie.

2.1 Introduction

Il existe des modèles qui s'intéressent aux changements d'hétérozygotie pour d'autres traits que l'empreinte parentale. Ces modèles permettent d'avoir des attendus sur une transition d'une sélection haploïde vers une transition diploïde : il s'agit des modèles qui permettent d'étudier l'évolution des durées relative des phases haploïdes et diploïdes dans les cycles de vie (Perrot *et al.* 1991; Otto & Goldstein 1992; Otto 1993). Le modèle d'Otto & Goldstein (1992) est un modèle qui suppose deux locus : un locus où apparaissent des mutations délétères (locus de viabilité) et un locus qui modifie la probabilité d'entrer en méiose, donc de subir la sélection en phase haploïde ou diploïde.

L'ensemble de ces modèles montrent que les conditions dans lesquelles une phase diploïde prédominante évolue dépendent du niveau de dominance des mutations délétères. L'état diploïde n'évolue que si les mutations délétères sont récessives. Mais cette condition n'est pas suffisante. Otto & Goldstein (1992) montrent que les mutations délétères récessives s'accumulent en déséquilibre de liaison avec les modificateurs de ploïdie (qui peuvent être des gènes impliqués dans la régulation de la méiose par exemple). Cela représente un coût, mais ce coût est d'autant plus faible que la recombinaison est forte. Il

faut donc également que la recombinaison soit forte pour que la phase diploïde prédomine dans le cycle de vie.

D'autres modèles ont permis d'évaluer l'effet du régime de reproduction sur l'évolution des cycles de vie (Otto & Marks 1996). Dans ces modèles, l'autofécondation diminue l'efficacité de la recombinaison, et limite donc les conditions dans lesquelles la diploïdie peut évoluer.

Nous nous inspirons de ces modèles pour étudier l'évolution de la double fécondation, du doublement des contributions maternelles, de la polysporie et du périsperme. Nous construisons des modèles à deux-locus. Un des locus modifie le type de tissu nourricier, et l'autre locus détermine la viabilité.

Dans une première section, nous étudions l'évolution de la double fécondation et des contributions maternelles dans une population panmictique. Dans une deuxième section, nous étudions l'effet du régime de reproduction sur l'évolution de la double fécondation. Dans une troisième section, nous réinterprétons le modèle de l'évolution de la double fécondation pour généraliser les prédictions existantes de l'évolution de l'empreinte parentale. Enfin, nous traitons l'évolution de la polysporie et du périsperme. L'ensemble de ces modèles ont été réalisés suivant la même méthodologie. Nous présentons celle-ci en premier lieu.

2.2 Méthodologie

Nous construisons des modèles à deux-locus. Le locus modificateur détermine la probabilité de la double fécondation, du doublement des contributions maternelles, de la participation de plusieurs produits méiotiques à la formation du gamétophyte (polysporie), et/ou la probabilité de stocker les ressources dans le périsperme plutôt que dans l'albumen, selon les modèles. En ce qui concerne la polysporie, nous avons choisi de traiter l'évolution d'un état $mm' : 1p$. Ce faisant, le locus modificateur détermine la probabilité que le tissu nourricier soit dans un état ($1m : 0p$, $1m : 1p$, $2m : 1p$, $mm' : 0p$ ou $mm' : 1p$) plutôt qu'un autre. Afin d'explorer au mieux les conditions de l'évolution des différents tissus nourriciers, nous confrontons chacun de ces états deux-à-deux, bien que certaines de ces confrontations (par exemple $1m : 0p$ vs $mm' : 1p$) correspondent à plusieurs transitions (ici la double fécondation et la polysporie).

On peut synthétiser les différents modèles possibles par le tableau 2.1, où tous les états sont confrontés deux à deux.

	Etat 2			
Etat 1	1m :1p	2m :1p	mm' :0p	mm' :1p
1m :0p	DF	DFbis	Perisp1	Polysp1
1m :1p	-	MCD	Perisp2	Polysp2
2m :1p	-	-	Perisp3	Polysp3
mm' :0p	-	-	-	PP

Table 2.1 – Ensemble des modèles correspondant à la confrontation d'états du tissu nourricier deux à deux.

Parmi les modèles de la Table 2.1, DF et MCD permettent d'étudier l'évolution de la double fécondation (transition $1m : 0p \rightarrow 1m : 1p$) et du doublement des contributions maternelles (transition $1m : 1p \rightarrow 2m : 1p$). Le modèle DFbis modélise la transition $2m :$

$0p \rightarrow 2m : 1p$, sous l'hypothèse que la double fécondation ne précède pas le doublement des contributions maternelles (cf. sous-section 1.2.1.2). Par construction, l'état $2m : 0p$ est exactement équivalent à un état $1m : 0p$ (d'où le placement de ce modèle dans le tableau). Les autres modèles correspondent à l'évolution du périsperme et de la polysporie.

Deux allèles, α et β, sont présents au locus modificateur. Pour chaque modèle, nous considérons trois modes d'expression à ce locus. Dans le premier mode d'expression, ou mode biparental, les deux allèles de ce locus sont exprimés. Dans ce cas l'état 2 (tableau 2.1) se produit avec une probabilité ν si le génotype de l'embryon est $\alpha\alpha$, avec une probabilité $\nu + \frac{\delta_\nu}{2}$ si le génotype de l'embryon est $\alpha\beta$, et avec une probabilité $\nu + \delta_\nu$ ($\in [0, 1]$) si le génotype de l'embryon est $\beta\beta$. Dans les deux autres modes d'expression, un seul allèle de ce locus est exprimé : l'allèle transmis par le gamétophyte femelle (mode maternel), ou l'allèle transmis par le pollen (mode paternel). Dans ces cas de figure, l'état 2 se produit avec une probabilité ν si l'allèle exprimé est α, et avec une probabilité $\nu + \delta_\nu$ si l'allèle exprimé est β.

Le locus de viabilité affecte la survie de la graine en fonction du génotype de l'albumen. Deux allèles sont présents à ce locus, l'allèle fonctionnel A, et un allèle délétère a, qui apparait par mutation à un taux μ. La dominance de a augmente avec le ratio "nombre de copies de a/ nombre de copies total". Ainsi, à l'état homozygote, l'allèle A à une valeur sélective de 1. L'allèle a à une valeur sélective de $1-s$ ($0 < s < 1$). A l'état hétérozygote, la valeur sélective est intermédiaire : elle est de $1-hs$ pour les hétérozygotes aA, $1-h_1 s$ pour les hétérozygotes aaA et $1 - h_2 s$ pour les hétérozygotes aAA, avec $0 < h_2 < h < h_1 < 1$ (Table 3.3).

Génotype	Valeur sélective
A	1
a	$1 - s$
AA	1
Aa	$1 - hs$
aa	$1 - s$
AAA	1
Aaa	$1 - h_1 s$
AAa	$1 - h_2 s$
aaa	$1 - s$

Table 2.2 – Table des valeurs sélectives en fonction du génotype du tissu nourricier.

Le génotype du tissu nourricier dépend de l'état du tissu nourricier ($1m : 0p$, $1m : 1p$, $2m : 1p$, $mm' : 0p$ ou $mm' : 1p$) et des allèles transmis par les parents. Par exemple, pour un tissu nourricier $2m : 1p$, le génotype AAa ne pourra être rencontré que si un allèle A est transmis par la femelle, et un allèle a est transmis par le mâle.

Nous suivons les fréquences des gamètes ou des sporophytes (selon les modèles), en temps discret (génération par génération), et en suivant le cycle de vie décrit dans le chapitre d'introduction (Fig. 1.1). Au moment de la méiose, les deux locus recombinent à un taux r. Certains points de méthodologie étant spécifiques aux différents modèles, nous les détaillons lors de leur présentation.

2.3 Évolution de la double fécondation et des contributions maternelles

Dans cette section, nous nous intéressons à l'évolution de la double fécondation (DF), qui modifie le ratio des contributions génétiques parentales de $1m : 0p$ à $1m : 1p$, et à l'évolution vers un doublement des contributions maternelles (MCD), qui modifie le ratio des contributions de $1m : 1p$ à $2m : 1p$.

2.3.1 Méthodes

Dans les modèles DF et MCD, nous suivons l'évolution des quatre fréquences de gamétophytes (AM, aM, Am, am). Nous écrivons les récursions (fréquence à un temps t en fonction des fréquences à un temps t-1) pour ces quatre fréquences. Les expressions résultantes étant complexes, nous utilisons deux méthodes pour les simplifier et les interpréter. L'une de ces méthodes est l'analyse de stabilité. Dans l'analyse de stabilité, on fait l'hypothèse que la population est initialement monomorphe pour le locus modificateur (l'allèle α est seul présent). On calcule la fréquence de l'allèle délétère dans cette population (ou équilibre mutation-sélection), puis on étudie les conditions de l'invasion initiale d'un allèle modificateur mutant (l'allèle β) en supposant que l'équilibre mutation-sélection (i.e. p_a, la fréquence d'équilibre de l'allèle délétère) est peu affecté par ce changement. L'autre méthode est l'approximation QLE (*Quasi Linkage Equilibrium*). Pour cette approche, il est nécessaire de transformer les quatre fréquences de gamétophytes en deux fréquences alléliques et un déséquilibre de liaison. On suppose que le déséquilibre de liaison évolue rapidement par rapport aux fréquences des allèles, ce qui permet de calculer un pseudo-équilibre (ou QLE) que l'on insère dans les expressions décrivant les changements de fréquences des allèles. Contrairement à l'analyse de stabilité, l'approche QLE permet d'étudier l'évolution de la fréquence des allèles au locus modificateur sans faire d'hypothèse sur leurs fréquences initiales. Cependant, l'approche QLE est fausse lorsque la recombinaison est trop faible.

2.3.2 Résultats

C'est le niveau d'exposition des allèles délétères à la sélection chez les hétérozygotes qui détermine en grande partie les conditions d'invasion de la double fécondation et des contributions maternelles (Fig. 2.1). Trois mesures d'expositions nous intéressent (E_{mat}, E_{pat}, E_{moy}). Soit E_{mat} le niveau d'exposition des mutations délétères lorsqu'elles sont héritées maternellement (donc exprimées par le chromosome maternel), et E_{pat} le niveau d'exposition des mutations délétères lorsqu'elles sont héritées paternellement (donc exprimées par le chromosome paternel). Lorsqu'un embryon hérite l'allèle a paternellement et A maternellement, l'allèles a a :

- une exposition nulle (0) dans un tissu nourricier $1m : 0p$;
- une exposition égale à sa dominance, h, dans un tissu nourricier $1m : 1p$;
- une exposition de h_2 dans un tissu nourricier $2m : 1p$ (dominance de a dans le génotype AAa, résultant de la transmission paternelle de a)

Lorsqu'un embryon hérite l'allèle a maternellement et A paternellement, l'allèle a a :

- une exposition totale (1) dans un tissu nourricier $1m : 0p$;
- une exposition égale à sa dominance, h, dans un tissu nourricier $1m : 1p$;

- une exposition de h_1 dans un tissu nourricier $2m : 1p$ (dominance de a dans le génotype aaA, résultant de la transmission maternelle de a)

La probabilité que l'allèle délétère soit hérité maternellement est la même que la probabilité que l'allèle délétère soit hérité paternellement. Par conséquent, on peut définir une troisième mesure d'exposition, E_{moy}, l'exposition moyenne des allèles délétères à la sélection :

$$E_{moy} = \frac{E_{mat} + E_{pat}}{2} \tag{2.1}$$

L'exposition moyenne des allèles délétères est donc $(0 + 1)/2 = 1/2$ dans un tissu nourricier $1m : 0p$, $(h + h)/2 = h$ dans un tissu nourricier $1m : 1p$, et $(h_1 + h_2)/2 = h_m$ dans un tissu nourricier $2m : 1p$.

L'étude des équations (cf. Annexe A) indique que deux termes déterminent la direction de la sélection. Un terme de sélection directe, et un terme de sélection indirecte. Lorsque la recombinaison est élevée, seule la sélection directe compte. Dans ce cas, l'état qui minimise le niveau moyen d'exposition évolue. Lors d'une transition $1m : 0p \rightarrow 1m : 1p$, l'état $1m : 1p$ évolue donc lorsque $h < \frac{1}{2}$, et lors d'une transition $1m : 1p \rightarrow 2m : 1p$, l'état $2m : 1p$ évolue lorsque $h_m < h$. Par contre, lorsque la recombinaison est faible, il existe un coût lié à l'accumulation des mutations délétères en déséquilibre de liaison avec le modificateur. Dans ce cas, un effet indirect s'ajoute à l'effet direct, et le niveau d'exposition respectif des chromosomes maternels et paternels (plus seulement le niveau moyen d'exposition) joue un rôle. Si l'exposition d'un chromosome est diminuée lors d'une transition donnée, l'innovation évolue dans des gammes de valeurs de dominance d'autant plus faibles que la recombinaison diminue. Par contre, si l'exposition d'un chromosome est augmentée, l'innovation évolue dans des gammes de valeurs de dominance d'autant plus grandes que la recombinaison diminue. Lors d'une transition $1m : 0p \rightarrow 1m : 1p$, l'état $1m : 1p$ diminue l'exposition du chromosome maternel (car $h < 1$), et augmente l'exposition du chromosome paternel ($h > 0$). L'état $1m : 1p$ évolue donc dans des gammes de valeurs de dominance plus restreintes sous contrôle femelle que sous contrôle mâle. Lors d'une transition $1m : 1p \rightarrow 2m : 1p$, l'état $2m : 1p$ augmente l'exposition du chromosome maternel ($h_1 > h$), et diminue l'exposition du chromosome paternel ($h_2 < h$). L'état $2m : 1p$ évolue donc dans des gammes de valeur de dominance plus petites sous contrôle mâle que sous contrôle femelle.

2.3.3 Interprétation et discussion

L'effet direct est simple à comprendre : il favorise l'état dans lequel les mutations délétères sont le plus masquées en moyenne à la sélection. L'effet indirect est plus complexe, car il dépend du déséquilibre de liaison, qui change selon que le contrôle soit mâle ou femelle. Le déséquilibre de liaison change d'une part parce que l'exposition des allèles délétères n'est pas le même sur le chromosome mâle et le chromosome femelle, et d'autre part parce que la probabilité qu'un modificateur reste physiquement associé à un allèle délétère non purgé dépend du fait qu'il est sur le même chromosome (i.e. en *cis*) ou non (i.e. en *trans*) par rapport à cet allèle délétère. En effet, imaginons un allèle délétère qui n'est pas purgé coté paternel du fait d'un tissu $1m : 0p$ dans le modèle DF. Le modificateur femelle est en *trans* par rapport à cet allèle délétère. Ainsi, lors de la formation de nouveaux gamètes, le modificateur femelle ne reste associé à l'allèle délétère que si les

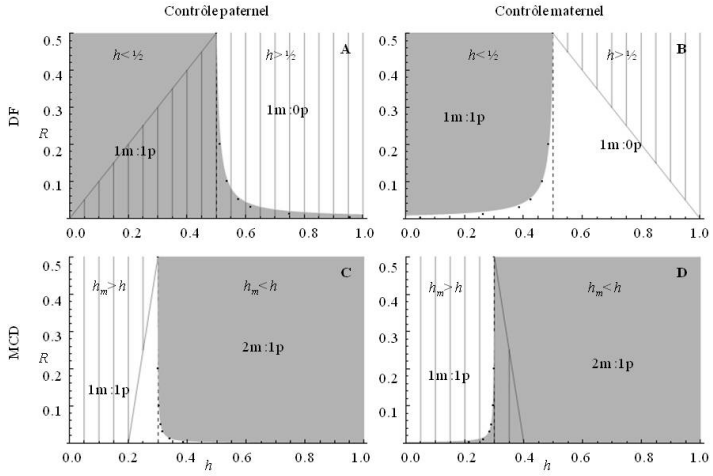

Figure 2.1 – Évolution de la double fécondation et des contributions maternelles d'après la résolution analytique en QLE. La direction de la sélection est indiquée pour différentes valeurs de h (axe x) et R (axe y). Dans A et B (modèle DF), $1m : 1p$ est favorisé dans l'aire grisée, alors que $1m : 0p$ est favorisé dans l'aire blanche. Dans C and D (modèle MCD), $2m : 1p$ est favorisé dans l'aire grisée, alors que $1m : 1p$ est favorisé dans l'aire blanche. A, C et B, D illustrent respectivement une expression paternelle et une expression maternelle du modificateur. Les aires hachurées correspondent à un déséquilibre de liaison négatif (le modificateur est associé au chromosome purgé). La sélection directe favorise $1m : 1p$ à gauche de la ligne pointillée (qui représente $h = 1/2$, A et B), et favorise $2m : 1p$ à droite de la ligne pointillée (qui représente $h = h_m = 0.3$, C et D). L'approximation QLE n'est plus appropriée lorsque r est très petit (les points représentent les valeurs critiques de h obtenues à partir de simulations exactes). Les valeurs de paramètre sont $s = 5.10^{-2}$, $\mu = 10^{-3}$, $\nu = 0.1$, $\delta_\nu = 0.01$, $h_1 = 0.2$ et $h_2 = 0.4$.

deux locus recombinent, ce qui a lieu à un taux r. Le modificateur mâle, en revanche, est en *cis* par rapport à cet allèle délétère. Ainsi, lors de la formation de nouveaux gamètes, le modificateur mâle reste associé à l'allèle délétère si les deux locus ne recombinent pas, ce qui a lieu à un taux $1 - r$. Lorsque $r = 1/2$, la probabilité d'être associé à l'allèle délétère est la même pour les deux modificateurs, et il n'y a plus de différence dans le coût lié à l'accumulation des allèles délétères. C'est pourquoi quand $r = 1/2$ les conditions dans lequelles la double fécondation évolue ne sont plus sexe-dépendantes. L'issue de la sélection dépend donc d'un équilibre entre le bénéfice direct, et le coût indirect et sexe-dépendant quand r est différent de $1/2$, de masquer les allèles délétères.

D'après nos résultats, la double fécondation devrait évoluer lorsque les mutations délétères sont récessives, au moins lorsque la recombinaison est forte. C'est une condition réaliste puisque les valeurs de dominance vont généralement de 0.1 à 0.4 (voir Garcia-Dorado & Caballero 2000; Vassilieva *et al.* 2000; Szafraniec *et al.* 2003). On s'attend donc à ce que la double fécondation soit très répandue. Le doublement des contributions maternelles est susceptible d'évoluer si l'exposition moyenne des mutations délétères dans un tissu $2m : 1p$ est inférieure à l'exposition moyenne des mutations délétères dans un tissu $1m : 1p$. On s'attend à ce que l'exposition moyenne d'une mutation délétère a dans un tissu aA soit quasiment médiane à son exposition dans un tissu aaA et dans

un tissu aAA, donc à ce que les deux expositions moyennes soient proches. On peut donc s'attendre à ce que $h < h_m$ chez certaines espèces, et $h > h_m$ chez d'autres, donc à trouver du polymorphisme chez les angiospermes. L'ensemble de ces prédictions sont compatibles avec la distribution observée des tissus nourriciers, puisque la double fécondation se produit chez quasiment toutes les angiospermes, et il existe à la fois des angiospermes pour lesquelles les contributions maternelles sont doublées, et d'autres pour lesquelles elles ne le sont pas (bien que les espèces $2m : 1p$ soient les plus fréquentes, cf. chapitre d'introduction).

Sous l'hypothèse que la double fécondation ait évolué postérieurement au doublement des contributions maternelles, la double fécondation correspondrait à la transition $2m : 0p \rightarrow 2m : 1p$. Dans le modèle DF2, l'état $2m : 0p$ évolue en $2m : 1p$ lorsque $h_m < 1/2$ lorsque la recombinaison est forte. Lorsque la recombinaison est faible, cette condition est réduite et $2m : 1p$ n'évolue plus que pour des valeurs de h_m qui sont encore plus petites sous contrôle maternel, et cette condition est élargie et $2m : 1p$ évolue également pour des valeurs de h_m qui sont plus grandes que $1/2$ sous contrôle paternel. Les mécanismes impliqués dans le modèle DF2 sont les mêmes que pour DF et MCD, c'est pourquoi nous ne les détaillons pas.

Le fait que les conditions dans lesquelles la double fécondation et le doublement des contributions maternelles évoluent changent selon que le contrôle du modificateur est mâle ou femelle nous permet de décrire un **nouveau conflit mâle-femelle**, conflit qui porte sur l'exposition des allèles délétères à la sélection. Selon les conditions, ce conflit résulte dans un état optimum du tissu nourricier qui n'est pas le même selon le parent (Figure 2.2).

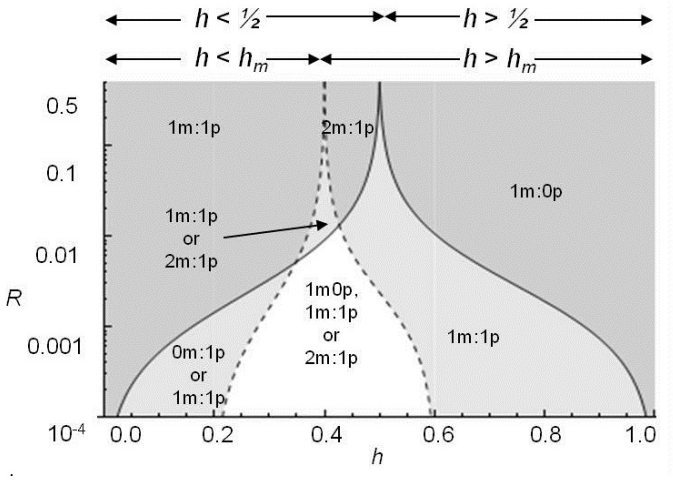

Figure 2.2 – Évolution couplée de la double fécondation des contributions maternelles lorsque $h_m < 1/2$ ($h_1 = 0.6$ et $h_m = 0.2$). Les lignes pleines indiquent les valeurs de h critiques du modèle DF en cas de contrôle maternel et en cas de contrôle paternel. Les lignes pointillées indiquent les valeurs de h critiques du modèle MCD en cas de contrôle maternel et en cas de contrôle paternel. Par ailleurs, $s = 10^{-1}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$.

Si les conditions dans lesquelles ce conflit est prédit sont remplies (r faible, valeurs de dominance intermédiaires), on pourrait s'attendre à ce que les gènes impliqués dans

la double fécondation et le doublement des contributions maternelles montrent les signes d'une coévolution antagoniste, ce qui est également une des prédictions de la théorie du conflit sur l'allocation des ressources. Si des signes de coévolution antagoniste sont observés pour ces gènes, ils ne pourront donc pas servir à discriminer la théorie du conflit sur l'allocation des ressources de la théorie de l'hétérozygotie.

L'article correspondant à l'étude détaillée des modèles DF et MCD ayant été publié, voir le manuscrit en Annexe A pour davantage de détails sur les mécanismes, et plusieurs autres points de discussion relatifs à ces modèles.

2.4 Évolution de l'empreinte parentale

Il existe deux types d'empreintes parentales : celles pour lesquelles c'est l'allèle paternel qui est sous exprimé, et celles pour lesquelles c'est l'allèle maternel qui est sous exprimé. Bien que l'empreinte soit *a priori* une innovation très différente de la double fécondation ou d'un changement dans le nombre de contributions maternelle, l'ensemble de ces traits ont pour point commun de changer le ratio des contributions maternelles et paternelles.

En ce qui concerne l'empreinte, lorsque l'allèle paternel est éteint, l'expression du gène à empreinte est du type $1m : 0p$ ou $2m : 0p$ selon que l'empreinte est animale ou végétale, et si elle est végétale, selon le tissu nourricier. Lorsque c'est l'allèle maternel est éteint, l'empreinte est du type $0m : 1p$. Dans les modèles DF et MCD nous avons choisi de modéliser l'évolution de la probabilité de la double fécondation ou du doublement des contributions maternelles. Pour un modèle d'évolution de l'empreinte, c'est l'évolution du taux d'expression du locus sous exprimé qu'il faut modéliser ($0m : 1p$ correspond par exemple à une extinction maternelle complète, mais l'extinction a probablement été progressive). Il est possible de réinterpréter les résultats des modèles DF et DF2 pour étudier l'évolution de l'empreinte, à condition que ν et δ_ν soient vus comme des taux d'expression, au lieu de probabilités.

2.4.1 Extinction paternelle

En cas d'extinction de l'allèle paternel, l'empreinte parentale correspond à un état $1m : 0p$ ou $2m : 0p$ locus-spécifique. L'évolution de l'empreinte pour laquelle l'extinction est paternelle ($1m : 1p \leftrightarrow 1m : 0p$ ou $2m : 1p \leftrightarrow 2m : 0p$) est donc l'inverse de l'évolution de la double fécondation ($1m : 0p \leftrightarrow 1m : 1p$, modèle DF, ou $2m : 0p \leftrightarrow 2m : 1p$, DF2).

Dans le modèle DF, l'extinction paternelle (état $1m : 0p$) n'évolue que si les mutations délétères sont dominantes ou, sous contrôle maternel, si la recombinaison est faible (fig. 2.1).

Dans le modèle DF2, l'état $2m : 1p$ évolue en $2m : 0p$ lorsque $h_m > 1/2$ lorsque la recombinaison est forte. Lorsque la recombinaison est faible, cette condition est élargie et $2m : 0p$ évolue également pour des valeurs de h_m plus faibles en cas de contrôle femelle, et cette condition est réduite et $2m : 0p$ n'évolue plus que pour des valeurs de h_m élevées en cas de contrôle mâle.

2.4.2 Extinction maternelle

En cas d'extinction de l'allèle maternel, l'empreinte parentale correspond à un état locus-spécifique $0m : 1p$. La transition $1m : 1p \leftrightarrow 0m : 1p$ a été étudiée par Spencer & Williams (1997) dans un modèle où le modificateur est sous contrôle maternel. Dans ce modèle, l'état $0m : 1p$ évolue, comme l'état $1m : 0p$ sous contrôle paternel dans le modèle DF, lorsque les mutations délétères sont dominantes et que la recombinaison n'est pas trop faible. Si $0m : 1p$ évolue de la même manière qu'évolue $1m : 0p$ sous contrôle paternel, c'est parce que dans ces deux modèles l'allèle est éteint en *cis* par rapport à l'allèle modificateur exprimé.

Sous contrôle paternel, l'allèle maternel est éteint en *trans* par rapport à l'allèle modificateur exprimé. Par conséquent, sous contrôle paternel, l'empreinte $0m : 1p$ devrait évoluer à partir d'un état $1m : 1p$ dans les mêmes conditions que l'état $1m : 0p$ sous contrôle maternel dans le modèle DF. Ainsi, on s'attend à ce que $0m : 1p$ évolue sous contrôle paternel quand les mutations sont dominantes ou lorsqu'elles sont récessives mais que la recombinaison est faible.

Enfin, l'empreinte $0m : 1p$ devrait évoluer à partir d'un état $2m : 1p$ lorsque $h_m > 1/2$ si la recombinaison est forte. Si la recombinaison est faible, cette condition devrait être élargie (et $0m : 1p$ devrait évoluer pour des valeurs de h_m inférieures à $1/2$) en cas de contrôle mâle, et cette condition devrait être restreinte (et $0m : 1p$ devrait évoluer pour des valeurs de h_m encore plus grandes que $1/2$) en cas de contrôle femelle.

2.4.3 Discussion

Comme on l'a précisé, il est rare que les mutations délétères soient dominantes. De même, il est peu probable que $h_m > 1/2$, puisque h_m devrait être proche de h. L'empreinte parentale devrait donc rarement évoluer d'après ces réinterprétations des modèles. Pour expliquer pourquoi l'empreinte évolue, les hypothèses suivantes sont possibles :

- L'empreinte ne touche que les locus dont l'ensemble des mutations délétères sont dominantes. Cette hypothèse est compatible avec le fait que peu de gènes sont à empreinte, mais l'idée qu'il puisse y avoir des locus pour lesquels la plupart des mutations délétères sont dominantes reste à démontrer.
- L'empreinte résulte d'un conflit mâle-femelle sur le ratio d'expression des contributions mâles et femelles. Une telle hypothèse requiert que l'extinction maternelle soit sous contrôle paternel, et inversement.
- L'empreinte est le résultat d'une autre pression de sélection (conflit sur l'allocation des resources, ajustement des effets doses après l'évolution des contributions parentales, etc. cf. sous-section 5.2), suffisante pour contrebalancer le coût qu'elle représente en termes d'exposition des allèles délétères.

Si l'empreinte est l'objet d'un conflit mâle-femelle sur l'allocation des ressources, l'allèle sous-exprimé devrait finir par être complètement éteint (Mochizuki *et al.* 1996). Or ce n'est pas toujours le cas. Dans certains cas, l'expression des deux allèles est différente, mais l'extinction de l'allèle sous-exprimé n'est pas complète.

Si on suppose que la théorie du conflit sur l'allocation est bien responsable de l'évolution des empreintes, il est raisonnable de penser que le coût lié à l'expression haploïde des allèles délétères est peut être la pression de sélection qui contraint l'évolution d'une extinction complète de la copie sous-exprimée. De même, parfois, il y a empreinte ou non selon les allèles présents aux locus cibles (empreintes alléliques, voir Table 1.1). Il serait

intéressant de tester si ce polymorphisme d'empreintes est corrélé à des différences de dominance pour ces allèles.

2.5 Éffet de l'autofécondation

Nous réalisons deux modèles pour la prise en compte de l'effet de l'autofécondation : un modèle dans lequel l'autofécondation est gamétophytique, et un modèle où l'autofécondation est sporophytique. Bien que l'autofécondation gamétophytique ne soit pas un mécanisme d'autofécondation que l'on rencontre dans les populations naturelles d'angiospermes, l'analyse du modèle dans lequel l'autofécondation est gamétophytique permet de comprendre les effets de l'augmentation d'homozygotie. Ces modèles sont en tous points similaires à ceux décrits ci-dessus, à part que les modèles précédent sont en régime panmictique. L'effet de l'autofécondation dans les modèles d'évolution du cycle de vie est de diminuer l'efficacité de la recombinaison (Otto & Marks 1996).

2.5.1 Résultats

Dans le modèle gamétophytique, la double fécondation évolue dans une gamme de valeurs de dominance de plus en plus grande quand l'autofécondation devient forte en cas de contrôle paternel, et dans une gamme de plus en plus petite en cas de contrôle maternel. L'étendue du conflit mâle-femelle sur l'exposition des allèles délétères décrit section 2.3 est donc de plus en plus grande au fur et à mesure que l'autofécondation gamétophytique augmente (Figure 2.3).

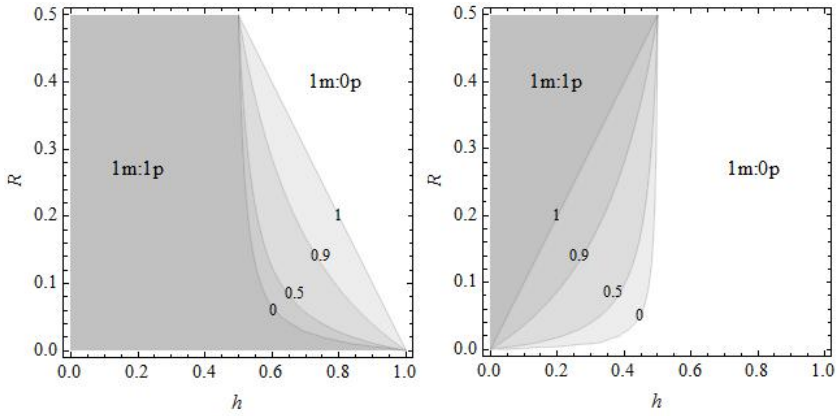

Figure 2.3 – Résultats de simulations pour l'évolution de la double fécondation, quand l'autofécondation est gamétophytique. Les chiffres qui sont sur le graphique indiquent les taux d'autofécondations utilisés pour les différentes simulations. Par ailleurs, $s = 5.10^{-2}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$.

Dans le modèle où l'autofécondation est sporophytique, les résultats sont inversés par rapport au modèle gamétophytique : quand le taux d'autofécondation sporophytique augmente, les stratégies maternelle et paternelle convergent vers la stratégie biparentale définie dans les modèles DF et MCD (voir Annexe A). La seule et unique condition

pour que la double fécondation évolue est donc $h < 1/2$. Ainsi, il existe un conflit mâle-femelle sur l'exposition des allèles délétères si le père et la mère sont distincts, mais ce conflit disparait quand il s'agit du même individu (Figure 2.4). Nous n'avons pas pu dériver de résultats analytiques pour ce modèle, et nous ne pouvons donc pas connaître l'origine de ses différences avec le modèle gamétophytique, bien qu'on puisse soupçonner qu'elle est liée à la production de 50% d'hétérozygotes dans le cas de l'autofécondation sporophytique.

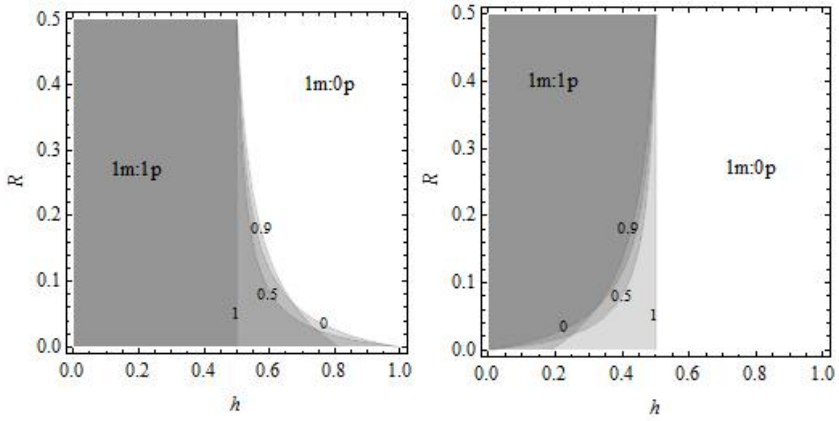

Figure 2.4 – Résultats de simulations pour l'évolution de la double fécondation quand l'autofécondation est sporophytique. Les chiffres qui sont sur le graphique indiquent les taux d'autofécondations utilisés pour les différentes simulations. Par ailleurs, $s = 10^{-1}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$.

2.5.2 Discussion

L'autofécondation gamétophytique est plus simple à modéliser et à étudier de manière analytique que l'autofécondation sporophytique. C'est pourquoi les modèles d'autofécondation gamétophytique sont souvent préférentiellement utilisés pour étudier l'effet de l'autofécondation. Cependant, les résultats du modèle d'autofécondation gamétophytique ne concordent pas avec les résultats du modèle sporophytique dans le cas spécifique de l'évolution de la double fécondation. C'est également le cas pour l'évolution de la recombinaison (Roze & Lenormand 2005). De ce fait, d'une manière générale, il parait important d'interpréter les résultats des modèles gamétophytiques avec prudence si ces modèles ne visent pas à étudier des traits qui évoluent chez des taxons qui pratiquent bien l'autofécondation gamétophytique.

Puisque l'autofécondation est sporophytique chez les angiospermes, nous devons utiliser les résultats du modèle sporophytique pour faire une prédiction biologique des conditions de l'évolution de la double fécondation. Étant donné que les mutations délétères sont très généralement récessives, nous pouvons prédire le maintien de la double fécondation quelque soit le régime de reproduction de l'espèce à laquelle on s'intéresse. Cette prédiction est intéressante, car la théorie du conflit sur l'allocation des ressources ne prédit pas le maintien de ces traits pour des plantes autofécondantes.

L'article correspondant à cette section étant en cours de rédaction, voir le manuscrit

en Annexe C pour davantage de détails techniques.

2.6 Évolution du périsperme et de la polysporie

Nous avons réalisé sept modèles pour étudier les effets de masquage en cas de polysporie (albumen $1m1m' : 1p$) et dans le périsperme ($1m1m'$). Nous nous attendions à ce que l'expression d'une contribution qui n'est pas transmise à l'embryon (m') se traduise par des coûts liés à l'expression d'allèles (potentiellement délétères) qui ne sont pas transmis à la génération suivante et la présence à la fois de coûts et de bénéfices liés à la diminution de l'exposition d'allèles qui eux seront transmis (m et p).

2.6.1 Méthodes

L'ajout d'une contribution qui n'est pas portée par l'embryon (m') est une difficulté technique, car du point de vue de la modélisation, la valeur sélective d'un embryon donné dépend du génotype du sporophyte de la génération précédente. Ainsi, au moment d'attribuer les valeurs sélectives nous ne devons plus distinguer seize catégories correspondant aux types d'embryons possibles, mais beaucoup plus (détail de l'attribution des valeurs sélectives, voir Annexe E).

Du fait de la complexité du modèle, nous devons suivre, au long du cycle de vie, seize fréquence sporophytiques (au lieu de quatre fréquences gamétophytiques dans les modèles de la section 2.3, DF et MCD), et nous n'avons donc pas pu dériver de résultats analytiques comme pour les modèles DF et MCD. Cependant, comme dans les modèles DF et MCD, on s'attend à observer un effet direct, lié à l'exposition moyenne des allèles délétères, et un effet indirect, qui dépend de la recombinaison et de l'exposition respectives des chromosomes (donc potentiellement du fait que le contrôle est mâle ou femelle). Pour les modèle DF et MCD, l'exposition moyenne des allèles délétères déterminait les conditions d'invasion lorsque la recombinaison était forte, et le niveau d'exposition des allèles délétères sur chacun des chromosomes modifiait ces conditions lorsque la recombinaison était faible.

2.6.1.1 Equilibres mutation-sélection et exposition moyenne des allèles délétères

Dans les modèles DF et MCD, les différents types de tissus nourriciers hétérozygotes étaient produits en fréquences égales, ce qui permettait d'obtenir E_{moy} par un calcul trivial ($E_{moy} = (E_{pat} + E_{mat})/2$). Déterminer E_{mat}, E_{pat} et E_{moy} pour le périsperme ou l'albumen polysporique n'est pas aussi trivial que pour les autres états, car i) a peut être hérité via m' en plus de m et p, ce qui fait que le tissu nourricier peut être hétérozygote si le sporophyte maternel est hétérozygote, même si l'embryon ne l'est pas ; ii) *a priori* les différents tissus nourriciers hétérozygotes ne sont pas en fréquences égales, puisque leurs fréquences dépend de la fréquence des différents sporophytes qui ont été l'objet de la sélection à la génération précédente. Cependant, dans les modèles DF et MCD les équilibres mutation-sélection étaient de la forme $\mu/E_{moy}s$ pour les tissus $1m : 0p$, $1m : 1p$, et $2m : 1p$. Calculer les équilibres mutation-sélection devrait donc nous permettre de déterminer E_{moy}.

Pour déterminer l'équilibre mutation-sélection, nous calculons la valeur sélective moyenne

de l'allèle a dans un population où la probabilité d'être dans l'état 1 (tel que défini Table 2.1) est ν et la probabilité d'être dans l'état 2 est $1-\nu$, puis nous déterminons la valeur de cette valeur sélective moyenne après mutation.

Il n'est pas garantit que dans les modèles de l'évolution de la polysporie et du périsperme, l'exposition moyenne des allèles délétères détermine seule, comme dans les modèles DF et MCD, la direction de la sélection lorsque r est fort. Pour vérifier si c'est le cas, nous réalisons des simulations.

2.6.1.2 Simulations

Pour nos simulations, nous utilisons la méthodologie suivante. Pour chaque modèle, nous commençons par simuler l'évolution des fréquences (initialement toutes fixées à 1/16) pour les paramètres $\mu = 10^{-3}$, $\nu = 0.1$,$\delta_\nu = 0.01$, $h = 0.5$, $r = 0.25$, $h_1 = 0.4$, $h_2 = 0.2$ et $s = 10^{-1}$ ou $s = 10^{-2}$, sur un nombre de générations suffisant pour que les trois premières décimales de la fréquence de l'allèle délétère ne varient plus (en général, 400 générations lorsque $s = 10^{-1}$, 1500 générations lorsque $s = 10^{-2}$). Pour chacune de ces simulations, nous relevons cette valeur (qui est l'équilibre mutation-sélection, p_a).

Nous utilisons cette valeur pour calculer les fréquences sporophytiques initiales de nouvelles simulations (fréquence des génotypes $AA = (1-p_a)^2$, $Aa = (1-p_a)p_a$, $aa = p_a^2$). Pour ces simulations, nous établissons une grille pour r d'une part (simulations pour des valeurs de 0 à 1/2 par pas de 0.01) et soit h soit h_m (h_m pour les modèles Polysp1 et Polysp3 dans lesquels il n'y a pas de tissu diploïde) d'autre part (simulations pour des valeurs de 0 à 1 par pas de 0.1). Chaque simulation tourne sur 300 générations. En ce qui concerne les autres paramètres, nous gardons les mêmes valeurs que celles utilisées pour déterminer l'équilibre mutation sélection : $\mu = 10^{-3}$, $\nu = 0.1$, $\delta_\nu = 0.01$ et $s = 10^{-1}$ ou $s = 10^{-2}$, ainsi que $h_1 = 0.4$ et $h_2 = 0.2$ pour les simulations où h est dans la grille et où un tissu triploïde est présent ($2m : 1p$ ou $mm' : 1p$), et $h_d = (h_1 - h_2)/2 = 0.2$ pour les simulations où c'est h_m qui est dans la grille.

Enfin, pour chaque simulation, nous déterminons la direction du changement de fréquence au locus modificateur entre la 300e et la 385e génération, ce qui nous permet de tracer les graphiques présentés en résultats (zone blanche, augmentation de la fréquence de α, donc évolution de l'état 1 tel que défini Table 2.1, et zone grisée, augmentation de la fréquence de β, donc évolution de l'état 2 tel que défini Table 2.1).

2.6.2 Résultats

2.6.2.1 Equilibres mutation-sélection et exposition moyenne des allèles délétères

Les équilibres mutation-sélection sont :

$$p_a^{DF} = \frac{\mu}{\left((1-\nu)\frac{1}{2} + \nu h\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.2}$$

$$p_a^{MCD} = \frac{\mu}{\left((1-\nu)h + \nu\frac{h_1+h_2}{2}\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.3}$$

$$p_a^{Perisp1} = \frac{\mu}{\left((1-\nu)\frac{1}{2} + \nu\frac{h}{2}\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.4}$$

$$p_a^{Perisp2} = \frac{\mu}{\left((1-\nu)h + \nu\frac{h}{2}\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.5}$$

$$p_a^{Perisp3} = \frac{\mu}{\left((1-\nu)\frac{h_1+h_2}{2} + \nu\frac{h}{2}\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.6}$$

$$p_a^{Polysp1} = \frac{\mu}{\left((1-\nu)\frac{1}{2} + \nu h_2\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.7}$$

$$p_a^{Polysp2} = \frac{\mu}{\left((1-\nu)h + \nu h_2\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.8}$$

$$p_a^{Polysp3} = \frac{\mu}{\left((1-\nu)\frac{h_1+h_2}{2} + \nu h_2\right) s} + o\left(\frac{\mu}{s}\right)^1 \tag{2.9}$$

De ces équilibres mutation-sélection, qui concernent des populations mixtes, on peut déduire les équilibres mutation sélection de populations pour lesquelles le tissu nourricier est fixé. Ces équilibres mutation-sélection sont de la forme $\mu/E_{moy}s$ pour les tissus $1m : 0p$, $1m : 1p$, et $2m : 1p$:

$$p_a^{1m0p} = \frac{\mu}{\frac{1}{2}s} + o\left(\frac{\mu}{s}\right)^1; \tag{2.10}$$

$$p_a^{1m1p} = \frac{\mu}{hs} + o\left(\frac{\mu}{s}\right)^1; \tag{2.11}$$

$$p_a^{2m1p} = \frac{\mu}{\frac{h_1+h_2}{2}s} + o\left(\frac{\mu}{s}\right)^1; \tag{2.12}$$

On pourrait donc penser que l'équilibre mutation-sélection devrait permettre de prédire la valeur de E_{moy} pour $1m1m' : 0p$ et $1m1m' : 1p$. Pour des populations pour lesquelles les états $1m1m' : 0p$ et $1m1m' : 1p$ sont fixés, les équilibres mutation-sélection sont :

$$p_a^{1m1m'1p} = \frac{\mu}{h_2 s} + o\left(\frac{\mu}{s}\right)^1; \tag{2.13}$$

$$p_a^{1m1m'0p} = \frac{\mu}{\frac{h}{2}s} + o\left(\frac{\mu}{s}\right)^1; \tag{2.14}$$

On peut en déduire que dans un tissu $1m1m' : 1p$, l'exposition moyenne des allèles délétères est h_2, quand dans un tissu $1m1m' : 0p$, l'exposition moyenne des allèles délétères est $h/2$.

Nous avons vu section 2.3 comment la valeur de E_{moy} était déterminée pour les tissus $1m : 0p$, $1m : 1p$, et $2m : 1p$. En ce qui concerne le périsperme $1m1m' : 0p$, l'exposition moyenne est $h/2$. On peut interpréter cette exposition moyenne de la manière suivant :

$h/2 = (0 + h)/2$, zero étant l'exposition des allèles délétères sur le chromosome paternel, et h l'exposition des allèles délétères sur le chromosome maternel.

En ce qui concerne l'albumen tétrasporique, l'équilibre mutation-sélection ne dépend que de la valeur de dominance dans les hétérozygotes AAa, h_2. Cela est probablement dû au fait que les tissus nourriciers aaA sont plus rares que les tissus nourriciers AAa. En effet, les tissus nourriciers triploïdes hétérozygotes peuvent se former de deux façons. Premièrement, deux contributions maternelles m et m' différentes l'une de l'autre (l'une est A, l'autre a) sont fécondées par une contribution paternelle, qui est a avec une fréquence p_a et A avec une fréquence $1 - p_a$. Comme a est délétère, $p_a < 1 - p_a$, donc les tissus AAa qui en résultent sont plus fréquents que les tissu aaA. Deuxièmement, deux contributions maternelles m et m' identiques sont fécondées par une contribution paternelle différente d'elles. Soit ces contributions sont aa et sont en fréquence inférieure à p_a^2 (car les sporophytes ont subit la sélection à la génération précédente), et sont fécondées par A avec une fréquence $1 - p_a$, soit les contributions maternelles sont AA et sont en fréquence supérieure à $(1-p_a)^2$, et sont fécondées par a avec une fréquence p_a. Or comme p_a est petit (nécessairement inférieur à 1/2) $p_a^2(1 - p_a) < p_a(1 - p_a)^2$.

Quelque soit la façon dont les tissus triploïdes hétérozygotes sont formés, on peut prédire que AAa est plus fréquent que aaA.

2.6.2.2 Résultats des simulations

Nous commençons par présenter les résultats des simulations qui concernent l'évolution du périsperme, poursuivons par l'évolution de l'albumen tétrasporique, et terminons par la confrontation de ces deux états l'un contre l'autre.

Perisp1 La transition $1m : 0p \leftrightarrow mm' : 0p$ (modèle Perisp1) correspond au simple ajout d'une contribution maternelle. Pour ce modèle, on ne détecte pas de différence selon que le contrôle du modificateur est mâle, femelle, ou biparental. Par contre, on observe un effet important de s sur les conditions de dominance dans lesquelles l'état $mm' : 0p$ évolue (Figure 2.5). Ainsi lorsque $s = 10^{-1}$, l'état $mm' : 0p$ n'évolue que dans une gamme restreinte de valeurs de dominance ($h < 0.1$). La gamme de valeurs de dominance dans laquelle $mm' : 0p$ évolue est cependant plus grande si l'effet des mutations délétères est moindre ($s = 10^{-2}$).

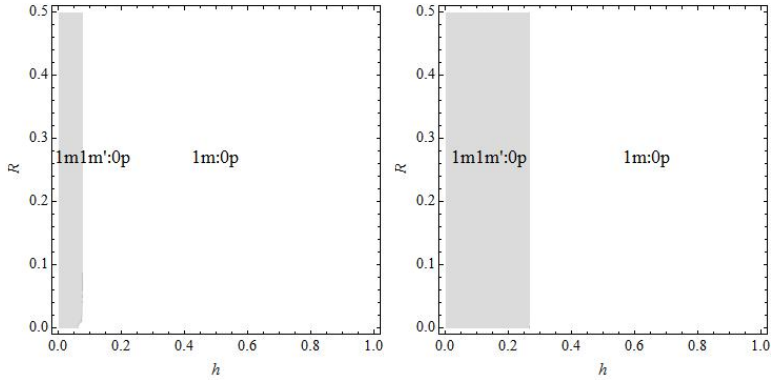

Figure 2.5 – Résultat de simulation pour le modèle Perisp1. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$

Perisp2 La transition $1m : 1p \rightarrow mm' : 0p$ (modèle Perisp2) correspond à la suppression d'une contribution paternelle et à l'ajout d'une contribution maternelle. Dans ce modèle, le résultat des simulations est différent selon que le contrôle du modificateur est mâle, femelle, ou biparental. L'état $mm' : 0p$ évolue dans des conditions de dominance plus larges si le contrôle du modificateur est paternel que si le contrôle du modificateur est maternel. On observe un effet important de s sur les conditions de dominance dans lesquelles l'état $mm' : 0p$ évolue (Figure 2.6).

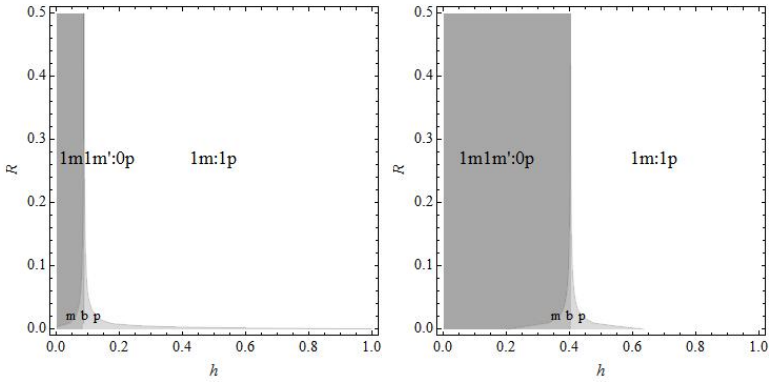

Figure 2.6 – Résultat de simulations pour le modèle Perisp2. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$

Perisp3 La transition $2m : 1p \rightarrow 1m1m' : 0p$ (modèle Perisp3) correspond à l'ajout d'une contribution m' et la suppression d'une contribution maternelle et de la contribution paternelle. Dans ce modèle, nous ne détectons pas de différence selon que le contrôle du modificateur est mâle, femelle, ou biparental, et $1m1m' : 0p$ semble évoluer lorsque $h < h_m$. Contrairement aux résultats des autres simulations, on n'observe pas d'effet de s sur la condition limite quand r est fort (Figure 2.7).

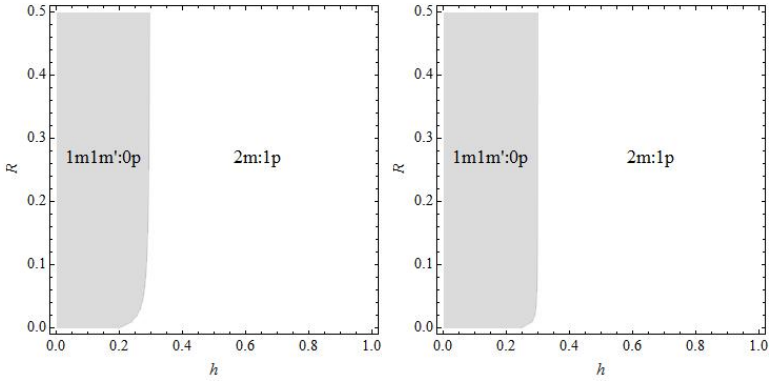

Figure 2.7 – Résultat de simulations pour le modèle Perisp3. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$, $h_1 = 0.4$, $h_2 = 0.2$

Polysp1 La transition $1m : 0p \rightarrow mm' : 1p$ (modèle Polysp1) correspond à l'ajout d'une contribution paternelle et d'une contribution m'. Dans ce modèle, l'état $1m1m' : 1p$ évolue lorsque h_m est faible, que le contrôle soit biparental, mâle, ou femelle (Figure 2.8).

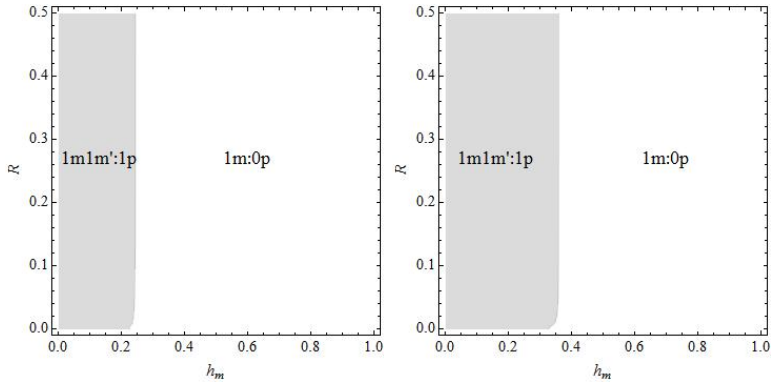

Figure 2.8 – Résultat de simulations pour le modèle Polysp1. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$, $h_1 = 0.4$, $h2 = 0.2$.

Polysp2 Comme la transition $1m : 0p \leftrightarrow mm' : 0p$ (modèle Perisp1), la transition $1m : 1p \leftrightarrow mm' : 1p$ (modèle Polysp2) correspond au simple ajout d'une contribution maternelle. Dans ce modèle, le résultat des simulations est différent selon que le contrôle du modificateur est mâle, femelle, ou biparental. Là encore, on observe un effet important de s sur les conditions de dominance dans lesquelles l'état $mm' : 1p$ évolue (Figure 2.9). Par ailleurs, l'état $1m1m' : 1p$ évolue lorsque h n'est pas trop faible, ce qui suggère que comme dans les autres modèles, il faut que l'exposition moyenne des allèles déltétères dans le tissu qui est résident soit un plus élevée que celle dans dans le tissu dont on étudie l'évolution pour que ce dernier évolue.

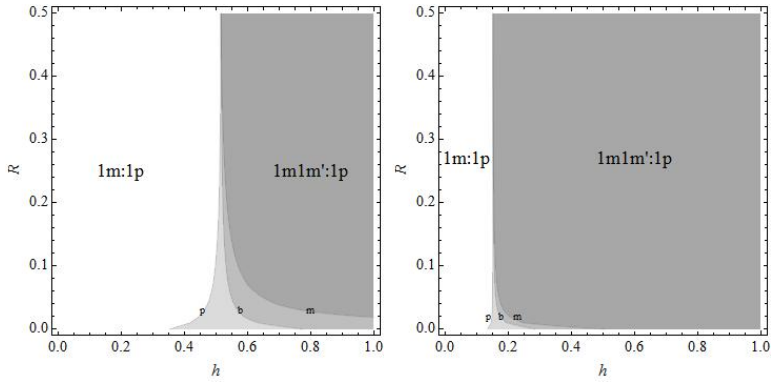

Figure 2.9 – Résultat de simulations pour le modèle Polysp2. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$, $h_1 = 0.4$, $h2 = 0.2$.

Polysp3 La transition $2m : 1p \rightarrow 1m1m' : 1p$ (modèle Polysp3) correspond à l'ajout d'une contribution m' et la suppression d'une contribution maternelle. Dans ce modèle,

le résultat des simulations est différent selon que le contrôle du modificateur est mâle, femelle, ou biparental, mais la différence est infime. L'état $1m1m' : 1p$ évolue pour des mutations délétères récessives dans un tissu triploïde.

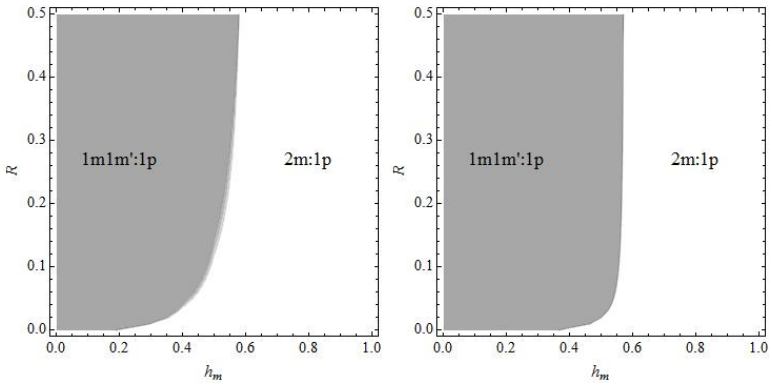

Figure 2.10 – Résultat de simulations pour le modèle Polysp3. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$, $h_1 = 0.4$, $h2 = 0.2$.

PP La transition $mm' : 0p \rightarrow mm' : 1p$ (modèle PP) correspond au simple ajout d'une contribution paternelle. Dans ce modèle, nous ne détectons pas de différence selon que le contrôle du modificateur est mâle, femelle, ou biparental. Le périsperme évolue lorque $h < h_m$, et la polysporie lorsque $h_m < h$ si la sélection est faible. Cependant, là encore, on observe un effet de s sur la condition limite (Figure 2.11).

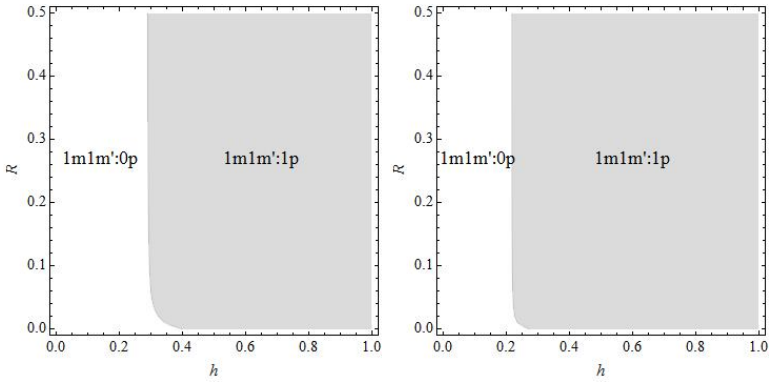

Figure 2.11 – Résultat de simulations pour le modèle PP. A gauche : $s = 10^{-1}$. A droite : $s = 10^{-2}$. Par ailleurs, $h_1 = 0.2$, $h_2 = 0.4$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$

2.6.3 Discussion

Les résultats de ces simulations indiquent que le périsperme peut évoluer pour des valeurs de dominance récessives dans un tissu diploïde, dans tous les modèles où il est impliqué. Les conditions dans lesquelles le périsperme évolue sont cependant d'autant

plus larges que les mutations délétères ont peu d'effet (s faible) quand le périsperme évolue face à des tissus nourricier $1m : 0p$ ou $1m : 1p$. Face à un tissu nourricier $2m : 1p$, le périsperme semble évoluer pour le critère $h_m < h$, qui ne dépend pas de s. Les conditions de l'évolution du périsperme sont plausibles en termes de dominance, puisque les mutations délétères sont généralement récessives. Cependant, les valeurs de s sont probablement beaucoup plus variables, d'un locus à l'autre, que les valeurs de dominance. De ce fait, il parait nécessaire d'intégrer les effets d'hétérozygotie à l'échelle du génome, en fonction de la distribution des valeurs de s, pour faire une prédiction plus précise.

La polysporie évolue pour de plus ou moins faible valeurs de h_m. Là encore, du fait de l'effet de s, il parait nécessaire d'intégrer les effets d'hétérozygotie à l'échelle du génome, en fonction de la distribution des valeurs de s, pour faire une prédiction plus précise.

Dans ces modèles, contrairement aux précédents, les conditions de l'évolution des tissus lorsque $r = 1/2$ ne dépendent pas que de la dominance, ils dépendent également beaucoup de s, la comparaison des valeurs de E_{moy} ne donne donc pas les conditions d'évolution lorsque $r = 1/2$. Une des perspectives de ce travail est de comparer les valeurs sélectives moyennes de populations dont les tissus nourriciers sont $1m : 0p$, $1m : 1p$, $2m : 1p$, $mm' : 0p$ ou $mm' : 1p$, ce qui devrait permettre d'obtenir ces conditions.

2.7 Hétérozygotie et évolution d'autres traits

Notre travail théorique nous a permis de montrer que l'hétérozygotie influe l'évolution de la double fécondation, du doublement des contributions maternelles, de la polysporie et probablement de l'empreinte parentale et de l'évolution de tissus nourriciers alternatifs tels que le périsperme : lorsque l'hétérozygotie augmente, il y a un bénéfice lié au masquage des mutations délétères, qui est couplé à un coût lié à l'accumulation des ces mutations délétères.

Les effets d'hétérozygotie pourraient affecter le mode d'expression génétique dans d'autres cas de figure que ceux que nous avons étudiés. Par ailleurs, si une pression de sélection qui n'est pas liée à l'hétérozygotie met en jeu l'évolution de l'expression des contributions génétiques dans un individu (pas forcément dans le tissu nourricier), tel qu'un conflit sur l'allocation des ressources, ou autre, on s'attend à ce que les effets liés aux changements d'hétérozygotie soient présent, et limitent ou facilitent l'évolution d'un trait.

Dans cette section, nous nous intéressons à trois cas de figure où les effets liés à l'hétérozygotie pourraient jouer un rôle : les effets maternels, la régulation génique, et l'expression mono-allélique aléatoire.

2.7.1 Effets maternels

Nous avons vu que l'empreinte parentale générait des effets parentaux par le biais d'une méthylation différentielle des copies maternelles et paternelles. Mais le contrôle du développement des embryons d'une génération par la précédente (par exemple, par la transmission de transcrits) présente les mêmes effets que l'empreinte parentale : l'expression d'un allèle est parent-spécifique. Cela génère des effets maternels, c'est-à-dire que le phénotype de l'embryon ou de son tissu nourricier dépend du génotype maternel.

Il existe deux types d'effets maternels : les effets maternels sporophytiques, et les

effets maternels gamétophytiques. Comme leurs dénominations l'indiquent, ils sont respectivement liés à l'expression de gènes du sporophyte et du gamétophyte. Les gènes dont l'expression est sporophytique conduisent à une expression $1m1m' : 0p$, alors que les gènes qui ont une expression gamétophytique conduisent à une expression $1m : 0p$.

D'après notre travail, il devrait exister des pressions de sélection qui favorisent ou non le contrôle maternel du développement, selon que les mutations délétères sont plutôt récessives ou non, et selon le taux de recombinaison dans le génome.

Gènes exprimés par le sporophyte : Il existe de nombreux gènes impliqués dans le développement de la graine et exprimés par le sporophyte. Le sporophyte femelle développe des tissus spécialisés dans la reproduction, ceux de l'ovule. A maturité, les téguments de l'ovule forment l'enveloppe de la graine. Par ce biais, le sporophyte maternel contrôle en partie le transfert de nutriments à la graine (comme en attestent les mutants *seg1*, *seg3*, *seg6*, *seg7* initialement identifiés chez l'orge, mais aussi les mutants *reduced grain filling-1* et *Miniature1* chez le maïs Chaudhury & Berger 2001). Le sporophyte maternel contrôle également la taille de la graine via la capacité d'élongation des téguments (Fitzgerald *et al.* 2008; Garcia *et al.* 2005). Ce contrôle s'effectue par le biais du gène *MET1* : un sporophyte femelle pour lequel le gène est non fonctionnel produit de plus gros grains (Fitzgerald *et al.* 2008). On peut constater que le contrôle du sporophyte concerne essentiellement des traits liés à l'accumulation des ressources, et pas, par exemple, à la différentiation de l'albumen. Pour expliquer ce patron, on peut imaginer qu'il existe une pression de sélection externe (par exemple liée au conflit sur l'allocation des ressources), qui joue en faveur de ce contrôle maternel sur l'allocation, mais que les autres traits ne sont pas sous le contrôle du sporophyte car cela représente un coût en termes d'hétérozygotie.

Gènes exprimés par le gamétophyte : Les transcrits du gamétophyte femelle, contenus dans le cytoplasme des gamètes, jouent un rôle après fécondation. Chez les animaux, le développement embryonnaire précoce est entièrement assuré par les ARN maternels. Ceux-ci se dégradent progressivement pendant différents stades du développement embryonnaire, dont le stade blastula et le stade gastrula. Le génome embryonnaire est initialement inactif. Il ne devient actif qu'à partir de la phase MZT (*Maternal to Zygotic Transition*), quand les gènes du zygote commencent à être transcrits (Schier 2007). Il est probable qu'il en soit de même chez les angiospermes (Vielle-Calzada *et al.* 2000; Pillot *et al.* 2010; Grimanelli *et al.* 2005). Les gènes du gamétophyte femelle semblent également contrôler le développement précoce de l'albumen (Vielle-Calzada *et al.* 2000; Pillot *et al.* 2010; Grimanelli *et al.* 2005), mais le mécanisme sous-jacent (transmission de transcrits ou bien transcription *de novo* restreinte au génome maternel) est sujet à débat. Si la double fécondation a bien évolué parce qu'elle augmente l'hétérozygotie, et que les conditions dans lesquelles la double fécondation a évolué se sont maintenues, la sélection qui résulte des effets d'hétérozygotie devrait conduire à une durée de latence la plus courte possible, et favoriser l'expression biallélique.

2.7.2 Régulation génique

Nous avons montré que lorsqu'un gène est régulé en *cis*, une baisse d'expression est plus couteuse pour le régulateur que lorsque le gène est régulé en *trans*.

Dans le génome, il existe de nombreux gènes dont la fonction est d'augmenter ou de diminuer l'expression d'autres gènes : les régulateurs. D'après nos résultats, on pourrait imaginer qu'il soit moins couteux qu'un inhibiteur exerce son contrôle en *trans*, et parallèlement, plus avantageux qu'un activateur exerce son contrôle en *cis*.

Si cela est exact, nous prédisons qu'un régulateur d'expression devrait contrôler l'allèle cible en *trans* plutôt qu'en *cis* si c'est un inhibiteur, et en *cis* plutôt qu'en *trans* si c'est un activateur.

2.7.3 Expression mono-allélique aléatoire

Comme nous l'avons précisé encadré 1.2 (en chapitre d'introduction), il existe des cas de figure où une seule des copies est exprimée à un locus donné, mais l'identité de la copie inactivée n'est pas pré-déterminée : l'inactivation est aléatoire. L'inactivation aléatoire peut être locus-spécifique, ou concerner des chromosomes entiers.

Prenons par exemple le cas de l'extinction du chromosome X chez les mammifères femelle. Un des X produit des ARN à un locus spécifique, le locus Xist. Ces ARN inhibent l'expression du chromosome, en *cis*. Il y aurait, en amont de l'expression monoallélique de Xist, un facteur autosomal, présent en quantité limitante, qui, au hasard, méthyle le promoteur du gène Xist sur l'un des chromosomes, et ce faisant l'éteint (Chaumeil *et al.* 2006; Ng *et al.* 2007).

Imaginons une mutation sur un chromosome donné, qui augmente, en *cis*, la probabilité que le facteur de blocage se fixe sur le gène Xist de ce chromosome. Ce chromosome a donc une probabilité plus élevée d'être celui qui reste actif. De ce fait, les chromosomes porteurs de cette mutation sont mieux purgés que les autres, ce qui comme nos modèles l'ont montré, peut constituer un avantage si les mutations délétères sont récessives et que la recombinaison est faible. On pourrait donc s'attendre à ce que la fréquence des chromosomes porteurs de cette mutation augmente en fréquence, et de ce fait, à ce que l'inactivation aléatoire ne respecte plus le ratio de 50/50 dans les individus où ils sont présents. La fréquence d'inactivation normale serais rétablie lors de la fixation de ces distorteurs d'expression dans la population, de la même manière que la ségrégation revient à 50/50 lorsqu'un distorteur de ségrégation se fixe dans une population.

Nous faisons donc la prédiction qu'on devrait trouver des distorteurs d'expression monoallélique aléatoire, c'est-à-dire des allèles modificateurs qui diminuent la probabilité, lorsque l'expression d'un locus ou chromosome est monoallélique, que le/les allèles éteints soient ceux qui sont présents sur le chromosome qui les porte.

2.7.4 Discussion

Le conflit sur les niveaux d'expression est donc susceptible d'affecter non seulement l'évolution de la double fécondation, du doublement des contributions maternelles, et de l'empreinte parentale, mais également l'évolution de la transmission de transcrits d'une génération à l'autre, ou plus généralement des effets maternels, de l'expression monoallélique aléatoire, et de la régulation génique.

Des modèles spécifiques sont nécessaires pour valider ces prédictions, car il est possible que d'autres facteurs entrent en jeu dans ces cas de figure. Cependant, si ces prédictions sont validées, elles pourraient participer à une meilleure compréhension de la manière dont la régulation et l'expression des génomes évolue.

2.8 Discussion générale et perspectives

La double fécondation, l'évolution du nombre de noyaux polaires (contributions maternelles), l'empreinte parentale et la polysporie modifient les relations de dominance dans l'albumen. Ces modifications sont toutes parent-spécifique. La principale différence entre l'empreinte parentale et les autres modifications est qu'elle est locale.

D'après nos résultats, la double fécondation ($1m : 0p \rightarrow 1m : 1p$) devrait évoluer et se maintenir très facilement, qu'une espèce soit autogame ou allogame. Le doublement des contributions maternelles ($1m : 1p \rightarrow 2m : 1p$), le périsperme, et la polysporie peuvent évoluer dans des conditions plausibles, mais probablement pas toujours remplies. Enfin, l'empreinte parentale (que l'extinction soit paternelle ou maternelle) évolue dans des conditions qui apparaissent peu réalisables.

Nos modèles ont donc généré de nombreuses prédictions en ce qui concerne la théorie de l'hétérozygotie. D'autres modèles mériteraient d'être étudiés :

Évolution corrélée de l'apomixie et de la polysporie : Carman (1997) montre que la polysporie est corrélée à l'apomixie. Dans le cas de l'apomixie, les embryons peuvent être issus de cellules du sporophyte, et porter des contributions m ou m', quelque soient les contributions reçues par le gamétophyte. On pourrait imaginer que la polysporie est un moyen de mieux exposer les contributions portées par l'embryon apomictique. Un modèle spécifique mérite donc d'être étudié pour mettre en évidence l'effet de l'apomixie sur l'évolution de la polysporie.

Phases de sélection haploïde et diploïde : Certains gènes sont spécifiquement exprimés dans l'albumen, quand d'autres ont des fonctions importantes à tous les stades et sont donc également exprimés dans le gamétophyte ou le sporophyte. Les gènes qui ne sont pas spécifiquement exprimés dans l'albumen subissent une phase de sélection haploïde et/ou diploïde en plus de la phase de sélection intra-albumen. Cela devrait changer l'équilibre mutation-sélection dans la population, donc probablement également les conditions de l'évolution des tissus nourriciers. L'étude de ces effets et de la proportion des gènes exprimés spécifiquement dans l'albumen est donc une des perspectives importantes de l'étude de l'évolution des contributions parentales dans le tissu nourricier.

Superdominance : Dans le modèle d'imprinting de Spencer & Williams (1997), les effets de superdominance rendent quasi-impossible l'évolution de l'empreinte. On s'attend donc à ce que la superdominance favorise très fortement l'évolution de la double fécondation.

Hétérozygotie et réallocation des ressources : Les théories de l'hétérozygotie et du conflit sur l'allocation pourraient être complémentaires. En effet, on peut imaginer que l'hétérozygotie soit le mécanisme par lequel la contribution paternelle augmente l'allocation des ressources à la graine dans le cadre de la théorie du conflit sur l'allocation. Cette hypothèse, proposée par Willson & Burley (1983), n'a pas fait l'objet d'un modèle malgré sa simplicité. Pour réaliser ce modèle, il suffirait d'ajouter de la réallocation des ressources dans le modèle DF : lorsqu'un tissu nourricier contient des mutations délétères, la graine reçoit moins de ressource et ces ressources sont disponibles pour d'autres graines, dont le tissu nourricier n'est pas porteur de mutations délétères. Pour réaliser ce modèle, il suffit de moduler le taux d'acquisition des ressources en fonction du génotype du tissu nourricier. Les prédictions d'un tel modèles devraient alimenter les

discussions concernant les théories de l'hétérozygotie et du conflit sur l'allocation : comme on l'a signalé en chapitre d'introduction, aucun modèle n'a été consacré à l'évolution de la double fécondation dans le cadre de la théorie du conflit sur l'allocation.

Chapitre 3

Analyses comparatives

La double fécondation, le doublement des contributions maternelle, la polysporie, le stockage dans l'embryon ou le périsperme sont des traits qui ont évolué sur des échelles de temps longues. Les patrons associés à l'évolution de ces traits ont pourtant été peu étudiés. D'après Carman (1997), la polysporie est corrélée à l'apomixie et la polyembryonie (production de plusieurs embryons pour un même ovule). D'après Uma Shaanker *et al.* (1996), l'albumen persiste en tant que tissu de stockage davantage chez les espèces multiovulées que chez les espèces uniovulées. Enfin d'après Forbis *et al.* (2002), la proportion de l'embryon dans la graine a maturité à tendance à augmenter au fil de la phylogénie des plantes à graine et est corrélée à la durée de la dormance.

Nous avons souhaité explorer les données disponibles pour mettre à jour d'autres patrons. Pour réaliser cet objectif, nous avons mis en place une base de donnée contenant des informations portant sur l'embryologie, les tissus nourricier de la graine, les systèmes de reproduction, le niveau de compétition entre embryons, les niveaux de dépression de consanguinité, la ploïdie et quelques autres traits d'histoire de vie des angiospermes (mode de dispersion des graines, forme de vie, etc.). Plusieurs références bibliographiques et bases de données existantes ont été utilisée pour construire cette base de données (Maheshwari 1963; Gachet *et al.* 2005; Spichiger *et al.* 2000; Davis 1966; Klotz *et al.* 2002; Carman 1997; Uma Shaanker *et al.* 1988). Nous avons notamment passé au crible Maheshwari (1963) et Davis (1966) pour les données concernant l'embryologie et le développement de la graine, et Spichiger *et al.* (2000) pour les données concernant l'état albuminé ou exalbuminé de la graine. Nous avons également élaboré des prédictions théoriques, relatives aux théories du conflit de l'allocation des ressources, et à la théorie de l'hétérozygotie. Nous présentons d'abord rapidement les prédictions et les variables auxquelles nous nous sommes intéressées, la base de données, puis deux exemples d'analyse.

3.1 Choix des traits et prédictions

La graine présente des différences d'une espèce à l'autre. Dans une première sous-section, nous listons les traits du tissus nourriciers qui sont variables et présentent donc un intérêt pour une analyse comparative. Ces traits ont probablement évolué dans des conditions précises. Dans une seconde section, nous listons quelles prédictions peuvent être faîtes en ce qui concerne ces conditions, et nous en déduisons les données qui seront nécessaire pour tester les prédictions. Enfin, l'évolution des traits du tissu nourricier a

aussi des conséquences. Dans une troisième section, nous listons les variables et prédictions relatives aux conséquences de l'évolution des traits du tissu nourricier auxquels on s'intéresse.

3.1.1 Traits du tissu nourricier

En plus de l'évolution de la double fécondation, des contributions maternelles, et de l'empreinte parentale (chez les angiospermes mais aussi chez les mammifères), le conflit sur l'allocation des ressources est supposé être responsable de l'évolution de l'hypostase (tissu lignifié au niveau du chalaze, c'est-à-dire au niveau de la zone où les ressources sont transférées à l'ovule, Westoby & Rice 1982), des haustoria (des structures embryonnaires qui envahissent le tissu sporophytique, Queller 1983; Bhatnagar & Kallarackal 1980) et du "strike" (tendance des noyaux polaires à devenir non-fonctionnels dans la cellule centrale, Haig 1986).

Nous avons décrit la variabilité de certains de ces traits en chapitre d'introduction. Ils ne présentent pas tous le même degré de variabilité. Notamment, la double fécondation n'est apparue qu'une fois pour exercer la fonction qu'elle exerce chez les angiospermes (comme on l'a précisé en chapitre d'introduction, la double fécondation chez les gymnospermes est rare et n'est pas équivalente d'un point de vue fonctionnel). Il n'y a que trois familles pour lesquelles soit il n'y a plus de double fécondation, soit l'albumen n'exerce aucune fonction. Il s'agit des Orchidées, Podostemacées, et Trapacées. Chez les Orchidées, aucune ressource n'est allouée aux graines qui sont nourries par les mycorhizes. Ainsi, quasiment tous les réplicas dont nous disposons pour étudier les conditions de présence et d'absence de double fécondation sont groupés dans la phylogénie (présence chez les angiospermes, absence chez les gymnospermes). Cela représente une contrainte majeure pour étudier l'évolution de ce trait. Les données que nous avons donc intégrées dans la base de donnée concerne donc seulement l'ensemble des autres traits énumérés ci-dessus.

3.1.2 Conditions de l'évolution des tissus nourriciers

D'après la théorie du conflit, on s'attend à ce que les traits énumérés section 3.1.1 évoluent plus souvent lorsque la compétition entre les embryons est forte et que les embryons sont peu apparentés (voir chapitre d'introduction).

D'après plusieurs auteurs Arathi *et al.* (1999, 1996); Bawa *et al.* (1989); Ganeshaiah & Uma Shaanker (1988); Krishnamurthy *et al.* (1997); Mohana *et al.* (2001); Uma Shaanker *et al.* (1996), le taux d'avortement (ou nombre de graines produites sur nombre d'ovules initialement présents) est un indice de la compétition entre embryons. Des données étant disponibles, nous les avons intégrées à notre base de données.

L'apparentement entre les embryons est modulé par le système de reproduction. La graine peut se reproduire sexuellement ou non, et lorsque la reproduction est sexuée, plusieurs facteurs peuvent faire varier le taux de multipaternité (Fig. 3.1). Nous avons donc collecté pour notre base de données des données relatives au système de reproduction et à ces différents facteurs.

Le type de tissu nourricier pourrait ne dépendre ni d'un conflit sur l'allocation des ressources, ni de l'hétérozygotie. Il apparait intéressant d'étudier la corrélation entre le type de tissu nourricier et son contenu (lipides, protéines, ou sucres), car il est probable que le tissu nourricier évolue en réponse à des contraintes liées à la nutrition, en particulier

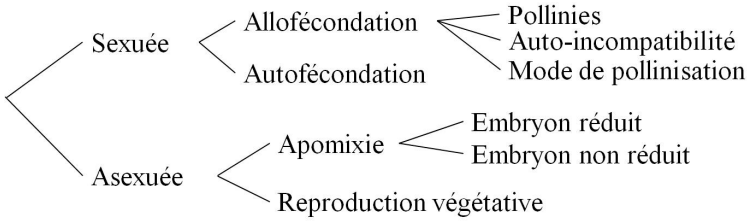

Figure 3.1 – Systèmes de reproduction

en ce qui concerne les transitions vers un stockage des ressources dans les cotylédons ou le périsperme.

Enfin, il est possible que l'hétérozygotie, donc le type de tissu nourricier le plus adapté, soit différent selon que la plante est autotetraploïde ou di- ou allotétraploïde.

3.1.3 Conséquences de l'évolution du tissu nourricier

Le type de tissu nourricier pourrait avoir des conséquences sur au moins quatre types de données : la dépression de consanguinité, la taille des graines, la variabilité dans la taille de la graine, et la prépondérance des effets mâles sur le déterminisme de la taille de la graine.

Nos modèles (chapitre 2) ont montré que la fréquence d'équilibre des mutations délétères dépend du type de tissu nourricier. Il est donc probable que la dépression de consanguinité dépende du type de tissu nourricier. On sait par ailleurs qu'un régime de consanguin augmente la purge des allèles délétères (Lande & Schemske 1985), ce qui implique de prendre en compte le régime de reproduction comme co-variable dans ce test.

La théorie du conflit d'allocation prédit que l'augmentation relative du nombre de contributions maternelles par rapport au nombre de contributions paternelles diminue le contrôle paternel sur la taille des graines. On s'attend donc à trouver des effets paternels sur le phénotype de la graine de moins en moins importants dans le sens tissu $1m : 1p \rightarrow$ tissu $2m : 1p \rightarrow$ tissu polysporique.

La théorie du conflit prédit que le tissu $2m : 1p$ correspond à un contrôle maternel plus important de la taille de la graine, et que l'optimum de taille de graine maternel est plus petit que l'optimum paternel. Toutes choses étant égales par ailleurs, on s'attend donc à ce que les graines soient plus petites lorsque le tissu nourricier est $2m : 1p$ que lorsqu'il est $1m : 1p$.

La théorie du conflit prédit également que si les graines sont en compétition, un plus grand contrôle paternel correspond à une plus grande variabilité de tailles de graines. Toutes choses étant égales par ailleurs, on s'attend donc a ce que la taille des graines soit plus variable lorsque le tissu nourricier est $1m : 1p$ que lorsqu'il est $2m : 1p$ ou $1m : 0p$.

Nous avons donc intégré des données relatives à la dépression de consanguinité, la taille des graines, et sa variabilité dans la base de données. Nous avons collecté des informations sur la prépondérance des effets mâles sur le déterminisme de la taille de la graine, mais elles étaient trop disparates pour être intégrées à la base. Elles sont données Table 3.4.

3.2 Base de données

L'objectif de l'analyse de données bibliographique est d'étudier les corrélations entre les caractéristiques du tissu nourricier et différent traits d'histoire de vie. Pour cela, il était nécessaire de compiler des données à la fois concernant les tissus nourriciers et les traits d'histoires de vie.

Les données relatives aux tissu nourriciers ont été obtenues à partir de *Recent advances in the embryology of angiosperms* (Maheshwari 1963), *Systematic embryology of the angiosperms* (Davis 1966), Botanique systématique des plantes à fleur (Spichiger *et al.* 2000) et un article (Carman 1997). Les données du Maheshwari (1963) se présentent sous la forme d'une description générale du développement de la graine (chaque chapitre ou section se concentre sur le développement d'une des parties de la graine). Des indications spécifiques sont données seulement pour les espèces sortant du patron général. Nous avons récolté les informations portant sur le type de tissu nourricier principal (périsperme, cotylédons, albumen), la présence d'un hypostase, le type de gamétophyte (duquel il est possible de déduire le niveau de sporie du gamétophyte), le nombre de noyaux polaires, le nombre de noyaux formant le gamétophyte femelle, le nombre de mitoses impliquées dans la formation du gamétophyte femelle, le type de développement de l'albumen (hélobial, syncitial, cellulaire, cf. encadré 3.1), l'occurrence d'apomixie, et lorsque la plante était apomictique, le type d'apomixie. Les données de Davis (1966) sont organisées par famille. Pour chaque famille est donnée une description verbale allant des organes reproducteurs au développement de la graine, en passant par la fécondation. Des informations supplémentaires sont données pour les genres faisant exception au sein des familles. Nous avons tiré de cet ouvrage des informations sur le nombre de cellules dans le grain de pollen au moment où il est relaché, le type d'ovule (ana, ortho ou campylotrope [1]), le nombre de téguments (uni ou bi tégumenté), le type de nucelle (tenui ou crassinucellé [2]), la présence d'un endothelium, l'occurrence de *strike*, le type de gamétophyte femelle (Polygonum, Oenothera, etc.), le nombre de noyaux formant le gamétophyte femelle, le nombre de noyaux polaires, l'absence de double fécondation, le nombre de synergides, le nombre d'antipodes, le type de développement de l'albumen, la présence de haustoria, le mode de développement de l'embryon (onagrad, asterad, solanad, etc. [3]), l'occurrence d'apomixie, le type d'apomixie, l'occurrence de polyembryonie, et la présence d'amidon dans la graine. Spichiger *et al.* (2000) est une flore botanique, les données y sont organisées par famille, et nous avons seulement récolté l'information portant sur le type de graine (albuminée ou exalbuminée). A partir de Carman (1997), nous avons obtenu des informations relatives à la polysporie, l'apomixie, et la polyembryonie, allant au minimum au niveau du genre, parfois l'espèce.

Les données relatives aux traits d'histoires de vie ont été obtenues à partir de bases de données écologiques existantes : BASECO (Gachet *et al.* 2005) et BiolFlor (Klotz *et al.* 2002). Il s'agit de données relatives à la ploïdie, le système de reproduction, le mode de dispersion des graines, ou la forme de vie. En ce qui concerne le système de reproduction, nous avons également utilisé la base de données de C.G. Eckert, qui contient des informations relatives au taux d'allofécondation.

Nous avons récolté des données indicatrices du niveau de compétition entre embryons. Uma Shaanker *et al.* (1988) contenait des données sur le nombre de graines par fruit, les

1. Orthotrope : ovule droit, Anatrope : ovule renversé et Campylotrope : ovule courbé.

2. Tenuinucellé : le nucelle est formé au maximum d'une ou deux couches de cellules, Crassinucellé, le nucelle est épais, formé de plusieurs couches de cellules.

3. Ces dénominations correspondent à des descriptions précises qu'il n'est pas nécessaire de détailler ici.

Encadré 3.1 – Le développement de l'albumen

Chez un grand nombre d'espèces, l'albumen se développe en deux phases. Une phase syncitiale et une phase cellulaire. La transition entre ces phases s'effectue par une cytokinèse particulière appelée cellularisation (Berger 2003). Le mode de développement de l'albumen est important pour l'acquisition des ressources, notamment parce que le nombre de mitoses syncitiales détermine le nombre de cellules de l'albumen, donc la force du puit (Brocklehurst 1977; Jones *et al.* 1996). Le mode de cellularisation varie en fonction des espèces (Judd *et al.* 2001) :

- Développement nucléaire. Dans ce type de développement, présent chez la plupart des espèces, l'albumen ne cellularise que tardivement ou jamais (Judd *et al.* 2001). Par exemple, chez la noix de coco, l'albumen ne cellularise pas, il est liquide, c'est le lait de coco.
- Développement cellulaire. Les parois des cellules se forment dès le départ (Judd *et al.* 2001). Dans l'albumen cellularisé, différents tissus se différentient le long d'un axe antéro-postérieur qui est orienté du micropyle (site d'entrée du tube pollinique) au chalaze (où les tissus vasculaires maternels sont connectés aux téguments (Berger 2003).
- Développement hélobial. Caractéristique de la plupart des monocotylées. Le noyau primaire de l'albumen produit deux cellules. L'une se divise peu, alors que les cellules dérivées de l'autre produisent la plus grande partie de l'albumen (Judd *et al.* 2001).

Chez les céréales, la phase syncitiale correspond à phase de latence, pendant laquelle la structure de l'albumen se met en place, la phase cellulaire, correspond à une phase de maturation, pendant laquelle les ressources sont accumulées. L'albumen peut être ou non consommé par l'embryon avant la maturité de la graine, produisant une graine exalbuminée ou albuminée, respectivement.

Catégorie	Variable	Description	Valeurs	Nb d'espèces	Nb de genres	Nb de familles
	FgamType	Type de gamétophyte femelle	Oenothera/ Polygonum/ etc.	212	269	111
	FgamSpory	Sporie du gamétophyte femelle	Monosporic/ Bisporic/ Tetrasporic	399	556	169
Sac	NoyPnb	Nombre de noyaux polaires	-n-	71	102	58
Embryonnaire	Nucellus	Type de nucelle	Tenuinucellar/ Crassinucellar	69	134	68
Embryologie	Haustoria	Présence d'un haustoria	Yes/No	48	100	51
	Megasporo	Description de la megasporogénèse	-texte-	76	133	84
	NoyP_fusion	Moment de la fusion des noyaux polaires	Jamais/ Avantfec/ AprèsFec	118	220	125
	MainStorageTissue	Tissu nouricier principal de la graine mature	Albumen/ Cotyledons/ etc.	13	24	24
	Albumen2	Fonction de l'albumen	Storage/Transfert	7	13	13
	MST_Content	Contenu du tissu nouricier principal	Oil/Starch/etc.	7	13	13
Tissu	StarchGrain	Présence d'amidon dans la graine	Yes/No	29	65	42
nourricier	Autonomy	Si la graine se développe sans tissu nourricier	Mycotrophy/ Photosynthese/ etc.	10	14	12
	Alb_dvpt	Mode de développment de l'albumen	Nucellar/ Cellular/ Helobial	120	198	104
	Endothelium	Presence d'un endothelium	Yes/No	19	34	22
	SeedMass_out	Seed mass for outbreds	Valeurs	27	25	22
	SeedMass_in	Seed mass for inbreds	Valeurs	27	25	22
	ReproType	Mode de reproduction	Vegetative/ BySeed	3147	836	155
	VegMode	Mode de reproduction végétatif	Rhizome/ etc.	1267	498	115
	SeedProd	Mode de formation des graines	Apomict/ Sexual	3531	1200	224
	BreedingSyst	Système de reproduction	Inbred/ Outbred	2482	729	148
	ApomixisType	Type d'apomixie	Aposporous/ Diplosporous	461	188	59
Système	InbreedingCoeff	Coefficient de consanguinité (F)	-n-	98	92	52
de	SelfingRate	Taux d'autofécondation	-n-	21	12	8
reproduction	PolMode	Mode de polinisation	Insects/Wind/etc.	4259	921	161
	SI	Systèmes d'autoincompatibilité	Yes/No	2276	640	122
	SI_Mecas	Mécanismes du système d'autoincompatibilité	-	479	261	80
	Dichogamy	Dichogamie	Protandry/ Protogyny	1996	585	116
	Heterostyly	Heterostylie	Yes/No	39	16	11

Table 3.1 – Base de données, Partie 1. Certaines variables sont informées seulement pour la famille ou le genre, quand d'autres le sont pour l'espèce. -n- indique que les données sont numériques.

Catégorie	Variable	Description	Valeurs	Nb d'espèces	Nb de genres	Nb de familles
Compétition	Polyemb	Occurence de polyembryonie	Yes/No	353	524	162
	Polyemb_type	Type de polyembryonie	Sexual/Apomict	156	255	118
	NbSeedPerFruit	Nombre de graine par fruit	-n-	209	167	54
	SeedSet_out	Nombre de graine par plante en cas d'allofécondation	-n-	43	37	30
	SeedSet_in	Nombre de graine par plante en cas d'autofécondation	-n-	44	37	30
	Ratio_NbSeed_NbOv	Ratio du nombre de graines sur le nombre d'ovules	-n-	213	163	52
	FruitSet_out	Nombre de fruit par plante en cas d'allofécondation	-n-	28	24	27
	FruitSet_in	Nombre de fruit par plante en cas d'autofécondation	-n-	28	24	27
	FloweringLength	Longueur de la phase de floraison	-n- (jours)	3040	828	154
	LifeSpan	Durée de vie	Annual/Perennial	4326	987	177
	Fecundity_out	Fécondité d'un outbred	-n-	31	20	14
	Fecundity_in	Fécondité d'un inbred	-n-	31	20	14
Dépression de consanguinité	3stageIBD	Mesure cumulative de dépression de consanguinité	-n-	57	42	29
	FruitSet_RelPerf	DC sur le nombre de fruit	-n-	28	24	21
	SeedMass_RelPerf	DC sur la masse des graines	-n-	27	25	22
	SeedSet_RelPerf	DC sur le nombrede graines	-n-	54	44	33
	Germination_RelPerf	DC sur le taux de germintation	-n-	57	44	30
	Survival_RelPerf	DC sur la survie	-n-	45	38	27
	Fecundity_RelPerf	DC sur la fécondité	-n-	32	21	15
	VegSize_RelPerf	DC sur la taille du sporophyte	-n-	47	39	27
	EQLmoy	Equivalents lethaux, valeur moyenne	-n-	7	8	7
Divers	Ploidy	Niveau de ploïdie	Diploid/ Tetraploid	2525	790	152
	AncientPloidy	Ploïdie ancestrale supposée	-	2694	802	124
	EvolutionaryOrigin	Espèce domestiquées	Natural/ Domesticated	3147	836	155
	DispMode	Mode de dispersion des graines	Water/ Wind/ Birds/ etc.	1266	484	118
	DispUnit	Unité de dispersion des graines	Seed/ Fruit/ etc.	154	133	42
	LifeForm	Forme de vie	Woody/ Herbaceous	99	82	38
	LifeForm_detailed	Forme de vie détaillée	Phanerophyte/etc.	99	82	38
	PlantSize	Taille du sporophyte	-n-	1209	486	115
	Germination_out	Taux de germination d'un outbred	-n-	51	39	27
	Germination_in	Taux de germination d'un inbred	-n-	50	39	27
	Survival_out	Survie d'un outbred	-n-	40	33	25
	Survival_in	Survie d'un inbred	-n-	41	33	25
	VegSize_out	Taille du sporophyte outbred	-n-	38	35	26
	VegSize_in	Taille du sporophyte inbred	-n-	39	36	27
	POratio_max	Ratio du nombre de grains de pollen par nombre d'ovules, valeur maximum	-n-	55	24	12
	POratio_in	Ratio du nombre de grains de pollen par nombre d'ovules, valeur minimum	Valeurs	233	121	39

Table 3.2 – Base de données, Partie 2.

taux d'avortement (nombre de graine par ovule), le mode de dispersion des graines, et le mode de pollinisation. Wiens (1984) contenait des données sur le taux d'avortement, le nombre d'embryons par plante, et pour les espèces correspondantes, sur la forme de vie (annuelle ou pérenne, herbacée ou ligneuse) et le régime de reproduction (autofécondant ou allofécondant). Nous disposions d'une base de données (non publiée) portant sur les mesures de la dépression de consanguinité à différents stades. Pour les espèces dont nous disposions déjà, nous avons enfin interrogé la *Seed Information Database* de KewGarden sur la masse des graines, le mode de dispersion, et le contenu de la graine (pourcentages de graisses et de protéines).

Il existe un grand nombre d'homonymies et de synonymies en ce qui concerne les noms de taxons de plantes. Nous nous sommes particulièrement confrontés à ce problème avec les références anciennes (Maheshwari 1963; Davis 1966). C'est pourquoi nous avons mis en place une base taxonomique, en fusionnant deux bases de synonymie existantes (Complete Plant Checklist [4] et BDNFF [5]). Le principe des bases taxinomiques est d'associer à chaque nom de taxon une clef unique et une clef qui fait référence à l'espèce auquel ce nom réfère. Un même nom peut référer à plusieurs espèces (en cas d'homonymie), et plusieurs noms peuvent se référer à la même espèce (en cas de synonymie). Les clefs d'identifications n'étaient pas les mêmes dans les deux bases, nous avons donc fusionné les bases sur les noms taxinomiques et créé nos propres clefs pour que chaque ligne ait un identifiant unique.

Après avoir affecté au maximum de lignes de chaque jeu de donné un identifiant taxinomique (référant à une clef de notre base), nous avons compilé l'ensemble de ces données sous *Access*. L'ensemble des variables récoltées et le nombre de taxons représentés pour ces variables et identifiés par notre base taxonomique est donné Table 3.1.

Il était également nécessaire d'acquérir des données phylogénétiques pour étudier les corrélations entre les caractéristiques du tissu nourricier et différent traits d'histoire de vie en tenant compte de l'inertie phylogénétique. L'ensemble des données phylogénétiques que nous utilisons ont été obtenues par l'outil Phylomatic, qui permet de récupérer la phylogénie, sous formats standards, de familles d'angiospermes que l'on a choisies (http ://www.phylodiversity.net/phylomatic/phylomatic.html).

Les données et les questions que nous nous sommes posées sont nombreuses, mais nous n'avons pas eu le temps de toutes les explorer. Nous présentons donc ci-après deux analyses, l'une concerne la théorie de l'hétérozygotie, et l'autre concerne la théorie du conflit sur l'allocation des ressources.

3.3 Tissus nourriciers et dépression de consanguinité

Lors d'une transition d'un état haploïde vers un état diploïde dans l'évolution des cycles de vie, l'augmentation de l'hétérozygotie a pour conséquence d'augmenter l'accumulation des mutations délétères qui, masquées dans les hétérozygotes, sont moins purgées par la sélection. Nos modèles ont montré que les changements dans le nombre de contributions mâles et femelles dans les tissus nourriciers modifie également l'équilibre mutation-sélection (voir chapitre 2) .

Partant de ce constat, nous avons souhaité étudier la relation entre type de tissu nourricier et dépression de consanguinité. Pour cela, nous avons tout d'abord émis des

4. `http://plants.usda.gov/dl_all.html`
5. `http://bbock.free.fr/botanica/BDNFF.php`

Genotype	Fitness
A, AA, AAA, $AAAAA$	1
$AAAAa$	$1 - h_A s$
AAa	$1 - h_2 s$
$AAaaA$	$1 - h_B s$
Aa	$1 - hs$
$AAaaa$	$1 - h_C s$
Aaa	$1 - h_1 s$
$aaaaA$	$1 - h_D s$
a, aa, aaa, $aaaaa$	$1 - s$

Table 3.3 – Valeur sélective pour chaque génotype, en fonction de l'effet des allèles délétères (s), et de la dominance de l'allèle délétère a dans ces génotypes.

prédictions théoriques. Nous avons exprimé la dépression de consanguinité pour un locus donné en fonction de la fréquence des allèles délétères. Les modèles réalisés chapitre 2 nous permettent de prédire cette fréquence, qui dépend du taux d'apparition des mutations délétères (μ), de leurs dominances dans les tissus nourriciers auxquels on s'intéresse (Table 3.3), et de leur effet (s). Cela permet de prédire la dépression de consanguinité $\delta = 1 - W_o/W_s$ (où W_o est la valeur sélective moyenne de la population après une génération d'allofécondation, et W_s est la valeur sélective moyenne de la population après une génération d'autofécondation) en fonction de ces paramètres pour les différents types de constitutions génétiques possibles :

$$\delta^{1m0p} = 0 \tag{3.1}$$

$$\delta^{1m1p} = \frac{\mu(1-2h)}{2h} \tag{3.2}$$

$$\delta^{2m1p} = \frac{\mu(1-h_1-h_2)}{h_1+h_2} \tag{3.3}$$

$$\delta^{1m1m'1p} = \frac{\mu(h_1-2h_2)}{h_2} \tag{3.4}$$

$$\delta^{2m2m'1p} = \frac{2\mu(h_C-h_A-h_B)}{h_A+h_B} \tag{3.5}$$

Partant de ces expressions et en utilisant les hypothèses de la théorie métabolique de la dominance, cela nous permet de prédire que pour un taux de mutation génomique U, on s'attend à ce que :

$$\delta^{Gymnosperme} \approx 0 \tag{3.6}$$

$$\delta^{Oenothera} \approx U \tag{3.7}$$

$$\delta^{Polygonum} \approx 0.8U \tag{3.8}$$

$$\delta^{Adoxa} \approx 1.4U - 1.7U \tag{3.9}$$

$$\delta^{Allium} \approx 1.4U - 2U \tag{3.10}$$

$$\delta^{Penea} \approx 1.6U \tag{3.11}$$

Ces expressions permettent de prédire que la dépression de consanguinité des tissus nourriciers de la graine devrait être affectée par sa constitution génétique, dans la direction tissu nourricier des gymnospermes → tissu monosporique → tissu polysporique. Puisque l'embryon se nourrit de ce tissu nourricier, nous nous attendons par ailleurs à ce que ces effets de dépression de consanguinité se répercutent sur le développement du sporophyte, et ce d'autant plus que le stade du cycle de vie est précoce.

Nous utilisons ensuite des données pour nous intéresser à la relation entre type de tissu nourricier et dépression de consanguinité à différents stades du cycle de vie. Pour cela, nous compilons un jeu de données contenant des informations sur le type de tissu nourricier, des valeurs de dépression de consanguinité à plusieurs stades, et enfin sur le taux d'allofécondation, ce pour 27 espèces de plantes à graine. En ce qui concerne le type de tissu nourricier, les gymnospermes ont un tissu $1m : 0p$ et les angiospermes sont classées comme monosporique, bisporique ou tétrasporique selon l'espèce. En ce qui concerne les valeurs de dépression de consanguinité, nous disposons de valeurs compilées pour 4 stades : stade graine, germination, survie adulte, et fécondité. Pour chaque stade, nous construisons un modèle ANOVA dans lequel la dépression de consanguinité est exprimée en fonction du type de tissu nourricier et du taux d'allofécondation. Sans préjuger des valeurs de h, h_1, h_2, h_A, h_B, h_C, et h_D, nous faisons la prédiction que l'effet du type de tissu nourricier doit être plus marqué en début de cycle.

D'après ces modèles, le tissu nourricier a un effet significatif sur la dépression de consanguinité aux stades graine, germination et fécondité. Comparativement, le taux d'allofécondation n'a d'effet significatif qu'au stade germination (cependant, la plupart des espèces on un taux d'allofécondation supérieur à 0.7, nous avons donc peu de variabilité sur cet effet pour réellement le tester). D'après nos résultats, il existe donc bien un effet du type de tissu nourricier sur la dépression de consanguinité. Cependant, cet effet ne suit pas la relation croissante 'dépression d'un tissu nourricier des gymnospermes < tissu monosporique < tissu polysporique' prédite à partir de la théorie métabolique de la dominance (Fig 3.2, Annexe D).

Par ailleurs, si la présence d'effets aux stades graine et germination ainsi que l'absence d'effet au stade "survie" sont conformes à notre attendu, il est surprenant d'observer un effet au stade fécondité, qui est un stade tardif. Il est possible de supposer que la dépression de consanguinité sur le nombre de graines (qui est une des composantes de la dépression associée au stade fécondité) est plus dépendante du tissu nourricier que le stade survie parce que l'homozygotie du sporophyte est héritée par les tissus nourriciers polysporiques (i.e. porteurs à la fois de m et m'), et pas par les tissus nourriciers monosporiques.

Ces analyses sont encore en cours d'interprétation, mais ces résultats appuient l'idée qu'il est important de s'intéresser aux effets de l'hétérozygotie pour comprendre les pressions de sélection qui agissent sur l'évolution des tissus nourriciers dans les graines.

Un manuscrit correspondant aux résultats de cette analyse étant en cours de rédaction, voir celui-ci Annexe D pour les détails sur l'obtention des prédictions et sur l'analyse.

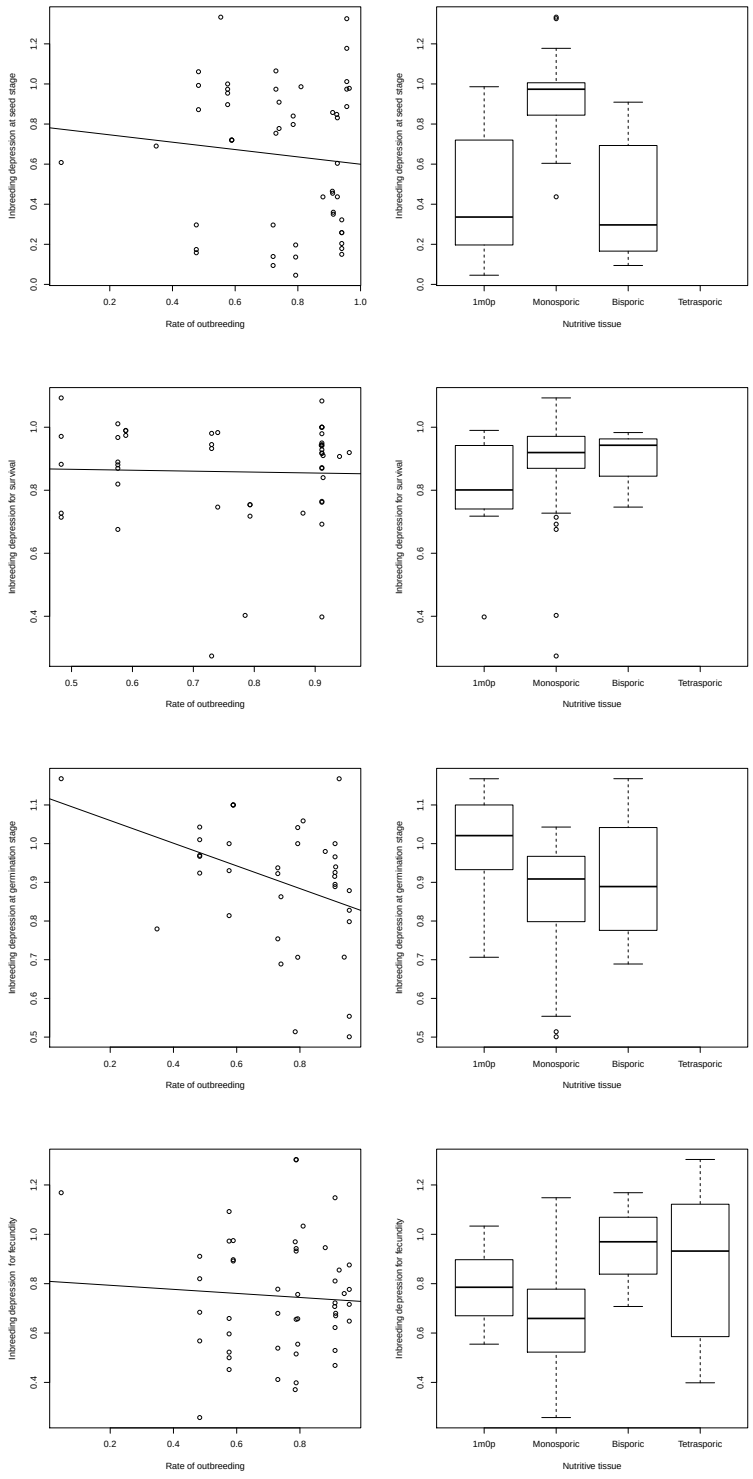

Figure 3.2 – Dépression de consanguinité en fonction du régime de reproduction et du type de tissu nourricier.

3.4 Tissus nourriciers et effets parentaux

La théorie du conflit sur l'allocation des ressources prédit que l'évolution des tissus nourricier change l'importance respective des contrôles maternels et paternels sur l'allocation des ressources. On s'attend par exemple à ce que, toutes choses étant égales par ailleurs, une graine dont le tissu nourricier est $1m : 1p$ présente un contrôle paternel plus important qu'une graine dont le tissu nourricier est $2m : 1p$. Par ailleurs, la théorie prédit que les indices du conflit devraient être d'autant plus marqués que la plante est allogame. En effet, chez les espèces autogames, le conflit sur l'allocation des ressources est relâché, il y a donc moins de pression de sélection pour faire disparaître les effets mâles. Si des effets paternels étaient observés plus souvent chez les autogames que les allogames, l'absence d'effet paternels chez les allogames pourrait être interprétée comme la résolution d'un conflit où la femelle aurait pris le contrôle. La théorie du conflit prédit que le tissu $2m : 1p$ correspond à un contrôle maternel plus important de la taille de la graine, et que l'optimum de taille de graine maternel est plus petit que l'optimum paternel. Toutes choses étant égales par ailleurs, on s'attend donc à ce que les graines soient plus grosses lorsque les effets paternels sont importants comparés aux effets maternels.

Pour vérifier ces prédictions, nous avons donc inventoriées les données disponibles concernant l'étude des effets paternels (Table 3.4), et croisé ces données avec les informations de notre base de données concernant les types de tissu nourricier, le régime de reproduction, et la taille de graine. Les données disponibles sont trop peu nombreuses et trop variables pour conclure, mais cette analyse permet de présenter les prédictions qui peuvent ou non être faîtes et le type de tests qui pourraient être menés lorsque d'avantages d'études portant sur les effets paternels seront disponibles.

Nous disposons de deux types d'analyse des effets paternels, qui sont 1) la décomposition de la variance liée aux effets males et femelles et 2) l'étude de l'effet du nombre de pères fécondants sur le coefficient de variation de la masse des graines ou sur la masse des graines. Dans le premier type d'analyse, les effets des différents males sont contrôlés, mais deux males ne sont jamais présents sur la même plante, si bien qu'un certain nombre d'effets 'plante mère' bruitent la mesure des effets males. Dans le second type d'analyse, les mâles sont mélangés sur les plantes mères, mais il n'est pas possible de contrôler la proportion de graines fécondée par chaque male ni de mesurer les masses de graine génotype par génotype. Une seule de ces analyses mélange les mâles sur une même plante tout en gardant l'information de paternité, celle de House *et al.* (2010).

Parmi l'ensemble de ces analyses de l'effet paternel, douze détectent un effet mâle sur la taille moyenne des graines, bien que pour deux d'entre elles cet effet soit indirect (lié a l'interaction mâle-femelle). Deux seulement ne détectent pas d'effet mâle sur la taille moyenne des graines. Quatre détectent un effet mâle sur le coefficient de variation de la taille des graines, et trois n'en détectent pas (Table 3.4).

Parmi les analyses impliquant une espèce pour laquelle une information sur le type de tissu nourricier est disponible (10 expériences), une seule ne détecte pas d'effet mâle, et le tissu nourricier de l'espèce de cette analyse est exalbuminé. Or, on s'attend à observer davantage d'effets mâles pour les tissus nourriciers exalbuminés qui sont $1m : 1p$ que pour les tissus albuminés qui sont très probablement $2m : 1p$. Neuf analyses détectent des effets mâles, et ce à la fois pour des tissus albuminés et des tissus exalbuminés. Il n'apparait donc pas de corrélation évidente entre la détection d'effets paternels dans le tissu nourricier et le type de tissu nourricier.

Nous disposons de données sur la taille des graines pour 15 analyses. La figure 3.3 représente la taille de graine en fonction du type d'effet détecté. La variance est trop im-

Famille	Espèce	Type d'expérience	Resultats	Détails	Ref	Breeding system	Tissu nourricier	Masse 100 graines
Polemoniaceae	*Ipomopsis aggregata*	1	i (MSW)	Pat ns, Mat 62%	Waser *et al.* (1995)			1,79g
Hydrophyllaceae	*Nemophila menziesii*	1	i (MSW)	Pat ns, Mat 20%	Byers *et al.* (1997)			2,8g
Asclepiadaceae	*Asclepias incarnata*	1	i* (MSW)	Pat ns, Mat 31.3-68.4%, Inter 7.3-46.5%	Lipow & Wyatt (1999)	Allogame ?		3,65g
Leguminosae-caesalpinioideae	*Chamaecrista fasciculata*	1	i* (SW)		Fenster (1991)			8,4g
Caryophyllaceae	*Lychnis flos-cuculi*	1	ii (SW)		Biere (1991)	Allogame	Exalbuminée ?	0.019g
Cruciferae	*Raphanus raphanistrum*	1	ii (SW)		Mazer *et al.* (1986)	Allogame	Exalbuminée ?	21,58g
Gramineae	*Anthoxanthum odoratum*	1	ii (SW)	Pat 2.6%, Mat 16,6%	Antonovics & Schmitt (1986)		Albuminée	0,53g
Polemoniaceae	*Phlox drummondii*	1	ii (SW)	Pat 1%, Mat 56%	Schwaegerle & Levin (1990)	Allogame		1,92g
Cruciferae	*Raphanus sativus*	1	ii (MSW)		Nakamura & Stanton (1989)	Allogame	Exalbuminée ?	19g
Campanulaceae	*Campanula americana*	1	ii (MSW)	Pat 9%, Mat 15%	Richardson & Stephenson (1991)	Allogame		0,17g
Cruciferae	textitArabidopsis thaliana	1	ii (MSW)		House *et al.* (2010)	Autogame		0,02g
Compositae	*Crepis tectorum*	1,2	ii (MSW), - (CV of MSW)		Andersson (1990)	Allogame	Exalbuminée	
Hydrophyllaceae	*Hydrophyllum appendiculatum*	1,2	0 (MSW), + (CV of MSW)		Wolfe (1995)	Allogame		
Cactaceae	*Pachycereus schottii*	2	- (CV of SW)		Holland *et al.* (2009)		Exalbuminée	2,2g
Nyctaginaceae	*Mirabilis jalapa*	2	- (CV of SW)		Niesenbaum (1999)	Allogame ?		57,1g
Rhamnaceae	*Rhamnus alpinus*	2	+ (CV of SW)		Banuelos & Obeso (2003)			
Leguminosae-papilionoideae	*Vigna unguiculata*	2	+ (CV of SW, MSW)		Uma Shaanker & Ganeshaiah (1997)	Autogame ?	Exalbuminée	103,6g
Ericaceae	*Vaccinium corymbosum* L.	2	+ (CV)		Vanderkloet & Tosh (1984)	Pas apomictique ?	Albuminée	0,44g
Cruciferae	*Cochlearia bavarica*	2	+ (SW)		Paschke *et al.* (2005)	Allogame	Exalbuminée ?	
Costaceae	*Costus allenii*	2	+ (MSW)		Schemske & Pautler (1984)	Allogame	Albuminée ?	

Table 3.4 – Effets mâles sur la taille des graines. Abréviations : SW taille des graines, MSW taille moyenne des graines, TSW masse totale des graines, CV coefficient de variation de la taille des graines. Expériences de type 1 : diallèles. i effet maternel, pas d'effet paternel, ii effet maternel, faible effet paternel, 0 pas de variance génétique additive, *effet paternel dans l'interaction male-femelle. Expériences de type 2 : mélanges de pollen. "+" détection d'un effet de la diversité des pollens, "-" pas d'effet détecté.

portante pour conclure, mais il n'apparait pas que les espèces pour lesquelles un effet mâle direct a été détecté aient des graines plus grosses. Il est connu que la taille de la graine varie principalement, d'une espèce à l'autre, en fonction de facteurs écologiques (Leishman *et al.* 2000; Turnbull *et al.* 1999; Moles *et al.* 2004, 2005). Une éventuelle analyse plus poussée de ce type de corrélation devra prendre en compte ces facteurs ou être corrigée par la phylogénie.

La plupart des analyses pour lesquelles nous disposons d'une information sur le régime de reproduction implique des espèces allogames. Les deux analyses correspondant aux deux espèces autogames détectent un effet paternel. Ce résultat est donc compatible avec la théorie du conflit sur l'allocation des ressources.

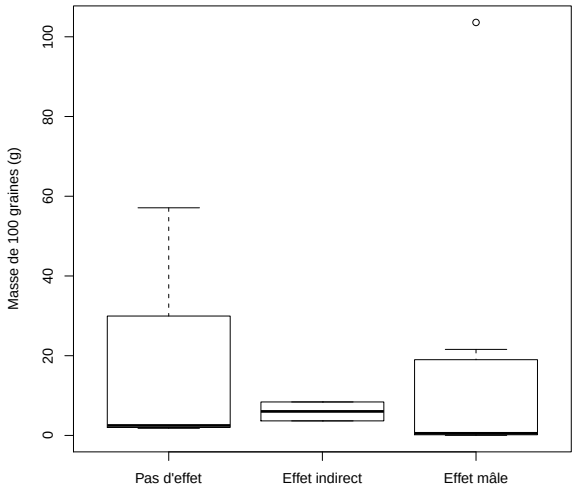

Figure 3.3 – Effets mâles sur la taille des graines.

Globalement, nous avons à la fois mis en évidence des éléments allant en faveur de la théorie du conflit sur l'allocation des ressources, et des éléments allant en défaveur de cette théorie. La corrélation entre effets paternels et type de tissu nourricier devrait être informative si un grand nombre de données étaient disponibles. La corrélation entre taille de graine et effet paternel pourrait être informative si les facteurs qui sont les plus importants dans le déterminisme de la taille des graines pour une espèce, c'est-à-dire les facteurs écologiques, sont pris en compte. Enfin, la corrélation entre effet paternel et le régime de reproduction (allogame ou autogame) pourrait devenir informatif si davantage de données étaient disponibles.

Chapitre 4

Conflits mâle-femelle et allocation des ressources chez Zea mais

La théorie du conflit sur l'allocation des ressources suggère que l'allocation des ressources aux graines est l'objet d'un conflit du fait de la compétition entre embryons pour des ressources qui sont limitées. L'objet de cette section est l'étude empirique des patrons d'allocation et de réallocation, ce qui permettra de déterminer (1) si les embryons sont effectivement en compétition pour les ressources et (2) quels sont les acteurs de cette compétition. Nous présentons deux analyses réalisées à partir du même matériel biologique : le maïs. Nous commençons par présenter les méthodes d'acquisition de ce matériel, puis présentons une analyse de l'allocation des ressources à l'échelle de l'épi, pour terminer sur une analyse de l'allocation des ressources à l'échelle du grain.

4.1 Introduction

Quelques études se sont intéressées à l'effet des mâles sur l'allocation des ressources aux graines. Nous en avons répertorié trois types. Une première catégorie de travaux s'intéresse à la décomposition de la variance de taille de graine liée aux effets mâles et femelles. Une seconde catégorie s'intéresse à l'étude de l'effet du nombre de pères fécondants sur le coefficient de variation de la masse des graines ou sur la masse des graines. Dans le premier type d'analyse, les effets des différents mâles sont contrôlés, mais deux mâles ne sont jamais présents sur la même plante. Dans le second type d'analyse, les mâles sont mélangés sur les plantes mères, mais il n'est pas possible de contrôler la proportion de graines fécondée par chaque mâle ni de mesurer les masses de graine génotype par génotype. Dans l'ensemble de ces analyses, il n'est donc pas possible d'étudier l'interaction entre les mâles. Une seule analyse mélange les mâles sur une même plante tout en gardant l'information de paternité, celle de House *et al.* (2010). Cependant, tous les mâles sont toujours présents sur toutes les plantes dans cette étude, si bien qu'il n'est pas non plus possible d'étudier l'effet des interactions entre mâles sur la taille des graines. Une troisième et dernière catégorie sont les travaux qui s'intéressent aux effets mâles en termes de dépression de consanguinité. Cependant, ces travaux ne permettent pas de discuter des effets mâles tels qu'on les attend en conditions naturelles.

Plusieurs facteurs peuvent être à l'origine de différences de tailles de graines. On peut imaginer des différences entre génotypes du fait de différences dans les capacités à utiliser

efficacement les ressources, ou du fait de différentes formes de stockage. Alternativement, on peut imaginer des différences entre génotypes qui résultent de différences de compétitivité, auquel cas sur une plante donnée, la quantité de ressources acquises par un génotype (caractérisé par le mâle fécondant) dépendra de la quantité de ressources acquises par les autres génotypes.

L'objet de cette étude est de déterminer dans quelle mesure les différences de masse entre graines fécondées par différents mâles résultent de différences de compétitivité. Pour atteindre cet objectif, nous fécondons des épis de maïs avec plusieurs génotypes mâles, deux par deux. Nous identifions ces mâles grâce à des marqueurs phénotypiques (couleur de l'aleurone) et étudions leurs interactions.

4.2 Matériel et méthodes

4.2.1 Matériel

Six lignées de *Zea mays*, toutes homozygotes, ont été obtenues au 'Maize Genetic CooperationStock Center' (http ://maizecoop.cropsci.uiuc.edu) : A619, W23, W22, B73, Mo17, et une ligne non-identifiée que nous dénommerons B1-Peru. Le locus R détermine la couleur de l'aleurone. Chaque lignée est représentée à la fois par une isolignée récessive au locus R (grain jaune ou blanc), et par une isolignée dominante au locus R (noir ou moucheté), sauf B1-Peru, pour laquelle seul l'isolignée dominante était disponible (la Table 4.1 inventorie chaque isolignée avec le nom de stock et du marqueur). Ces marqueurs nous permettent de reconnaitre les lignées fécondantes sur les épis où des pollens de différentes lignées étaient mélangés : partout où le pollen est mélangé, la femelle porte le marqueur récessif, un mâle porte également le marqueur récessif, et l'autre mâle porte le marqueur dominant.

Lignée	Marqueur	Phénotype	Reference du stock
A619	R1-st	Molted	X18C
A619	c1r1	Yellow	218DB
Mo17	R1-scm2	Black	X26B
Mo17	c1r1	Yellow	218DB
W23	R1-r	Black	M142A
W23	r1-g	Light yellow	M142T
W22	R1-sc	Black	X19A
W22	r1-g	Yellow	X14L
B73	R1-nj	Navajo	X17D
B73	c1r1	Yellow	218DB
NA	B1-Peru	Black	218GA

Table 4.1 – Lignées de *Zea mays* utilisées pour les croisements.

4.2.2 Plan experimental

Le plan expérimental et les croisements effectués sont décrits Table 4.2. Quatre types d'épis peuvent être distingués : (a) les épis 'autofécondés' (fécondation par un mâle de

la même lignée), (b) les épis fécondés par un seul mâle, dont la lignée est différente de celle de la femelle, (c), les épis fécondés par un mélange de deux lignées mâles, toutes les deux différentes de la lignée femelle et (d) les épis fécondés par deux lignées mâles, dont l'une est identique celle de la femelle. Les fécondations sont effectuées selon la méthode classique : le pollen a été récolté dans des sacs en papier et placé sur l'épi choisi. Les croisements ont été pratiqués dans les 10 à 15 minutes suivant le prélèvement du pollen. Dans le cas de croisements en mélange, un des sac contenant du pollen était simplement versé dans l'autre.

		Female						
Male 1	Male 2	A619	B1-Peru	B73	Mo17	W22	W23	Somme
A619	A619	16	11	5	5	13	11	61
A619	B1-Peru	13	0	0	0	15	8	36
A619	B73	5	0	0	1	11	8	25
A619	Mo17	8	0	0	8	17	13	46
A619	W22	19	0	0	6	19	15	59
A619	W23	12	0	0	0	21	20	53
B1-Peru	B1-Peru	10	7	6	5	7	8	43
B1-Peru	B73	13	0	0	5	11	7	36
B1-Peru	Mo17	7	0	0	4	12	9	32
B1-Peru	W22	11	0	0	0	15	10	36
B1-Peru	W23	10	0	0	1	13	15	39
B73	B73	7	10	10	7	8	8	50
B73	Mo17	8	0	0	5	11	7	31
B73	W22	9	0	0	2	13	8	32
B73	W23	11	0	0	2	13	10	36
Mo17	Mo17	9	10	4	8	8	9	48
Mo17	W22	10	0	0	8	12	13	43
Mo17	W23	10	0	0	4	20	16	50
W22	W22	13	14	6	4	13	14	64
W22	W23	8	0	0	3	14	15	40
W23	W23	5	16	7	6	14	14	62
Nombre d'épis		214	68	38	84	280	238	922
Nombre de mesures		368	68	38	132	497	412	1515

Table 4.2 – Nombre d'épis obtenus pour chaque modalité du plan de croisement. Pour chaque épi, les graines sont triées et pesées par génotype mâle. Pour chaque épi où deux pères sont présents, deux mesures de masse sont donc prises.

Le nombre d'ovules à été estimé pour chaque épi dont l'apex n'était pas cassé (nombre de graines sur une colonne plus nombre d'ovules non-fécondés, le tout fois le nombre de colonnes sur l'épi). Une partie des épis ont subi un traitement particulier (cf. 4.2.3). Pour chaque épi, le nombre de grains manquants a été compté, puis l'épi à été égrainé. Les graines ont été triées, comptées et pesées par phénotype. Nous disposons donc pour l'ensemble des épis d'une ou deux mesures de masse (selon qu'un ou deux phénotypes sont présents sur cet épi), chacune associée à un comptage des grains pesés.

4.2.3 Cartographie

Vingt-six épis de femelles W22 correspondant uniquement aux cases du plan expérimental où des pollens étaient mélangés ont été analysés pour l'étude des patrons de réallocation à l'échelle du grain. Sept colonnes et 16 à 21 lignes de grains ont été alignées sur une plaque adhésive pour chaque épi. Chaque ovule non-fécondé ou avorté était indiqué sur la carte par des points, afin de garder l'alignement (fig. 4.1). Les masses, les phénotypes (indicateur des génotypes) et les positions sur la carte (coordonnées en ligne, colonne) ont ensuite été collectés pour chacune des graines. Le nombre de lignes de grains présentes sur l'épi au dessus et au dessous de celles prélevées a été noté, pour calculer la distance (en nombre de grains) de chaque grain sur l'épi relativement à l'apex (variable RelPos pour "position relative").

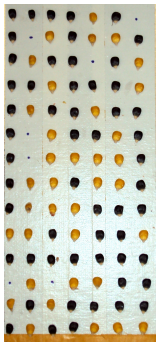

Figure 4.1 – Exemple d'une carte

4.2.4 Analyses statistiques

Nous étudions la masse de chaque grain en fonction de variables relatives à l'épi (nombre de grains par exemple), au grain lui-même (génotype, position), et de variables relatives à ses voisins (masses, génotypes, positions), à travers des analyses de variance (sous R). Toutes nos sélections de modèle sont basées sur des comparaisons d'AIC.

Le détail des analyses statistiques est donné pour chacune de nos deux analyses, au fur et à mesure de la description des résultats. La première analyse est basée sur les cartographies et consiste à étudier les variations de masse des grains en fonction de leur position sur l'épi, et de la masse des grains qui sont a proximité. La seconde analyse est basée sur les mesures de masses globales et consiste à étudier les variations de ces masses globales en fonction des génotypes. L'objectif de ces deux analyses est de détecter des effets de compétition entre grains.

4.3 Résultats

4.3.1 Allocation des ressources à l'échelle locale

L'objet de cette partie est d'étudier les patrons de réallocation à l'échelle du grain. Si les grains sont en compétition pour les ressources, nous nous attendons à ce que les

grains qui sont à proximité de trous et de petits grains soient plus gros.

Effets bloc Nous nous intéressons d'abord aux effets sans signification biologique, c'est à dire aux variations entre épis (variable "ear") et aux effets de la position horizontale (variable "Col", on peut s'attendre à ce que l'orientation du grain vis-à-vis de l'extérieur -pied du maïs, intempéries, etc.- affecte en partie sa taille). Nous nous intéressons également à l'effet de la position relativement à l'apex (variable "RelPos"), qui, elle, pourrait avoir un intérêt biologique (les ressources arrivant par la base).

Nous ajustons donc tout d'abord la masse des grains à ces variables, à l'aide d'un modèle linéaire (fig. 4.2, Eq.).

$$SeedMass_{pjkn} = \mu + \beta_p Ear + \beta_j RelPos + \beta_k Ear * Col + \epsilon_{pjkn}, \tag{4.1}$$

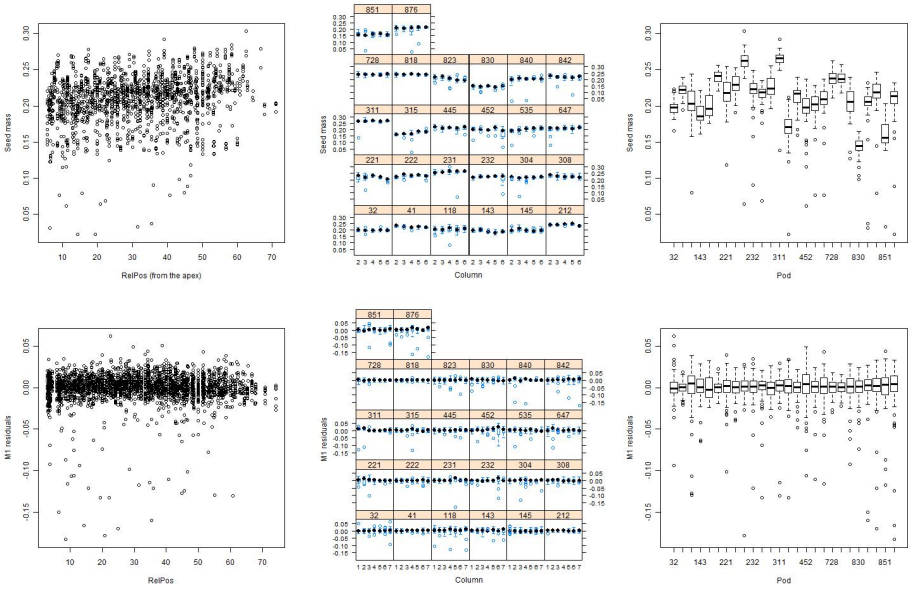

Figure 4.2 – Masse (en haut) et résidus (en bas) par épi et en fonction de la position sur l'épi

Chacune des ces variables a un effet important sur la masse des grains. Les grains situés à la base sont généralement plus gros que ceux situés à l'apex, probablement en raison de leur accès prioritaire aux ressources. L'étude des résidus de ce modèle constitue un point de départ à notre étude. Nous traçons à partir des résidus des variogrammes qui représentent les différences de masse des grains en fonction de leur distance (fig. 4.3).

Comme on peut le voir fig. 4.3, la masse de deux grains est moins corrélée si ces grains sont proches que si ces grains sont éloignés, ce qui indique qu'il y a des effets locaux sur la masse des grains. La majeure partie des effets (visibles dans le variogramme de gauche) disparaissent (variogramme du centre) lorsqu'on prend en compte la position. Cependant, un effet du voisin le plus proche (distance=1) se maintient, ce qui indique qu'il est possible que la masse d'un grain soit corrélée à la masse de son voisin immédiat.

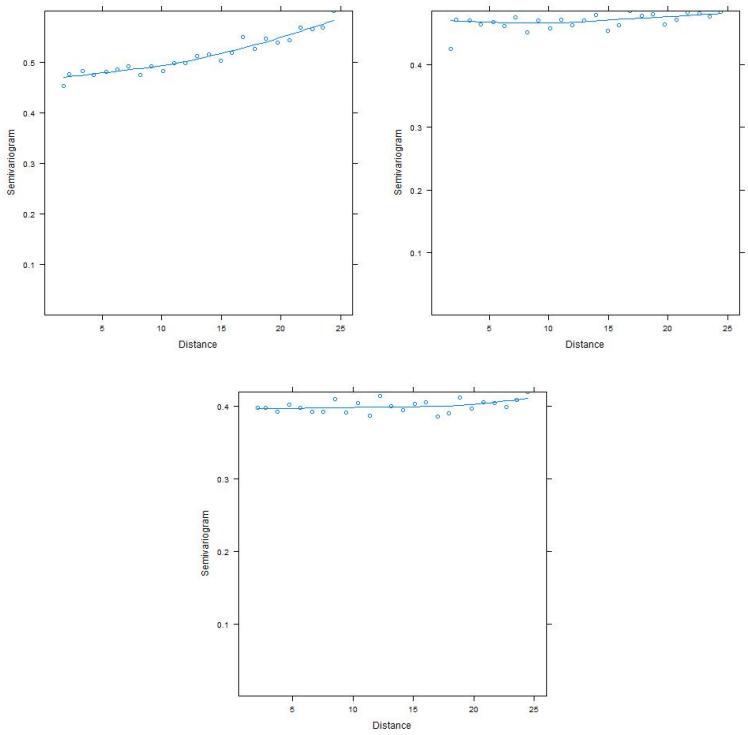

Figure 4.3 – Variogrames, i.e. semivariance moyenne de masse de couples de grains en fonction de la distance entre ces grains. En abscisse est représentée la distance entre grains. Pour chaque classe de distance d, la variance de masse de tous les couples séparés par une distance d est calculée, et moyennée, ce qui donne la semivariance, en ordonnée. Une faible semivariance correspond à une forte corrélation entre les masses des deux grains. En haut à gauche : variogramme des résidus d'un modèle ou seul l'effet épi est pris en compte. En haut à droite : variogramme des résidus du modèle portant sur les effets block. En bas : résidus du modèle complet.

Des modèles ANOVA vont nous permettre d'étudier plus précisément cette éventuelle corrélation. Dans tous ces modèles, les effets des variables 'Ear', 'RelPos' et 'Col' sont pris en compte (et hautement significatifs : p-value ≤ 0.001).

Effets des mâles Les génotypes des grains affectent leurs masses. Nous pouvons voir cet effet fig. 4.4.

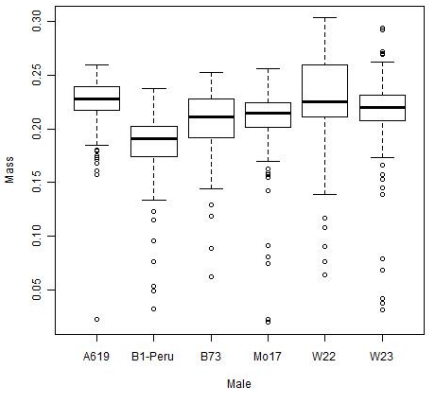

Figure 4.4 – Masse en fonction des génotypes mâles

Effet des voisins Nous prenons en compte plusieurs variables pour décrire les effets des premiers voisins. Une première catégorie de variables décrit l'occurrence de trous (i.e. d'ovules non-fécondés ou avortés, variables NH_T, NH_B, NH_L, et NH_R, NH pour *number of holes* avec T=Top, B=Bottom, L=Left, R=Right, variable qui peut prendre la valeur 0 ou 1). La deuxième catégorie de variables décrit la masse des voisins (zéro en cas de trous, variables SM_T, SM_B, SM_L, et SM_R, SM pour *seed mass*). Nous décrivons également des variables composites qui sont la moyenne des masses des voisins du dessus et du dessous (MSM_{BT}, MSM pour *mean seed mass*), la moyenne des masses des voisins de gauche et de droite (MSM_{LR}), et la moyenne des masses de l'ensemble des voisins (MSM_{BTLR}). Si les voisins du dessus et du dessous (de gauche et de droite) ont le même effet, remplacer les précédentes variables par celles-ci devrait diminuer l'AIC. La troisième catégorie de variables décrit le génotype des voisins (variables CM_T, CM_B, CM_L, et CM_R, CM pour *competing male*). Une variable décrivant non pas le génotype individuel des voisins, mais simplement le génotype du compétiteur présent sur l'épi, a également été testée (Variable $CompetingMale$).

Modèles Nous partons d'un modèle complet dans lequel on estime l'effet de la masse de chaque voisin individuellement. Nous comparons ce modèle complet à des modèles prenant en compte les masses moyennes (moyennes voisin haut+voisin bas et voisin gauche+voisin droite, puis moyenne de l'ensemble des voisins). Dans le modèle complet, on estime également l'effet du génotype de chaque voisin individuellement. Nous comparons ce modèle complet à des modèles pour lequel on ne prend plus en compte que $CompetingMale$ (génotype du compétiteur). Les effets épis et colonne (hiérarchisés dans

épi) sont pris en compte en tant qu'effets aléatoires. La sélection de modèle par AIC nous conduit au modèle 4.2 ($AIC = -8341.32$) :

$$\begin{aligned} Y_{pjkn} &= \alpha + \beta_p Ear + \beta RelPos + \beta_j Ear * Col \\ &+\beta_{q1} Male + \beta_{q2} CompetingMale \\ &+\beta_{k1} NH_B + \beta_{k2} NH_T + \beta_{k3} NH_L + \beta_{k4} NH_R \\ &+\beta_{l1} MSM_{BT} + \beta_{l2} MSM_{LR} + \beta_{l3} MSM_{BT} : MSM_{LR} \\ &+\beta_{m1} NH_L : NH_R \\ &+\beta_{m2} NH_B : NH_L + \beta_{m3} NH_B : NH_R \\ &+\beta_{m4} NH_T : NH_L + \beta_{m5} NH_T : NH_R \\ &+\epsilon_{pjklmn} \end{aligned} \tag{4.2}$$

Les estimations de ce modèle sont données Table 4.3 :

Effect	Estimate
$(Intercept)$	0.143
Ear	$random$
$Col : Ear$	$random$
$HauteurRel$	0.000
$Male = B1 - Peru$	-0.028
$Male = B73$	-0.003
$Male = Mo17$	-0.011
$Male = W22$	0.000
$Male = W23$	-0.003
$Compet = B1 - Peru$	-0.017
$Compet = B73$	-0.012
$Compet = Mo17$	-0.020
$Compet = W22$	0.009
$Compet = W23$	0.000
$NH_B = 1$	0.008
$NH_T = 1$	0.012
$NH_L = 1$	0.013
$NH_R = 1$	0.014
$NH_L = 1 : NH_R = 1$	-0.008
$NH_B = 1 : NH_L = 1$	0.015
$NH_T = 1 : NH_R = 1$	0.005
$NH_B = 1 : NH_R = 1$	0.008
$NH_T = 1 : NH_L = 1$	0.007
SM_{BT}	0.227
SM_{LR}	0.300
$SM_{BT} : SM_{LR}$	-0.825

Table 4.3 – Estimations du meilleur modèle d'analyse des cartographies

On remarque qu'on a à la fois un effet du génotype du grain focal, et du génotype du compétiteur présent sur l'épi. Les trous ont un effet quel que soit leur position (les quatre variables ont été conservées), et ont également un effet d'interaction deux-à-deux. Quasiment tous ces effets sont positifs, ce qui indique comme attendu que la présence d'un trou augmente la taille du grain. Les variables compilant les effets des voisins deux-à-deux ont été conservées, ce qui indique que les voisins du dessus et du dessous d'une part, et

les voisins de droite et de gauche d'autre part, ont le même effet. Les deux variables ont une importante interaction, négative, ce qui indique que si la masse moyenne des voisins des cotés et la masse moyenne des voisins du haut et du bas sont toutes les deux grandes, la masse du grain focal est diminuée (l'effet est de $0.227 + 0.3 - 0.825 = -0.595$ dans ce cas). On peut remarque que cet effet des voisins est plus important que l'effet du propre génotype de la graine (effet mâle de -0.028 au maximum).

La variable moyennant les masses des quatre voisins n'a pas été retenue, ce qui indique que les voisins des cotés ont un effet différent (moins important) que les voisins du dessus et du dessous. On peut imaginer que cela soit lié à la manière dont les ressources sont véhiculées dans l'épi (de bas en haut).

Comme on peut le voir dans le variogramme de droite Fig. 4.3, plus aucune corrélation n'est observée entre les résidus de cette analyse.

Globalement, ces résultats indiquent l'effet d'une compétition très locale : la taille d'un grain diminue fortement lorsque la taille des grains voisins est importante. Cette compétition locale est déterminante dans la taille du grain. L'un des pré-requis de la théorie du conflit sur l'allocation des ressources est que les embryons soient en compétition pour les ressources. Ces résultats sont cohérents cette hypothèse. L'étude qui suit s'intéresse aux patrons d'allocation des ressources à l'échelle de l'épi et doit permettre d'étudier plus précisément les modalités d'une éventuelle compétition.

4.3.2 Allocation moyenne des ressources

L'objet de cette partie est d'étudier les patrons de réallocation à l'échelle de l'épi. Nous étudions un modèle non-linéaire, dans lequel certains termes mettent en jeu de la compétition, quand d'autres sont indépendants de la compétition. Si les grains sont en compétition pour les ressources, nous nous attendons à ce que les termes qui mettent en jeu la compétition soient nuls. Cette étude est en cours de rédaction pour un article, nous ne donnons donc pas ici tous les détails, qui sont donnés Annexe B.

4.3.2.1 Modèle

Nous analysons la masse globale des grains d'un génotype donné, sur un épi donné, en fonction de plusieurs termes. Ces termes correspondent à différents modes d'allocation des ressources. Si M_{ijkl} est la masse globale des grains d'un génotype sur un épi l, le modèle peut être écrit :

$$M_{ijkl} = I_i f_j + I_j n_j + (D_i + D_j n_j + D_k n_k) f^*_{jki} \tag{4.3}$$

Nous détaillons les termes de ce modèle ci-après.

Termes indépendants de la compétition Nous considérons deux modes d'allocations qui sont indépendants de la compétition. La plante pourrait allouer une quantité donnée de ressources à chaque épi. Comme cette ressource est allouée par la plante, nous supposons qu'elle peut dépendre du génotype femelle (i). I_i est cette quantité de ressource (I pour Indépendant de la compétition). Une graine donnée reçoit une fraction $1/(n_j + n_k)$ de I_i (n_j et n_k étant le nombre de graines fécondées respectivement par les mâles j et k). De ce fait, le groupe de graines fécondées par le mâle j reçoit une fraction $f_j = n_j/(n_j + n_k)$ de I_i, c'est à dire une quantité de ressources égale à $I_i f_j$.

Une autre partie des ressources pourrait être amenée jusqu'à la graine par un effet puit. Un tel effet devrait dépendre du génotype du mâle fécondant, j. I_j est cette quantité de ressource. Le groupe de graines fécondées par le mâle j reçoit une quantité de ressources égale à $I_i f_j$.

Termes dépendants de la compétition Nous considérons également deux modes d'allocations qui sont dépendants de la compétition. Une plante de génotype i pourrait allouer une quantité de ressource non pas directement au grain, mais à un pool de ressources partagé en fonction de paramètres de compétition. D_i (D pour Dépendant de la compétition) est cette quantité de ressources. La fraction de D_i acquise par le mâle j en compétition avec le mâle k sur la femelle i s'écrit :

$$f^*_{jki} = \frac{f_j \alpha_{jki}}{f_j \alpha_{jki} + f_k \alpha_{kji}}; \tag{4.4}$$

où $f_k = n_k/(n_k+n_j)$, α_{jki} est la compétitivité de j face à k, et α_{kji} est la compétitivité de k face à j, sur la femelle i. Si nous supposons que α_{jki} et α_{kji} ne sont pas très différents, nous pouvons supposer que $\alpha_{jki}/\alpha_{kji}$ est proche de 1, auquel cas :

$$f^*_{jki} = f_j + a_{jki} f_j f_k + o\left(\frac{\alpha_{jki}}{\alpha_{kji}}\right)^1 \tag{4.5}$$

Le groupe de graines fécondées par le mâle j reçoit une quantité de ressources égale à $D_i f^*_{jki}$ par ce biais.

Enfin, la quantité de ressources présentes dans un pool partagé en fonction des paramètres de compétition pourrait être augmentée, elle aussi, par un effet puit mâle-dépendant. D_j est la quantité de ressource apportée dans ce pool par une graine fécondée par le mâle j et D_k est la quantité de ressource apportée dans ce pool par une graine fécondée par le mâle k. La quantité de ressource qui arrive dans le pool par effet puit est $D_j n_j + D_k n_k$. Comme précédemment, le génotype acquiert une fraction f^*_{jki} de cette ressource, soit une quantité $(D_j n_j + D_k n_k) f^*_{jki}$.

Nous ajustons différentes variantes de ce modèle par la minimisation des carrés moyens et effectuons une sélection de modèle en comparant les AIC.

4.3.2.2 Résultats et discussion de l'analyse de l'allocation moyenne des ressources

Le meilleur modèle indique :

1. qu'entre 50% et la totalité de la ressource allouée à l'épi l'est via l'effet du puit ;
2. que l'ensemble des ressources sont allouées par un mode de compétition (les termes indépendants de la compétition ne sont pas significatifs) ;
3. que les capacités de compétition des mâles sont transitives et ne dépendent que du génotype du mâle, pas de la femelle.

L'étude des estimations indique cependant que les différences de compétitivité des mâles ne sont pas très marquées, car elles ne créent, au maximum, qu'une différence de masse de 12%.

Ces résultats mettent en évidence que l'allocation globale des ressources repose essentiellement sur la compétition. Par ailleurs, le patron que nous décrivons souligne l'importance des effets de puit, il est donc compatible avec l'idée que la masse du grain de maïs est limité par l'effet puit dans la plupart des conditions (Borras *et al.* 2004).

4.4 Discussion

Les résultats de nos deux études mettent en évidence de la compétition entre grains. La première étude montre qu'il y a une compétition très locale, entre grains voisins, et la seconde montre qu'il y a une compétition globale entre génotypes sur un même épi. L'un des pré-requis de la théorie du conflit sur l'allocation des ressources est qu'il y ait de la compétition entre embryons sur une plante. Ce pré-requis est souvent implicitement considéré comme acquis, et l'hypothèse selon laquelle les embryons sont bien en compétition n'avait jamais été explicitement testée. Par notre travail, nous avons à la fois mis en place une méthodologie pour effectuer ce test, et confirmé que chez le maïs au moins, les différentes graines sont bien en compétition.

De plus, deux points de vue coexistent sur la théorie du conflit sur l'allocation des ressources dans la littérature. L'un de ces points de vue est que l'optimum de la femelle est d'allouer les ressources de manière équitable, alors que l'optimum du mâle est de biaiser l'allocation des ressources. L'autre de ces points de vue est que l'optimum de la femelle n'est pas d'allouer les ressources équitablement, mais plutôt en fonction de la qualité génétique des graines (Westoby & Rice 1982; Willson & Burley 1983; Haig 1990). De ce point de vue là, l'optimum du mâle est de ne pas être avorté (cf. chapitre d'introduction). Dans notre expérience, les ressources ne sont pas allouées équitablement entre les embryons. Deux interprétations sont possibles. Premièrement, l'optimum femelle pourrait être égalitaire, mais l'expression des gènes paternels dans l'albumen pourrait donner une part de contrôle aux mâles, ce qui leur permet de biaiser l'allocation des ressources. Cette interprétation est compatible avec l'observation que l'on a faites des patrons moyens d'allocation des ressources. Alternativement, l'optimum de la femelle pourrait être d'allouer les ressources en fonction de la qualité génétique des graines. Il a été proposé que le mécanisme de cet avortement sélectif soit la compétition ; les embryons les moins compétitifs n'acquérant pas suffisamment de ressources pour survivre. A première vue, cela pourrait sembler compatible avec nos patrons. Cependant, la qualité génétique devrait dépendre du nombre de mutations délétères dans les gamètes haploïdes, mais également de la combinaison entre ces deux génomes haploïdes, plus ou moins hétérotique. Or, d'après nos résultats, les capacités de compétition des mâles sont indépendantes du génome femelle, alors qu'on attendrait une dépendance si les capacités de compétition reflétaient l'hétérosis. Il est donc peu plausible que la compétitivité des mâles reflète leurs qualités génétiques, ce qui soutient plutôt le point de vue que l'optimum de la femelle est d'allouer équitablement les ressources, mais qu'elle n'a pas complètement le contrôle de cette allocation.

Chapitre 5

Quelle théorie pour l'évolution de l'empreinte parentale ?

Bien que la double fécondation et les variations dans le nombre de contributions maternelles d'une part, et l'empreinte parentale d'autre part soient des traits très différents, une seule et même théorie pourrait expliquer leurs évolutions : la théorie du conflit sur l'allocation des ressources. Comme nous l'avons précisé chapitre 3, la double fécondation telle qu'on la rencontre chez les angiospermes n'est apparue qu'une seule fois. L'empreinte parentale, bien qu'elle n'ait été étudiée que chez quatre espèces d'angiospermes (en comptant *A. lyrata* et le riz), est apparue plusieurs fois (voir section 1.2.2 en chapitre d'introduction), ce qui permet d'étudier l'évolution de l'empreinte plus facilement que l'évolution de la double fécondation. Or, s'il était démontré qu'une des théories proposées pour expliquer spécifiquement l'évolution de l'empreinte parentale est beaucoup plus explicative que la théorie du conflit sur l'allocation des ressources pour ce trait, cela indiquerait qu'il est pertinent de rechercher des théories alternatives aussi pour la double fécondation et les variations dans le nombre de contribution maternelle.

En ce qui concerne l'évolution de l'empreinte parentale, la théorie du conflit sur l'allocation des ressources est la théorie explicative qui a le plus été explorée par la littérature. Chez les animaux, la théorie semble en effet être la meilleure explication disponible à ce jour (Hurst 1997), bien que certains points ne concordent pas (McVean & Hurst 1997; Hurst 1997; Hurst & McVean 1997, 1998). En revanche, en ce qui concerne l'empreinte parentale chez les plantes, personne n'a réellement rassemblé les informations disponibles pour tester si la théorie du conflit sur l'allocation des ressources apparait comme la meilleure explication, bien que certains auteurs récusent l'idée qu'une théorie alternative puisse être plus explicative (voir Moore & Mills 2008).

En analysant les informations disponibles dans la littérature, nous estimons si les prédictions de la théorie du conflit sur l'allocation sont vérifiées par les données. Toutes ne le sont pas. Nous étudions donc ensuite la capacité des théories alternatives à valider leurs prédictions et à expliquer les données (section 5.2) pour enfin conclure quant à la question : 'la théorie du conflit d'allocation est-elle la meilleure théorie disponible à ce jour pour expliquer l'empreinte parentale chez les angiospermes ?'.

5.1 L'empreinte parentale et le conflit d'allocation des ressources

En cas de conflit sur l'allocation des ressources, on s'attend à (Hurst 1997; Hurst & McVean 1998; Charalambous *et al.* 2003) :

1. observer un patron de coévolution antagoniste (une empreinte sous contrôle maternel doit répondre à une empreinte sous contrôle paternel, McVean & Hurst 1997; Hurst & McVean 1998) ;
2. ce que l'empreinte ne se maintienne pas au delà de la période d'investissement parental (Hurst & McVean 1998) ;
3. ce que la modification du statut (actif ou éteint) d'un allèle qui est sous contrôle paternel provoque un excès de croissance quand la modification du statut d'un allèle qui est sous contrôle maternel entraîne un déficit de croissance (Hurst 1997; Hurst & McVean 1998) ;
4. ce que l'occurrence de l'empreinte soit corrélée à la polyandrie (plus le nombre de mâles fécondants est important, plus le conflit est supposé intense, Hurst & McVean 1998) ;
5. ce que lorsque les ressources ne sont pas allouées à un embryon (parce l'empreinte du gène est affecté par exemple), elles puissent être réallouées aux autres embryons (Charalambous *et al.* 2003).

Nous reprenons ces cinq critères un à un pour évaluer la mesure dans laquelle ils valident ou non la théorie du conflit sur l'allocation des ressources.

5.1.1 Coévolution antagoniste

L'une des prédictions de la théorie du conflit sur l'allocation des ressources est que l'on devrait observer un patron de coévolution antagoniste entre les empreintes. En effet, les inhibiteurs et activateurs de croissance sont sujets à des pressions de sélection différentes selon qu'ils sont exprimés par le mâle ou par la femelle. Cela devrait rendre leur niveau d'expression optimum parent-dépendant et générer des empreintes parentales. On s'attend, plus précisément, à ce qu'une empreinte qui est sous contrôle maternel réponde à une empreinte qui est sous contrôle paternel (McVean & Hurst 1997; Hurst & McVean 1998).

Par ailleurs, si les gènes impliqués dans la croissance sont sujets à des pressions de sélection différentes selon qu'ils sont exprimés paternellement ou maternellement, on s'attend également à ce que les régulateurs de ces gènes soient l'objet de fortes pressions, ce qui pourrait augmenter le taux d'évolution de leurs séquences.

5.1.1.1 Taux d'évolution des séquences

Chez les angiospermes, l'analyse des séquences du gène à empreinte *MEDEA* d'*A. lyrata* à pu montrer que le taux de substitution est non neutre au niveau de son promoteur (Miyake *et al.* 2009; Kawabe *et al.* 2007), contrairement à l'analyse des séquences de *MEA* d'*A. thaliana.* Cela est conforme aux prédictions du conflit male-femelle, car *A. lyrata* est allofécondante, alors qu'*A. thaliana* est autofécondante.

Encadré 5.1 – Régulateurs de l'empreinte parentale chez *A. thaliana*

Les régulateurs d'empreintes connus concernent quasi-exclusivement *A. thaliana*. Il s'agit de DEMETER (DME), DNA METHYLTRANSFERASE1 (MET1), le groupe Polycomb (PcG), et DECREASED DNA METHYLATION1 (DDM1).

- DME intervient dans l'**activation** d'allèles **maternels** (il code pour une glycosylase qui déméthyle l'ADN par remplacement des bases (Choi *et al.* 2002)). Il est contrôlé en amont par *AGL80*, dont l'expression est restreinte au gamétophyte maternel (et qui est potentiellement à empreinte). L'expression de *DME* est donc elle aussi restreinte au gamétophyte maternel, plus précisément à la cellule centrale (et à certains tissus tels que les méristèmes). Elle chute brutalement après la fécondation (Choi *et al.* 2002).
- Le PcG intervient dans l'**extinction** d'allèles **maternels** et **paternels** (il intervient dans la méthylation des histones). Il est formé de plusieurs sous-unités dont MEA, FIS2, FIE et MSI1. Les gènes du PcG semblent être exprimés paternellement, avant fécondation et dans le gamétophyte femelle, avant et après fécondation (Cf Table 5.1).
- MET1 intervient dans le maintien de l'**extinction** d'allèles **paternels** (il intervient dans la méthylation de l'ADN). Il n'est exprimé que dans les tissus végétatifs et les gamètes mâles, il n'est pas exprimé dans les gamètes femelles (Xiao *et al.* 2006; Fitzgerald *et al.* 2008; Jullien *et al.* 2008).
- DDM1 intervient dans l'**extinction** d'allèles **paternels**. Il est exprimé dans l'albumen et le zygote (Vielle-Calzada *et al.* 1999).

5.1.1.2 Coévolution des empreintes

Nous avons rassemblé l'ensemble des informations nécessaires à l'étude de la coévolution des empreintes Table 5.1. La plupart des gènes pour lesquels nous disposons d'une information, c'est-à-dire pour lesquels les mécanismes de régulation sont connus, concernent *A. thaliana* (encadré 5.1).

Le déséquilibre entre le nombre d'empreintes pour lesquelles c'est l'allèle paternel qui est actif (une seule, l'empreinte de *PHERES1*) et le nombre d'empreintes pour lesquelles c'est l'allèle maternel qui est actif (au moins cinq chez *A. thaliana*) suggère l'absence de coévolution antagoniste en ce qui concerne les empreintes chez les angiospermes.

Un autre critère important est souvent occulté. Pour qu'une empreinte soit l'indice d'une coévolution antagoniste, il faut que le contrôle du statut (actif *ou* inactif) des allèles soit uniparental. En effet, un contrôle signifie à la fois mise en place et maintien. Ainsi, si une extinction est mise en place paternellement et maintenue par des régulateurs exprimés maternellement, le contrôle est biparental. C'est justement le cas pour *MEA*. L'extinction de l'allèle paternel est mise en place dans les gamètes mâles par le PcG (encadré 5.1). Puis elle est maintenue, après la fécondation, par le biais de facteurs exprimés par les chromosomes femelle (rétrocontrôle par la copie maternelle de *MEA*, *MEAm*). *MEA* ne correspond donc pas à l'attendu d'un contrôle uniparental du statut de chaque allèle.

Espèce	Gène (Fiabilité)	Expression végétative	Expression embryonnaire	Contrôle de l'allèle maternel	Contrôle de l'allèle paternel	Ref
A. thaliana	*FIS2* (+)	Non (MET1)	Non	Act./DME* + NA	Ext./MET1	Jullien *et al.* 2006a; Choi *et al.* 2002; Portereiko *et al.* 2006; Luo *et al.* 2000
A. thaliana	*FWA* (+)	Non (MET1)	Non*	Act./DME	Ext./MET1, DDM1	Kinoshita *et al.* 2003, 2004a,b, 2007; Jullien *et al.* 2006a; Ikeda *et al.* 2007
A. thaliana, *A. lyrata*	*MEA* (+)	Non	Non	Act./DME	Ext./PcG, DDM1, Maintien/MEAm	Vielle-Calzada *et al.* 1999; Dickinson & Scott 2002; Gehring *et al.* 2006; Baroux *et al.* 2006; Wang *et al.* 2006
A. thaliana	*MPC* (+)	Non (MET1)	Très faible (Biall.)	Act./DME + NA	Ext./MET1	Tiwari *et al.* 2008
A. thaliana	*PHE1* (+)	Non*	Oui	Ext./PcG + meth. (region distante)	Act./NA	Köhler *et al.* 2003; Makarevich *et al.* 2006, 2008
A. thaliana	*AtFH5* (+)	Non (PcG)	Non	Act./NA	Ext./PcG	Ingouff *et al.* 2005; Fitzgerald *et al.* 2009
A. thaliana	*AGL80* (-)	Oui	Non	Act./NA	Ext./NA	Portereiko *et al.* 2006
A. thaliana	*HDG9* (-)	Non	Non	Act./NA	Ext./NA	Nakamura *et al.* 2006
A. thaliana	*HDG8* (-)	Non	Oui (précoce)	Act./NA	Ext./NA	Nakamura *et al.* 2006
A. thaliana	*HDG3* (-)	Non	Oui	Ext./NA	Act./NA	Nakamura *et al.* 2006
A. thaliana	*ATMYBR2* (-)	NA	NA	Act./NA	Ext./NA	
A. thaliana	*AT5G62110* (-)	NA	NA	Ext./NA	Act./NA	
Zea mais	*ZmFIE1* (+)	Non	Non	Act./Demeth.	Ext./îlots CpG*	Springer *et al.* 2002; Danilevskaya 2003; Hermon *et al.* 2007
Zea mais	*MEG1* (+, transit.)	Non	Non	Act./NA	Ext./NA	Gutierrez-Marcos *et al.* 2004
Zea mais	*Nrp1* (-)	NA	Non	Act./NA	Ext./PcG*	Guo *et al.* 2003
Zea mais	*Mez1* (+)	Oui (Biall.)	Oui (Biall.)	Act./NA	Ext./CpG, Meth.*	Haun *et al.* 2007; Haun & Springer 2008; Haun *et al.* 2009
Zea mais	*R* (ea, +)	AD	AD	Ext./NA	Act./NA	Kermicle 1970
Zea mais	*Tubulin* (ea, -)	NA	NA	Ext./Meth.	Act./NA	
Zea mais	*dzr1* (ea, -)	Non	Non	Act./NA	Ext./Meth.	Chaudhuri & Messing 1994
Zea mais	*Mee1* (-)	Non	Oui (Mat., transit. : 3-8 jap)	Act./NA	Ext./Meth.	Jahnke & Scholten 2009
Oryza sativa	*OsFIE1* (-)	Non	Non	Act./NA	Ext./NA	Luo *et al.* 2009

Table 5.1 – Contrôle de la mise en place des empreintes. L' "Expression végétative" correspond à une expression dans des tissus qui ne sont pas impliqués dans la reproduction. Abréviations : ea=expression allélique, AD=Allèle dépendant, Act.=Activation, Ext. = Extinction, transit.=transitoire, Meth.= méthylation, Mat. = Maternelle, Biall.= Biallélique, NA=Inconnu, *=à confirmer

5.1.1.3 Coévolution antagoniste : synthèse

Globalement, les données disponibles ne correspondent donc pas à la prédiction d'une coévolution antagoniste.

5.1.2 Timing d'expression des traits

L'empreinte ne doit pas se maintenir au delà de la période d'investissement parental pour correspondre aux prédictions de la théorie. Comme on l'a vu, chez les angiospermes, la plupart des empreintes sont dans l'albumen. Ce tissu meurt au fur et à mesure que les ressources y sont accumulées. La seule empreinte identifiée dans l'embryon (*Mee1*) est transitoire. L'empreinte ne se maintient donc pas au delà de la période d'investissement parental, conformément à la prédiction de la théorie du conflit sur l'allocation.

5.1.3 Effets parentaux

La théorie du conflit sur l'allocation prédit que les activateurs de croissance sont maternellement éteints et paternellement actifs, et inversement pour les inhibiteurs de croissance (3e prédiction).

Cette prédiction est probablement la plus importante, car en ce qui concerne les autres prédictions, il est souvent possible de trouver des explications palliatives lorsqu'elles ne sont pas vérifiées. Si la prédiction qui concerne les effets sur la croissance n'est pas vérifiée, les données ne correspondent pas au coeur de la théorie : l'optimum paternel est d'augmenter l'allocation des ressources à l'embryon, alors que l'optimum maternel est de contrebalancer les biais induits par les gènes d'origine paternelle.

Nous testons cette prédiction en faisant l'inventaire des fonctions connues des gènes qui ont une empreinte.

5.1.3.1 Effets doses

Le croisement réciproque de plantes de niveaux de ploïdie différentes (Fig. 5.1) altère le ratio des génomes parentaux dans l'albumen. Par exemple, le croisement $4n$ (femelle) x $2n$ (male) de plantes dont le ratio normal dans l'albumen est $2m$: $1p$ produit des albumens $4m$: $1p$ (excès maternel), alors que le croisement $2n$ (femelle) x $4n$ (male) produit des albumens $2m$: $2p$ (excès paternel comparé au ratio normal).

En cas d'altération du ratio des gènes maternels et paternels dans l'albumen, il se produit un certain nombre de modifications dans le phénotype de la graine. Ces modifications du phénotype n'ont pas lieu si on modifie le niveau de ploïdie de la plante sans modifier le ratio 2 : 1 dans l'albumen (c'est-à-dire si le ratio est 4 : 2 par exemple).

Ces effets indiquent que les contributions maternelles et paternelles ne sont pas équivalentes, et cette absence d'équivalence a rapidement été attribuée aux empreintes parentales. Les phénotypes observés lors de ces croisements sont donc généralement interprétés en termes d'effet des gènes à empreinte. Si cette interprétation est juste, ces croisements donnent des indications sur la fonction de ces gènes (cependant, voir encadré 5.2).

Dans ces expériences de croisement réciproque, qui ont été réalisées chez un grand nombre d'espèces, les graines avortent souvent précocement. Cependant, chez les espèces qui initient un développement (*A. thaliana*), un excès maternel conduit généralement à

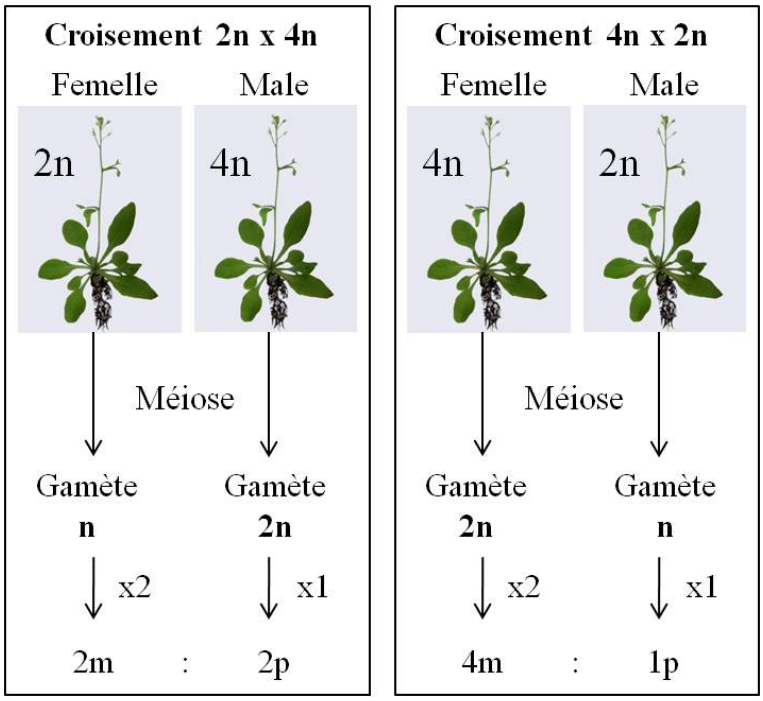

Figure 5.1 – Croisements altérant le ratio des gènes dans l'albumen.

Encadré 5.2 – Les altérations du ratio des contributions parentales sont-ils un indicateur de l'effet des empreintes ?

Certains résultats récents indiquent que lors des croisements conduisant à des ratios altérés, les changements de dosage dérégulent complètement les empreintes (Jullien & Berger 2010). Notamment, chez *A. thaliana*, dans le croisement 2n x 4n (excès paternel), l'extinction de *FIS2* disparait, conduisant à l'expression biparentale de ce gène (mais également, paradoxalement, à une expression globale moindre). Dans ce même croisement, l'expression de la copie maternelle de *MEA* est diminuée au point que la copie paternelle est exprimée de manière prédominante (Jullien & Berger 2010). D'après ces résultats, il est risqué de faire un lien direct entre les phénotypes observés (excès ou déficit de croissance) et l'empreinte parentale.
Cependant, le statut (expression préférentiellement maternelle *versus* paternelle) de la plupart des empreintes étudiées par Jullien & Berger (2010) n'était pas altéré.

Encadré 5.3 – L'autre fonction des FIS

Les *FIS* (Fertilization Independant Seed formation) d' *Arabidopsis thaliana* jouent un rôle crucial avant fécondation. En l'absence de fécondation, certaines des mutations affectant ces gènes peuvent initier le développement autonome de l'albumen (Ohad *et al.* 1996; Curtis & Grossniklaus 2008). Trois gènes sont concernés : *MEA* (ou *FIS1*), *FIS2* et *FIE* (ou *FIS3*, Chaudhury & Berger 2001). En fait, les gènes *FIS* sont répresseurs de gènes cibles spécifiques, impliqués dans le contrôle de l'initiation du développement de la graine (Grossniklaus & Schneitz 1998). Les deux premiers de ces gènes ont également une empreinte parentale (cf. paragraphe 1.1.4). Les gènes homologues de *FIE* d'*Arabidopsis thaliana* chez le riz et le maïs (*OzFIE1*, *OzFIE2*, *ZmFIE1*, *ZmFIE2*) ne semblent, eux, pas impliqués dans la répression du développement autonome de l'albumen (Danilevskaya 2003; Luo *et al.* 2009).

la production d'une graine plus petite, quand un excès paternel conduit généralement à la production d'une graine plus grosse. Les causes possibles de ces phénotypes ont été étudiées pour certaines plantes cultivées et sont revues par Vinkenoog *et al.* (2003). L'excès paternel semble retarder la cellularisation de l'albumen (Orge, Maïs), ce qui permet un plus grand nombre de mitoses et donc un plus grand nombre de cellules dans l'albumen. D'autres effets ont été observés : irrégularités mitotiques (Orge, Seigle), une activité méristématique anormale au niveau du micropyle (Orge, Maïs), dépôt d'amidon retardé et moindre (Orge, Maïs). L'excès maternel a simplement pour conséquence d'avancer la cellularisation et le dépôt d'amidon. Sachant que la quantité de ressources accumulées dépend principalement du nombre de cellules de l'albumen (Brocklehurst 1977; Jones *et al.* 1996) et du nombre et du volume des grains d'amidon, on comprend pourquoi la graine est plus grosse en cas d'excès paternel.

Il en a été déduit que les gènes à empreinte maternellement éteints ou paternellement actifs ont probablement pour conséquence un accroissement des ressources allouées à l'embryon (i.e. ce sont des activateurs de croissance), tandis que les gènes à empreinte paternellement éteints ou maternellement actifs ont probablement un effet inverse (i.e. ce sont des inhibiteurs de croissance), ce qui correspond à la prédiction du conflit sur l'allocation des ressources.

5.1.3.2 Étude des mutants

Les gènes ayant une empreinte parentale chez les angiospermes ont été inventoriés dans le chapitre d'introduction, Table 1.1. Certains ont un effet gamétophytique avant fécondation : *FIS*, *MEA* et *FIS2* (initiation autonome du développement de l'albumen, encadré 5.3). Dans ce paragraphe, nous nous concentrons sur les effets de ces gènes après fécondation.

Sur tous les gènes pour lesquels une information est disponible, on peut distinguer quatre catégories :

- Les gènes qui semblent avoir un rôle direct sur la croissance ;
- Les gènes qui présentent une fonction de régulation des empreintes, et dont le rôle dans l'accumulation des ressources est au mieux indirect ;
- Un gène qui correspond à l'opposé de l'attendu théorique ;

- Les gènes qui ne semblent pas avoir de fonction cruciale dans le développement ou dont l'expression biparentale n'a pas d'effet.

Les gènes qui semblent avoir un rôle direct sur la croissance sont *MPC* et *MEG1*. La réduction d'expression de *MPC* par ARNs interférents génère de plus gros embryons (Tiwari *et al.* 2008). Sachant que MPC est exprimé maternellement, cela correspond à l'attendu du conflit. Chez le maïs, *MEG1* est exprimé exclusivement dans les cellules de transfert, et pourrait donc être impliqué dans le transfert des ressources (Gutierrez-Marcos *et al.* 2004).

Les gènes dont le rôle dans l'accumulation des ressources est au mieux indirect sont *MEA*, *FIS2*, et *AGL80*.

Ces gènes présentent une fonction de régulation d'autres empreintes, mais certains ont également un rôle dans la différentiation de l'albumen. *MEA* et *FIS2* codent pour deux des sous-unités du PcG (groupe Polycomb). Le PcG contrôle les extinctions paternelles d'*AtFH5* (qui est un régulateur du cytosquelette, et serait également activateur de croissance, Fitzgerald *et al.* 2009) et *MEA* (on a donc un contrôle rétroactif pour ce gène, Jullien *et al.* 2006b; Baroux *et al.* 2006; Gehring *et al.* 2006), ainsi que l'extinction maternelle de *PHERES1* (dont la fonction précise n'est pas claire : la mutation *pheres1* n'a pas d'effet notable lorsque qu'elle est hérité paternellement, bien que l'allèle paternel est celui qui est paternellement actif, Köhler *et al.* 2003), probablement via la fonction de déméthylation de l'histone H3K27. Le PcG à également pour cible dans la cellule centrale un régulateur du cycle cellulaire (Skp1-like protein, Köhler *et al.* 2003).

Probablement du fait de leurs contrôles sur *AtFH5*, un effet de certaines mutations de *MEA* et *FIS2* est bien d'inhiber la croissance. Les mutations *mea* et *fis2* ont pour conséquence la dérégulation de *PHERES1*, *AtFH5* et le régulateur du cycle cellulaire, ce qui entraine des changements i) dans le timing de développement de l'albumen, ii) dans le degré de prolifération mitotique et iii) sur le moment où la cellularisation de l'albumen se produit (Köhler *et al.* 2003). L'augmentation de la durée de la phase mitotique syncitiale augmente la croissance de l'albumen, et ce-faisant la taille de la graine (Kiyosue *et al.* 1999; Ingouff *et al.* 2005). Les gènes *FIS* ont également un rôle dans la détermination de l'axe antéropostérieur de l'albumen via la migration orientée des noyaux dans le syncytium (Guitton *et al.* 2004). Vu son expression restreinte au pôle postérieur de l'albumen, le régulateur du cytosquelette *AtFH5* est probablement l'intermédiaire impliqué dans ce phénotype (Sorensen *et al.* 2001).

Enfin, *AGL80* contrôle l'expression de DME (Portereiko *et al.* 2006), qui est lui-même impliqué dans l'activation maternelle de plusieurs des gènes à empreinte.

Le gène qui correspond à l'opposé de l'attendu théorique est *AtFH5*, qui est exprimé maternellement, serait un activateur de croissance, ce qui est contraire à l'attendu du conflit. Comme on l'a précisé ci-dessus, c'est également un régulateur du cytosquelette (Fitzgerald *et al.* 2009).

Les gènes qui ne semblent pas avoir de fonction cruciale dans le développement ou dont l'expression biparentale n'a pas d'effet sont *FWA*, *HDG3*, *HDG8*, *HDG9* et *mez1*. Les gènes *HOMEODOMAIN GLABROUS* (*FWA*, *HDG3*, *HDG8* et *HDG9*) semblent être redondants (Kinoshita *et al.* 2004b; Nakamura *et al.* 2006). Leur rôle parait plus important dans le développement de la fleur (Les mutants *hdg3* présentent les phénotype feuilles recourbées, male stérile, et *HDG3* à également un rôle

dans la déhiscence des anthères, Li *et al.* 2007). Chez le maïs, l'expression biparentale de *mez1* n'a pas d'effet phénotypique.

Les fonctions des autres gènes à empreinte sont moins claires (Berger & Chaudhury 2009). Mais globalement, les données qui concernent la fonction des gènes apportent donc un soutien mitigé à la théorie du conflit sur l'allocation des ressources.

5.1.3.3 Effets parentaux : synthèse

Comme on l'a vu, l'excès de copies maternelles (croisements 4n x 2n) génère des graines sous-développées, tandis que l'excès de copies paternelles (2n x 4n) génère des graines plus grosses que la normale. Ce patron semble soutenir la théorie selon laquelle l'empreinte parentale est un produit du conflit sur l'allocation des ressources, si on admet que ces phénotypes (excès et déficit de croissance) sont dus à l'empreinte.

Cependant, lorsqu'on s'intéresse spécifiquement à la fonction des gènes à empreinte via des expérience de mutagénèse, on remarque que le soutien est très mitigé.

Chez les angiospermes, il est possible de produire des embryons parthénogénétiques malgré l'empreinte parentale (voir par exemple Guitton *et al.* 2004). Chez les mammifères, cela n'est pas possible (Hurst 1997). Par ailleurs, si chez les mammifères l'expression biparentale de certains gènes à empreinte donne lieu à des effets aussi brutaux que l'avortement embryonnaire (Hurst 1997), ça ne semble pas être le cas chez les angiospermes (au moins en ce qui concerne *MEA*, D. Grimanelli et D. Autran, comm. pers.). Ces éléments suggèrent que l'empreinte parentale a des effets moins importants chez les angiospermes que chez les mammifères, ce qui va plutôt en défaveur de la théorie du conflit pour les angiospermes.

En ce qui concerne les effets parentaux, il existe donc des éléments allant en faveur et des éléments allant en défaveur de la théorie du conflit sur l'allocation des ressources.

5.1.4 Corrélation aux traits d'histoire de vie

La corrélation entre l'occurrence de l'empreinte et la polyandrie est difficile à tester, du fait que l'empreinte n'a strictement été mise en évidence que chez trois espèces (en comptant *A. lyrata*). Le maïs est une plante allofécondante, et on s'attend donc d'après les prédictions théoriques à observer une empreinte parentale chez cette espèce. Au contraire, *A. thaliana* est une plante autofécondante. L'occurrence de l'empreinte chez cette espèce est donc avancée comme un élément n'étant pas compatible avec la théorie.

5.1.5 L'empreinte parentale des plantes est-elle correctement expliquée par la théorie du conflit sur l'allocation des ressources ?

En ce qui concerne les prédictions sur la corrélation aux traits d'histoire de vie et la coévolution antagoniste, il y a à la fois des éléments allant en faveur, et des éléments allant en défaveur de la théorie du conflit d'allocation des ressources. En ce qui concerne le point qui pourrait paraître le plus important, celui sur la corrélation entre la direction des empreintes et la fonction des gènes, les données soutiennent pas tellement la théorie. Aucune donnée n'était disponible sur les patrons de réallocation. Seul le point concernant le timing d'expression des traits correspond donc parfaitement à l'attendu de la théorie du conflit d'allocation.

L'étape suivante est de confronter les données aux autres théories proposées dans la littérature, afin d'évaluer si ces théories apparaissent plus vraisemblables.

5.2 Les théories alternatives

La plupart des théories énoncées pour expliquer l'empreinte parentale ont été compilées et discutées par Hurst (1997). Cependant, les arguments de cette revue sont assez euthérien-centrés, et méritent d'être réétudiés en prenant en compte les spécificités des angiospermes.

Nous nous concentrerons sur les théories qui permettent d'expliquer un tant soit peu les données pour les angiospermes. Les théories que nous avons écartées sont celles qui sont propres aux euthériens et celles écartées par Hurst (1997). Nous avons abordé les hypothèses relatives à l'hétérozygotie chapitre 2. Nous discuterons ici, en plus de ces hypothèses, des hypothèses d'évitement de la parthénogénèse (Solter 1987, 1988; Chandra & Nanjundiah 1990), de régulation du développement, de contrôle de la différentiation cellulaire (Holliday 1990), de tamponnage des effets dose (Jullien & Berger 2009), de minimisation de la variance dans les taux d'expression (Solter 1988), de l'espacement temporel (Hall 1990), et de régulation des hybridations (Birchler *et al.* 2003).

5.2.1 Théorie ayant trait à l'hétérozygotie

La théorie de l'hétérozygotie prédit que l'empreinte devrait évoluer sur des locus dont les mutations délétères sont dominantes. Il est difficile de tester cette hypothèse, car les allèles des gènes à empreinte sont actuellement exprimés de manière haploïde. Lorsqu'au locus *MEA* le mutant *mea* est hérité maternellement, mais que l'allèle paternel est fonctionnel *et* exprimé (ce qui est possible si l'empreinte est dérégulée, par exemple chez un mutant *ddm1/ddm1*), le phénotype de la graine n'est pas complètement normal. Cela suggère que la mutation *mea* n'est pas complètement récessive, bien que d'autres interprétations soient possibles.

Une autre possibilité est que l'empreinte soit le produit d'un conflit mâle-femelle sur le niveau d'expression des allèles délétères. Cette hypothèse requiert que l'extinction maternelle soit sous contrôle paternel, et inversement. Or, comme on l'a vu section 5.1.1, ce n'est pas le cas général.

Cette théorie n'apparait donc pas comme une explication satisfaisante. Cependant, comme on l'a vu chapitre 2, les effets liés aux changements d'hétérozygoties devraient permettre de mieux comprendre les cas de figure dans lesquels l'empreinte parentale n'évolue pas (par exemple $IGF2R$ n'est pas sous empreinte chez l'homme) ou ne correspond pas complètement aux prédictions des théories avancées (par exemple la théorie du conflit prédit que l'allèle sous-exprimé devrait être complètement éteint, ce qui n'est pas toujours le cas, voir par exemple Matsuoka *et al.* 1996, et que tous les allèles du locus devraient être sous empreinte, ce qui n'est pas non plus toujours le cas, voir Table 1.1).

Encadré 5.4 – Mise en place de l'apomixie chez les angiospermes

Dans le cas classique d'une plante diploïde à gamétophyte Polygonum, pour obtenir un embryon diploïde par apomixie, il faut que la réduction méiotique n'ait pas lieu. La conséquence est la production de noyaux polaires eux-aussi chacun diploïdes, et donc d'un albumen contenant quatre contributions maternelles. Après pseudo-fécondation (en général, la cellule centrale est fécondée mais pas l'oosphère), l'albumen est donc $4m : 1p$ ce qui ne correspond pas au ratio 2 : 1 requis pour le développement de ce tissu.
Les angiospermes qui pratiquent l'apomixie avec pseudo-fécondation ont mis en place deux stratégies de contournement de ce problème. La première possibilité est l'altération dans le développement du gamétophyte ou dans le processus de fécondation. Dans le premier cas, une mitose est supprimée, si bien que le gamétophyte n'est plus formé que de quatre cellules (un module *sensu* Friedman *et al.* 2008) et un seul noyau polaire diploïde est donc contenu par la cellule centrale (Raghavan 2006, p.175). Dans ce cas, l'albumen est $2m : 1p$. Dans le second cas, les deux cellules spermatiques fécondent la cellule centrale (chez *Ranunculus auricomus*, Köhler & Grossniklaus 2005). Dans ce cas, l'albumen est $4m : 2p$. La deuxième possibilité est le changement de sensibilité aux changements de ratios des génomes maternel et paternel. Ainsi, chez *Tripsacum dactyloïdes* et *Paspalum* spp., le ratio est modifié mais il n'y a pas d'impact négatif sur le développement (Köhler & Grossniklaus 2005).
Chez la plupart des Asteracées apomictiques, le développement peut être autonome, et avoir lieu sans la fécondation de l'albumen qui est donc 2 :0 (Köhler & Grossniklaus 2005). Les mécanismes qui diminuent la sensibilité du génome aux doses est inconnu. On est parvenu a le faire expérimentalement pour *A. thaliana* : il faut une mutation maternelle *ddm1* (hypométhylation) et un mutant de la classe fis (*fertilization independant seed development,* Vinkenoog 2000).

5.2.2 L'hypothèse d'évitement de la parthénogénèse

La parthénogénèse est supposée être un cul de sac évolutif, conduisant à l'extinction (voir par exemple Asher Jr 1970). Or l'empreinte parentale à probablement pour conséquence d'empêcher la parthénogénèse, car tous les gènes nécessaires au développement ne sont pas exprimés par un seul type de gamète (mâle ou femelle). C'est pourquoi il a été proposé que l'empreinte ait évolué parce qu'elle permet d'éviter la parthénogénèse (Solter 1987, 1988; Chandra & Nanjundiah 1990).

Cependant, un tel argument suppose qu'un trait dont les effets bénéfiques ont lieu à très long terme puisse évoluer. Cela est très improbable (voir par exemple Spencer 2000), d'autant plus vu les coûts associés à l'empreinte (perte de l'avantage hétérozygote, mais aussi impossibilité totale de se reproduire pour les femelles occasionnelles n'ayant pas la capacité de se reproduire sexuellement, Hurst 1997). Enfin, il existe des angiospermes parthénogénétiques (Köhler & Grossniklaus 2005, encadré 5.4). Cette théorie n'est donc pas non plus satisfaisante.

5.2.3 L'hypothèse de régulation du développement

L'empreinte entraîne l'avortement précoce de structures qui se développent sans fécondation. D'après certains auteurs, chez les mammifères, l'empreinte permet donc d'éviter le développement autonome et l'utilisation de ressources par une structure non viable. Ainsi, la régulation différentielle du développement par les gènes d'origine mâle et femelle devrait permettre d'éviter les maladies trophoblastiques de l'ovaire (Varmuza & Mann 1994), ou les placentas invasifs (Hall 1990) [1].

D'après nous, cette théorie peut être étendue pour les angiospermes. Chez ce groupe, l'empreinte permet d'éviter le développement autonome de l'albumen, donc l'utilisation de ressources par un gamétophyte dont la fécondation future n'est pas garantie. En effet, c'est bien l'expression gamétophytique de gènes sous empreinte dans l'embryon, les *FIS*, qui inhibent le développement précoce de l'albumen chez les angiospermes. Pour les angiospermes, l'empreinte a le contrôle de l'initiation du développement (c'est pourquoi nous appelons ci-après la version de cette hypothèse appliquée aux angiospermes "l'hypothèse de contrôle de l'initiation du développement").

5.2.3.1 Observations expliquées par la théorie

D'après Hurst (1997) (dont, pour rappel, la revue des données est assez euthériens-centrés), l'hypothèse de régulation du développement :

- ne peut expliquer l'occurrence d'extinctions dans les deux sexes ;
- n'explique pas la correspondance entre la direction des effets sur la croissance et la direction de l'empreinte ;
- n'explique pas la viabilité de souris pour lesquelles à la fois *IGF2* et *IGF2R* sont délétées ;
- n'explique pas pourquoi plus d'un gène est à empreinte (selon Hurst, 1997, il pourrait suffire d'éteindre un gène vital pour le développement pour obtenir le même patron).

Nous trouvons au contraire que la théorie explique très bien ces observations. On peut expliquer l'occurrence d'extinctions dans les deux sexes d'une part et la correspondance entre la direction des effets sur la croissance et la direction de l'empreinte d'autre part par l'idée que c'est l'équilibre subtil entre les doses d'inhibiteurs maternels et d'activateurs paternels qui permet d'initier le développement une fois que la fécondation a eu lieu. Si les activateurs étaient maternellement actifs, les inhibiteurs ne permettraient pas d'enrayer le développement autonome des tissus maternels. Si les inhibiteurs étaient paternellement actifs, la dose d'activateurs apportée par la contribution paternelle pourrait ne pas suffire à l'initiation du développement. Le rôle du génome paternel dans l'initiation du développement est clairement démontré chez *A. thaliana*. Nowack *et al.* (2006) montrent que la cellule centrale mutée *fis1* initie le développement de l'albumen avec une probabilité de 11% cinq jour après maturité de la fleur, et perd des ressources. Si l'oosphère est fécondé (mais pas l'albumen), la probabilité que l'albumen initie son développement est de 92%. Le développement de l'ovule est donc dépendant de deux signaux : un signal positif qui provient de la fécondation de l'oosphère, et le relâchement d'un signal négatif régulé par l'action des gènes de la classe *FIS*.

La viabilité de souris dont IGF2 et IGF2R ont été délétées peut être facilement expliquée en suivant la même logique que dans le paragraphe qui précède. La fonction

1. Maladies qui pourraient être causées par une surexpression maternelle de gènes du développement embryonnaire.

d'IGF2R est de maintenir un état latent, celle d'IGF2 de contrebalancer IGF2R. Si aucun de ces gènes n'est présent, le développement a lieu normalement. Enfin, on peut expliquer pourquoi plus d'un gène est sous-empreinte par l'idée que plusieurs gènes différents ont le potentiel, s'ils sont exprimés au mauvais moment, d'initier un gaspillage des ressources.

L'hypothèse de contrôle de l'initiation du développement est compatible avec deux autres faits importants. Premièrement, il est nécessaire pour que l'argument soit valide que les ressources puissent être réallouées. Cela n'est possible que pour des espèces pour lesquelles au moins une partie des ressources est allouée après la fécondation, ce qui est le cas des marsupiaux, euthériens, et angiospermes. Deuxièmement, l'hypothèse est compatible avec le fait que l'empreinte semble se réduire lorsque l'approvisionnement en ressources s'arrête (chez les angiospermes, c'est au moins le cas pour *MEG1*, Raghavan 2006 p.180, et *FIE2*, Gutierrez-Marcos 2009).

Enfin, l'hypothèse du signal pour l'initiation du développement permet d'expliquer pourquoi il est difficile de produire des individus apomictiques, et chez les plantes, des albumens dont le ratio n'est pas 2 : 1 : pour que les gènes exprimés par le père puissent contrebalancer les gènes exprimés par la mère, il faut que ces gènes génèrent des effets seuils, très sensibles aux effets de dosage.

5.2.3.2 Points faibles de la théorie

S'il est largement admis que les gènes *FIS* interviennent dans la régulation du développement de l'albumen (Voir par exemple Nowack *et al.* 2006; Spillane *et al.* 2000; Sorensen *et al.* 2001; Köhler *et al.* 2003), on peut se demander si les autres gènes qui sont à empreinte ont également cette fonction ou une fonction liée à l'accumulation des ressources.

Comme on l'a vu section 5.1.3, tous les gènes à empreinte ne semblent pas directement impliqués dans l'accumulation des ressources. Cependant, l'hypothèse de contrôle de l'initiation du développement est plus compatible avec les données qui concernent la fonction des gènes que la théorie du conflit sur l'allocation des ressources. En effet, la théorie du conflit prédit que seuls les gènes qui ont un effet sur la croissance doivent être sous empreinte, alors que la théorie de la régulation du développement est compatible le fait que des gènes à empreinte soient des régulateurs d'autres gènes et soient impliqués dans la différentiation de l'albumen.

Globalement, il n'existe donc aucun élément qui soit réellement incompatible avec cette théorie, et la plupart des données sont expliquées.

5.2.4 L'hypothèse de différentiation cellulaire

L'hypothèse de différentiation cellulaire (Berger & Chaudhury 2009; Berger 2010; Messing & Grossniklaus 1999; Baroux *et al.* 2002) est spécifique aux angiospermes. L'empreinte parentale est une différence de méthylation liée au regroupement de deux génomes, une cellule spermatique et un gamète femelle (la cellule centrale) dont les profils de méthylation sont différents. En général les marques épigénétique telles que la méthylation ont pour fonction la différentiation cellulaire (Surani 2001).

De plus, dans le cas des angiospermes, il est admis que les deux cellules spermatiques sont autant aptes l'une que l'autre à féconder l'oosphère ou la cellule centrale (voir par exemple Ingouff *et al.* 2009), à part chez Plumbago (Russell 1985). D'un point de vue épi-

génétique, elles semblent exprimer, chez *A. thaliana*, un même répertoire de gènes (Berger *et al.* 2008).

Or, les deux cellules spermatiques fécondent deux cellules aux destins bien différents : l'une des fécondations est censée produire un embryon tandis que l'autre est censée produire un tissu nourricier. Si les cellules spermatiques sont équivalentes, il est nécessaire que les cellules du gamétophyte femelle soient, elles, différentes l'une de l'autre pour que l'albumen soit différencié de l'embryon. Dans ce contexte, le profil de méthylation des cellules spermatiques ne peut correspondre, au mieux, qu'à celui d'un des deux gamètes femelles fécondés.

Du point de vue de cette hypothèse, l'empreinte n'est pas une innovation, qui joue un rôle particulier, mais une conséquence nécessaire de la double-fécondation. Un effet de bord.

Si cette hypothèse est valide, on s'attend à ce que les allèles d'un même gène hérités par le père soient exprimés de la même manière (tous les deux actifs ou tous les deux inactifs) dans l'embryon et l'albumen, et à ce que les allèles qui sont hérités par la mère soient exprimés de manière différente (l'un actif et l'autre inactif) dans l'embryon et l'albumen. Les informations nécessaires pour tester cette hypothèse sont données Table 5.1. Nous en faisons la synthèse Table 5.2.

Huit des gènes pour lesquels nous disposons d'une information fiable valident l'hypothèse, quand un seul (*Mez1*) la rejette.

Plusieurs interprétations sont possible en ce qui concerne *Mez1*. Soit les cellules spermatiques ne sont en réalité pas équivalentes (elles différent par leur profil épigénétique), soit elles sont reprogrammées post-fécondation. Plusieurs éléments indiquent qu'elles pourraient être reprogrammées. Par exemple, chez *Arabidopsis*, les copies paternelles de l'histone H3.3 ne sont pas régulées de la même manière dans l'embryon et l'albumen (Ingouff *et al.* 2007). Par ailleurs, le génome semble ne commencer à s'exprimer qu'au bout de 3 ou 4 jours après la fécondation, autant dans l'embryon que dans l'albumen (Vielle-Calzada *et al.* 2000). Ce délai pourrait être celui nécessaire à la reprogrammation du génome paternel (Pillot *et al.* 2010), en fonction de la cellule d'arrivée. Il se pourrait également qu'une partie du patron de méthylation des cellules spermatiques soit défini alors que les cellules spermatiques arrivent a proximité de l'oosphère et de la cellule centrale, par les cellules végétatives qui supportent la croissance du tube pollinique et apportent les cellules spermatiques au gamétophyte femelle. En effet, ces cellules végétatives subissent une reprogrammation épigénétique rapide, dont on pense qu'elle pourrait diriger des changements épigénétiques dans les cellules spermatiques via la production de siRNAs (Slotkin *et al.* 2009).

Notons que si les cellules spermatiques sont réellement reprogrammées, et si cette reprogrammation est sous contrôle maternel, ce qui est probable puisqu'elle est dépendante du gamète femelle (oosphère ou cellule centrale), l'hypothèse du conflit sur l'allocation des ressources ne saurait être une explication satisfaisante de l'empreinte parentale.

Si l'hypothèse de différentiation cellulaire était auto-suffisante, on s'attendrait à ce que plusieurs milliers de gènes soient à empreinte (F. Berger, comm. pers.), ce qui n'est pas le cas (voir section 1.1.4). D'après F. Berger (comm. pers.), il faut donc faire appel à l'hypothèse que différentes pressions de sélection maintiennent ou non l'empreinte parentale pour expliquer que les empreintes soient si peu nombreuses. D'après les modèles (chapitre 2), l'empreinte parentale ne devrait pas évoluer dans les conditions qu'on connait (la plupart des mutations délétères sont récessives). De ce fait, l'évolution de l'empreinte parentale représente un coût en termes d'hétérozygotie. On peut imaginer que c'est ce coût qui ait conduit à la réduction maximale du nombre d'empreintes, et

Espèce	Gène	Fiabilité	Expression embryonnaire	Expression dans l'albumen	Génomes paternels ont le même patronNA	Oosphère et cellule centrale ont le même patronNA	Valide l'hypothèseNA
A. thaliana	*FIS2*	+	Non	Maternelle	Oui	Non	Oui
A. thaliana	*FWA*	+	Non*	Maternelle	Oui	Non	Oui
A. thaliana , *A. lyrata*	*MEA*	+	Non	Maternelle	Oui	Non	Oui
A. thaliana	*MPC* (*At3g19350*)	+	Très faible (Biallélique)	Maternelle	NA	Oui	NA
A. thaliana	*PHE1*	+	Oui	Paternelle	Oui	Non	Oui
A. thaliana	*AtFH5*	+	Non	Maternelle	Oui	Non	Oui
A. thaliana	*AGL80*	-	Non	Maternelle	Oui	Non	Oui
A. thaliana	*HDG9*	-	Non	Maternelle	Oui	Non	Oui
A. thaliana	*HDG8*	-	Oui (stades précoces)	Maternelle	Non	Oui	Non
A. thaliana	*HDG3*	-	Oui	Paternelle	Oui	Non	Oui
A. thaliana	*ATMYBR2*	-	NA	Maternelle	NA		NA
A. thaliana	*AT5G62110*	-	NA	Paternelle	NA		NA
Zea mais	*FIE1*	+	Non	Maternelle	Oui	Non	Oui
Zea mais	*MEG1*	+	Non	Maternelle	Oui	Non	Oui
Zea mais	*Nrp1*	NA	Non	Maternelle	Oui	Non	Oui
Zea mais	*Mez1*	+	Oui (Biallélique)	Maternelle	Non	Oui	Non
Zea mais	*R* (empreinte allélique)	+	Non	Maternelle	Oui	Non	Oui
Zea mais	*dzr1* (empreinte allélique)	-	Non	NA	NA		NA

Table 5.2 – Éxpressions génétiques et hypothèse de différentiation cellulaire : comparaison de l'expression génétique des allèles d'origine mâle et femelle dans l'embryon et l'albumen. NA indique que la donnée n'est pas disponible.

que les empreintes restantes soient limitées au peu de gènes nécessaires pour contrôler en amont tous les patrons de différentiation cellulaires.

L'hypothèse de différentiation cellulaire est au minimum aussi explicative que la théorie du conflit sur l'allocation des ressources, et mérite donc d'être approfondie.

5.2.5 L'hypothèse de régulation des effets doses

L'hypothèse de régulation des effets doses suggère que l'empreinte a initialement évolué pour réguler des gènes dose-sensibles, dérégulés lors de l'évolution de la double fécondation et des changements dans le nombre de contributions maternelles (Dilkes & Comai 2004; Jullien & Berger 2009). Cette idée est compatible avec l'observation que l'empreinte parentale concerne essentiellement des gènes dont les effets-dose sont notables (voir par exemple Leblanc *et al.* 2002). Cependant, cette théorie n'explique pas davantage de données.

5.2.6 L'hypothèse de minimisation de la variance dans les taux d'expression

L'hypothèse de minimisation de la variance dans les taux d'expression suggère que l'empreinte parentale est un moyen d'affiner le contrôle, et, comme son nom l'indique, de diminuer la variance d'expression de gènes donnés (Solter 1988). D'après Hurst (1997), au moins deux raisons pourraient être envisagées. Premièrement, supposons qu'il y ait du bruit dans le contrôle de la transcription du gène, donc que la précision du contrôle soit mauvaise. Si un gène doit être exprimé à faible dose, et que l'extinction d'un gène est simple, il pourrait être plus précis d'éteindre l'un des gènes et de diminuer l'expression du second, par exemple de moitié, plutôt que de diminuer l'expression de deux gènes de trois-quarts. Deuxièmement, si le produit d'un gène est requis à haute dose, mais avec une variance minimale du taux d'expression, alors en assurant que l'un des allèles est exprimé, la transcription pourrait atteindre un taux moyen.

Cependant, on peut se demander pourquoi une précision importante est plus nécessaire pour certains gènes que pour d'autres. D'après Hurst (1997), une précision importante serait requise lors du développement embryonnaire essentiellement pour des raisons de valeur sélective inclusive, un développement excessif de l'embryon risquant de diminuer la survie de la mère et des autres embryons. Les points expliqués par cette théorie sont presque exactement les mêmes que ceux expliqués par la théorie du conflit d'allocation (Hurst, 1997). Les points qui diffèrent sont que :

- ce modèle est compatible avec un taux d'évolution normal des locus à empreinte ;
- ce modèle n'explique pas la correspondance entre la direction de l'empreinte et la direction de l'effet sur la croissance (ce qui n'est pas incompatible avec les données chez les angiospermes).

5.2.7 L'hypothèse de la régulation des hybridations

L'hypothèse de la régulation des hybridations (Birchler *et al.* 2003) suggère que les empreintes pourraient être une barrière à l'hybridation entre plantes d'une même espèce mais de ploïdies différentes. En tant que barrière à la reproduction, l'empreinte est d'après cette théorie un mécanisme adaptatif impliqué dans les événements de spéciation. Cette

théorie est compatible avec le fait qu'il est très difficile de croiser, pour la plupart des espèces, des plantes de ploïdies différentes.

5.3 Discussion

Deux hypothèses, l'hypothèse du signal pour l'initiation du développement, et l'hypothèse de différentiation cellulaire se démarquent dans leur capacité d'explication des données. Chez les mammifères, comme on l'a précisé, c'est le conflit sur l'allocation des ressources qui est la meilleure explication de l'empreinte parentale (Hurst 1997). Sachant que les empreintes parentales des angiospermes et des mammifères sont convergentes, la question se pose de savoir s'il est pertinent de vouloir les expliquer par la même théorie.

En fait, on rassemble souvent les deux empreintes sous le même chef parce qu'elles concernent toutes les deux des gènes exprimés au moment de l'allocation des ressources, chez des espèces qui allouent les ressources après la fécondation, ce qui correspond aux conditions nécessaires pour l'évolution d'un conflit. Mais si on regarde de plus près, on se rend compte que les empreintes des mammifères et des angiospermes présentent aussi quelques différences (gènes généralement actifs dans les tissus somatiques chez les mammifère, *versus* éteints chez les angiospermes, l'empreinte a lieu dans l'embryon chez les mammifères, pas chez les plantes, la parthénogénèse n'a pas été enregistrée à l'état naturel chez les mammifères, mais l'a été chez les angiospermes, chez les animaux l'initiation du développement est sous le contrôle de l'arrêt de méiose, et pas chez les angiospermes puisqu'il y a un stade gamétophytique).

Il est important de se demander si ces deux innovations, convergentes, doivent trouver leur explication dans la même théorie, quitte à ce que tous les faits ne concordent pas, ou bien s'il vaut mieux que chacune trouve son explication dans une théorie différente, ce qui devrait permettre d'améliorer pour chacune la concordance des faits. En l'occurrence, l'adéquation de la théorie du conflit sur l'allocation des ressources avec les faits qui concernent l'empreinte parentale chez les angiospermes est tellement faible, qu'il parait plus raisonnable d'envisager une autre théorie. L'apport de données supplémentaires, notamment concernant l'empreinte chez d'autres espèces telles que le maïs ou le riz, est nécessaire pour asseoir davantage les deux hypothèses que sont l'hypothèse du signal pour l'initiation du développement et l'hypothèse de différentiation cellulaire, mais nous pensons que ces hypothèses sont de bonnes candidates pour expliquer l'évolution de l'empreinte chez les angiospermes.

Si l'empreinte parentale des angiospermes n'est pas l'objet d'un conflit sur l'allocation, il est probable que le conflit sur l'allocation n'exerce pas non plus d'importantes pressions de sélection sur la double fécondation, le doublement des contributions maternelles, et la polysporie. De ce fait, cette revue de littérature renforce l'idée qu'il est important d'étudier les pressions de sélection qui portent sur ces traits en dehors d'un éventuel conflit sur l'allocation des ressources.

Chapitre 6

Conclusion générale

Les tissus nourriciers de la graine présentent une grande diversité de formes. Cette diversité est en partie caractérisée par une diversité de constitutions génétiques : le tissu nourricier de la graine des gymnospermes est haploïde et maternel, alors qu'un tissu nourricier d'une graine angiosperme est issu de la double fécondation (albumen), ou formé de tissus du sporophyte maternel (périsperme) ou embryonnaire (cotylédons). Ce dernier est donc au minimum diploïde et il est généralement biparental chez les angiospermes.

Plusieurs types de pressions de sélection sont supposément sous-jacents à cette diversité. La théorie du conflit sur l'allocation des ressources stipule que les tissus nourriciers des angiospermes sont l'objet d'un conflit mâle-femelle sur l'allocation des ressources, parce que les angiospermes présentent la particularité, comparées aux gymnospermes, d'allouer une importante partie des ressources à la graine postérieurement à la fécondation. Dans le cadre de cette théorie, la double fécondation est supposée être une stratégie paternelle dont l'effet serait de biaiser l'allocation des ressources, et le doublement des contributions maternelles et la polysporie sont vues comme des réponses à la double fécondation, dont les effets seraient de rééquilibrer l'allocation des ressources entre les embryons. La théorie du conflit sur l'allocation suggère également qu'un autre trait, l'empreinte parentale, évolue parce que les gènes d'origine paternelle sous-expriment les inhibiteurs de croissance du tissu nourricier, ce qui aurait pour effet de biaiser l'allocation des ressources, et que les gènes d'origine maternelle, en réponse, sous-expriment les activateurs de croissance du tissu nourricier, ce qui aurait pour effet de rééquilibrer l'allocation des ressources, et ce jusqu'à ce que tous les inhibiteurs et les activateurs de croissance aient une expression exclusivement maternelle et exclusivement paternelle, respectivement.

Bien que l'empreinte soit *a priori* une innovation très différente de la double fécondation ou d'un changement dans le nombre de contributions maternelle, l'ensemble de ces traits ont pour point commun de changer le ratio d'expression des contributions maternelles et paternelles, à la différence près que l'évolution de l'empreinte est restreinte à un locus particulier, alors que les changements dans le nombre de contributions parentales concernent tout le génome.

La théorie de l'hétérozygotie suggère que les changements dans les ratios d'expression des contributions maternelles et paternelles changent le niveau de masquage des allèles délétères. Cela peut être avantageux si le changement conduit à un masquage accru des allèles délétères, mais a pour conséquence l'accumulation de mutations délétères dans la population. Par conséquent, les changements dans les niveaux d'expression des contributions maternelles et paternelles peuvent ou non favoriser l'évolution d'un tissu nourricier

donné.

L'objectif principal de cette thèse était évaluer dans quelle mesure les théories du conflit sur l'allocation des ressources et de l'hétérozygotie sont capables d'expliquer l'évolution des tissus nourriciers de la graine angiosperme. La synthèse des prédictions et observations apportées par la littérature et cette thèse permet cette évaluation et ouvre plusieurs perspectives.

6.1 Évaluation des théories

Comme nous l'avons vu chapitre 5, en ce qui concerne l'empreinte parentale chez les angiospermes, aucune des prédictions de la théorie du conflit sur l'allocation des ressources (à part la prédiction qui concerne le timing de l'allocation) n'est réellement vérifiée par les données. Cela permet de questionner l'importance accordée à la théorie du conflit sur l'allocation des ressources, pas seulement pour l'évolution de l'empreinte parentale chez les angiospermes, mais également pour l'évolution de la double fécondation, du doublement des contributions maternelles, et de la polysporie. Ci-après, nous évaluons donc le pouvoir explicatif de la théorie du conflit sur l'allocation des ressources pour l'ensemble des traits qui sont supposés être sous son contrôle, et comparons ce pouvoir explicatif à celui de théories alternatives.

6.1.1 Évaluation de la théorie du conflit sur l'allocation des ressources

Une théorie satisfaisante doit permettre d'expliquer un grand nombre de données à elle seule, tout en ne laissant inexpliquées que le moins de données possible.

Un grand nombre de traits sont susceptibles d'être expliqués par la théorie du conflit sur l'allocation des ressources. En effet, en plus de l'évolution de la double fécondation, des contributions maternelles, et de l'empreinte parentale (chez les angiospermes mais aussi chez les mammifères), le conflit sur l'allocation des ressources est supposé être responsable de l'évolution :

- des téguments sporophytiques qui pourraient être impliqués dans le contrôle maternel de la taille de la graine (Westoby & Rice 1982) ;
- des septum dans les fruits (qui séparent physiquement les graines, Uma Shaanker *et al.* 1988) ;
- de l'hypostase (tissu lignifié au niveau du chalaze, c'est-à-dire au niveau de la zone où les ressources sont transférées à l'ovule, Westoby & Rice 1982) ;
- des haustoria (des structures embryonnaires qui envahissent le tissu sporophytique, Queller 1983; Bhatnagar & Kallarackal 1980) ;
- du "strike" (tendance des noyaux polaires à devenir non-fonctionnels dans la cellule centrale, Haig 1986) ;
- de la surproduction d'ovules et la polyembryonie (supposés êtres des adaptations maternelles permettant l'avortement sélectif, Sakai 2007; Uma Shaanker & Ganeshaiah 1996) ;
- des patrons d'expression des hormones (Queller 1983; Ravishankar *et al.* 1995).

Dans le chapitre 5, nous avons inventorié plusieurs critères auxquels devaient répondre les empreinte parentales pour valider la théorie du conflit sur l'allocation. Il est possible de généraliser ces critères. Si les traits précités sont dus à un conflit sur l'allocation des

ressources, on s'attend à observer les éléments suivants :

- ces traits sont l'objet d'une coévolution antagoniste ;
- ces traits ne doivent pas se maintenir au delà de la période d'investissement parental ;
- la suppression expérimentale d'un trait qui est sous contrôle paternel doit provoquer un excès de croissance quand la suppression expérimentale d'un trait qui est sous contrôle maternel doit entrainer un déficit de croissance ;
- ces traits doivent être d'autant plus fréquents et plus poussés que les conditions dans lesquelles le conflit est attendu sont remplies ;
- lorsque les ressources ne sont pas allouées à un embryon, elles doivent pouvoir être réallouées aux autres embryons (Charalambous *et al.* 2003).

Nous évaluons ci-après si les cinq critères énoncés ci-dessus sont respectés pour l'ensemble des traits précités (hormis l'empreinte parentale des angiospermes, puisque ce travail à été effectué chapitre 5).

6.1.1.1 Coevolution antagoniste

La double fécondation (DF) et le doublement des contributions maternelles (MCD) semblent être les deux premières étapes d'une coévolution antagoniste. Cependant, dans les étapes suivantes, le nombre de contributions maternelles augmente toujours (polysporie), alors que le nombre de contributions paternelles n'augmente plus. Chez les gymnospermes, il existe des grains de pollen contenant jusqu'à plus de dix cellules spermatiques (Willson & Burley 1983). Cela suggère qu'il n'y a pas de contrainte coté paternel à l'augmentation du nombre de cellules spermatiques fécondant la cellule centrale. Alternativement, une endomitose de la cellule spermatique aurait permis de doubler simplement le nombre de contributions paternelles. On comprend également mal pourquoi la femelle n'a pas éliminé les récepteurs nécessaires à la cellule spermatique pour féconder la cellule centrale, ou transféré la fonction d'acquisition des ressources vers une autre cellule du gamétophyte, plutôt que de doubler les contributions maternelles. En ce qui concerne le retour vers un contrôle mâle, une vision alternative est possible. Comme nous l'avons précisé dans l'introduction de la thèse, la graine exalbuminée est dérivée. Il est envisageable que cette étape soit un retour vers un ratio $1m : 1p$, c'est-à-dire l'étape suivant DF et MCD dans cette coévolution antagoniste. Ce scénario n'est pas très satisfaisant pour au moins deux raisons : i) les cotylédons ne remplacent pas l'albumen au point que celui-ci disparaisse, et ii) la polysporie est difficilement intégrable en tant que 4e étape dans ce scénario. Globalement, ces données n'apportent donc pas un support très solide à la théorie du conflit.

Beaucoup de gènes ont une empreinte parentale chez les mammifères. Il y a eu coévolution antagoniste entre *IGF2* et *IGF2R* chez la souris, conduisant à l'extinction totale d'*IGF2* coté maternel, et d'*IGF2R* coté paternel. Ceci est l'argument principal du conflit sur l'allocation des ressources. Cependant, plusieurs éléments vont plutôt à l'encontre d'un patron de coévolution antagoniste. L'analyse du taux d'évolution des séquences de plusieurs gènes à empreinte ne permet pas de mettre en évidence l'occurrence d'un taux aussi important que celui attendu (McVean & Hurst 1997; Hurst & McVean 1998). Par ailleurs, certaines empreintes ne sont pas poussées au maximum (l'allèle sous-exprimé n'est pas toujours complètement inactif) et sont dépendantes des allèles présent aux différents locus, ce qui ne correspond pas aux prédictions de la théorie du conflit (Mochizuki *et al.* 1996).

Les patrons d'expression hormonaux semblent présenter une coévolution antagoniste. D'après les éléments disponibles, les hormones impliquées dans l'augmentation de l'allocation des ressources (Acides gibbérelliques, Auxines) sont exprimées par les embryons, tandis que celles impliquées dans la régulation négative du flux des ressources (Acides abcissiques) sont exprimées par le sporophyte maternel (Ravishankar *et al.* 1995). Ceci est un patron tout à fait compatible avec la théorie du conflit.

Il n'y a pas, à notre connaissance, d'indice de coévolution antagoniste en ce qui concerne les autres traits énumérés ci-avant (septums, haustoria, strike, etc.). En ce qui concerne la prédiction de coévolution antagoniste, il y a donc des éléments allant en faveur, et des éléments allant en défaveur de la théorie.

6.1.1.2 Timing d'expression des traits

L'intégralité des traits que l'on a inventoriés chez les angiospermes sont exprimés dans la graine ou le fruit. Ils ne sont donc exprimés, conformément aux prédictions de la théorie, que pendant la période d'allocation des ressources.

Chez les mammifères, l'empreinte parentale est généralement restreinte à des gènes exprimés dans l'embryon est dans le placenta, ce qui correspond aux prédictions de la théorie. Cependant, il y a aussi des empreintes dans certains organes adultes (cerveau, foie, voir par exemple Keverne *et al.* 1996; Wang *et al.* 2004; De los Santos *et al.* 2000). Ces empreintes concernent des gènes impliqués par exemple dans le comportement, et sont donc susceptibles d'affecter l'allocation des ressources dans le cadre de la compétition intra-fratrie, après la naissance (Hurst 1997).

En ce qui concerne la prédiction concernant le timing d'expression des traits, les différents éléments vont donc en faveur de la théorie du conflit.

6.1.1.3 Effets parentaux

La théorie prédit que les traits qui sont sous contrôle paternel augmentent la croissance, tandis que ceux qui ceux sont contrôle maternels inhibent la croissance.

Seules les empreintes et les patrons d'expression hormonaux ont été étudiées pour ce critère. Comme on l'a vu les hormones activatrices de croissance sont exprimées par l'embryon, quand les hormones inhibitrices de croissance sont exprimées par la mère, ce qui correspond à l'attendu de la théorie.

La revue des effets parentaux des gènes à empreinte des mammifères à été faite par Hurst & McVean (1997). D'après ces auteurs, seulement certains gènes sont conformes aux prédictions, si bien que les données ne soutiennent globalement pas la théorie. Par ailleurs, tous les gènes impliqués dans la croissance chez les mammifères ne sont pas à empreinte (Hurst & McVean 1998). L'ensemble de ces éléments vont plutôt en défaveur de la théorie du conflit sur l'allocation des ressources.

6.1.1.4 Corrélation aux traits d'histoire de vie

Dans la littérature, plusieurs études s'intéressent à la relation entre phénotype de la graine et traits d'histoires de vie. Deux études présentent des résultats qui sont compatibles avec la théorie, mais pourraient être expliqués autrement.

De Jong *et al.* (2005) prédisent que le régime de reproduction doit être corrélé à la taille des graines, car l'optimum des embryons converge vers l'optimum plus faible du sporophyte en cas de régime autofécondant. Ils montrent par l'analyse des données d'une flore que conformément à leur prédiction, les graines autofécondantes ont des graines plus

petites, en moyenne, que les plantes allofécondantes. Cela est congruent avec la théorie du conflit, mais pourrait également être expliqué par une différence d'hétérosis par exemple.

D'après Uma Shaanker & Ganeshaiah (1997), il semblerait que chez plusieurs espèces (*Phaseolus latheroides*, *Epilobium*), le taux d'avortement des ovules augmente lorsque les embryons sont peu apparentés. Ces auteurs interprètent un taux d'avortement élevé comme étant le résultat de la compétition entre les embryons. Cependant, la diminution de l'apparentement diminue la variance de qualité génétique des graines, donc potentiellement l'écart de compétitivité. Dans ces conditions, on peut imaginer que moins de graines seraient susceptibles de passer en dessous du seuil de ressources provoquant l'avortement.

Une étude présente des résultats qui sont incompatibles avec la théorie du conflit. D'après Carman (1997), la polysporie, l'apomixie et la polyembryonie (production de plusieurs embryons pour un même ovule) se produisent plus souvent ensemble que ce qu'on attendrait par hasard. Il est compréhensible que la polyembryonie soit corrélée avec l'apomixie dans la mesure où dans les cas de polyembryonie, les embryons surnuméraire sont généralement apomictiques (Maheshwari 1963). Cependant, la théorie du conflit sur l'allocation prédit que la polysporie devrait être corrélée à des systèmes de reproduction sexués, allofécondants. Les observations de Carman (1997) ne sont donc pas compatibles avec l'explication avancée par la théorie du conflit d'allocation des ressources pour la polysporie.

Nous n'avons pas pu mettre en évidence, en raison d'une quantité insuffisante de données, d'éventuelles corrélations entre la présence d'effets paternels et les traits d'histoire de vie ou les différents types de tissus nourriciers. Cependant, ce type d'analyse devrait être informatif lorsque davantage de données seront disponibles.

Aucune de ces études n'apporte réellement d'éléments allant en faveur la théorie du conflit sur l'allocation des ressources. Le fait que la polysporie soit corrélé à l'apomixie est un point vraiment critique, qui remet véritablement en question les causes de l'évolution de la polysporie.

6.1.1.5 Réallocation

Les embryons doivent être en compétition pour qu'un conflit sur l'allocation des ressources ait lieu. Un certains nombre d'éléments, tels que la présence de haustoria, ou le mode d'expression des hormones, sont souvent considérés comme des indices de compétition. Cependant, il s'agit d'indices indirects. Sadras & Denison (2009) ont montré que les plantes utilisent souvent des mécanismes pouvant être interprétés comme des indices de compétition pour allouer les ressources d'une unité à l'autre (branches et leurs sous-parties), et ce en l'absence de conflit. Le cas de la dominance apicale est un exemple parlant. Le méristème apical émet des hormones qui inhibent le développement des méristèmes qui sont à proximité. Ainsi, ce qui peut apparaître comme des indices de compétitions ne sont pas nécessairement les indices d'un conflit.

Le seul moyen d'étudier le degré de compétition entre les embryons est donc d'étudier la réallocation, c'est-à-dire le fait que lorsqu'une ressource n'est pas acquise pas un embryon (ou la graine qui le contient), elle est acquise par d'autres. Notre étude sur les patrons de l'allocation des ressources chez le maïs est la seule à s'intéresser sérieusement à la réallocation. Nous montrons dans cette étude que l'allocation des ressources est l'objet d'une réelle compétition chez le maïs. Plus précisément, nous montrons que différents génotypes mâles ont des compétitivités différentes, et transitives.

Ce type d'étude devra être répliqué sur différentes espèces pour vérifier si ce patron, observé chez le maïs, est représentatif de l'ensemble des angiospermes. Cependant, notre

étude vérifie la prédiction du conflit sur l'allocation selon laquelle les embryons sont en compétition pour les ressources.

6.1.1.6 Discussion

Les observations qui concernent le timing de l'expression des traits et la réallocation soutiennent la théorie du conflit sur l'allocation des ressources. Les observations qui concernent la coévolution antagoniste, les effets parentaux, et la corrélation soutiennent en partie seulement cette théorie.

Les théories alternatives à la théorie du conflit sur l'allocation ont été moins explorées qu'elle, probablement du fait que ces théories sont trait-spécifiques. Dans cette thèse, nous avons particulièrement étudié la théorie de l'hétérozygotie, qui se veut explicative pour l'évolution de la double fécondation, du doublement des contributions maternelles, et la polysporie. Ci-après, nous confrontons ces théories alternatives aux données, afin de comparer leurs pouvoirs explicatifs à celui de la théorie du conflit sur l'allocation des ressources.

6.1.2 Évaluation de la théorie de l'hétérosis

Dans sa version initiale, la théorie de l'hétérosis prédit :

- que l'albumen est sujet à l'hétérosis ;
- que l'augmentation du nombre de contributions parentales différentes augmente l'hétérosis.

En ce qui concerne la première prédiction, Brink & Cooper (1940, 1947) montrent par la comparaison du développement de graines allo et autofécondées que les graines autofécondées interrompent leur développement plus souvent que les graines allo-fécondées, et que les changements physiologiques auxquels sont imputables cet arrêt du développement impliquent le sac embryonnaire et les tissus maternels qui y sont adjacents. Ils en concluent que l'hétérosis de l'albumen joue un rôle important dans le développement de la graine. Cependant, dans ces expériences, il existe des effets confondants car les embryons subissent également la dépression de consanguinité dans les croisements auto-fécondants.

Nos modèles théoriques nous ont permis d'affiner la seconde prédiction. On s'attend à ce que :

- la double fécondation évolue lorsque les mutations délétères sont globalement récessives, et ce d'autant plus facilement que la plante est auto-fécondante (autofécondation sporophytique) ;
- le doublement des contributions maternelles augmente l'hétérosis si l'exposition moyenne des mutations délétères dans l'albumen $2m : 1p$ (h_m) est inférieure à celle d'un albumen $1m : 1p$ (h). Ainsi, contrairement à ce qui prédisent Friedman *et al.* (2008), il n'est pas nécessaire que les contributions supplémentaires soient différentes pour modifier l'hétérosis.
- la polysporie peut évoluer lorsque les mutations délétères plutôt récessives en moyenne dans un tissu triploïde (h_m petit), et dépendent beaucoup de s ;
- la polysporie peut évoluer lorsque les mutations délétères plutôt récessives en moyenne dans un tissu diploïde (h petit), et dépendent beaucoup de s ;
- l'empreinte parentale évolue dans les conditions dans lesquelles la double fécondation n'évolue pas, c'est-à-dire pour des locus pour lesquels les mutations délétères sont dominantes.

D'après les données disponibles, la plupart des mutations délétères sont récessives (les valeurs de dominance vont généralement de 0.1 à 0.4, voir par exemple Garcia-Dorado & Caballero 2000; Vassilieva *et al.* 2000; Szafraniec *et al.* 2003). Cela suggère que la double fécondation a pu évoluer parce qu'elle augmente l'hétérosis de la graine.

Il n'existe pas de données qui permettent de savoir si les conditions sont réellement remplies pour que le doublement des contributions maternelles évolue en augmentant l'hétérosis. Il est probable que h_m soit proche de h, donc qu'il soit supérieur ou inférieur à sa valeur selon les taxons, ce qui pourrait expliquer le polymorphisme (il existe à la fois des tissus nourricier $2m : 1p$ et $1m : 1p$).

En ce qui concerne le périsperme et la polysporie, un modèle intégrant différents locus plus ou moins importants pour la viabilité (différents s) est nécessaire pour savoir si ces traits peuvent évoluer dans des conditions réalisable. Il est possible que ces conditions soient réalisées dans certains taxons et ne le soient pas dans d'autres, ce qui devrait, là encore, rendre compte du polymorphisme. Cependant, bien que nous n'ayons pas fait de modèle dans lesquels une fraction des embryons est apomictique, il est possible d'émettre, dans le cadre de la théorie de l'hétérosis, une autre hypothèse simple pour expliquer l'évolution de la polysporie : il est possible qu'elle optimise l'exposition des allèles délétères si les embryons sont apomictiques. En effet, les embryons apomictiques peuvent être formés de contributions m et/ou m'. La polysporie expose à la fois les allèles m et m', ce qui apparaît comme le meilleur moyen d'exposer à la sélection les allèles portés par ce type d'embryon.

Enfin, en ce qui concerne l'empreinte parentale, il est possible d'envisager qu'elle soit restreinte à de rares locus pour lesquels la plupart des mutations délétères étaient dominantes au moment où ces empreintes ont évolué, bien que cette hypothèse soit difficilement testable. Les prédictions de la théorie de l'hétérosis concernant l'empreinte parentale ne sont de toutes manières pas inutiles. La théorie du conflit sur l'allocation des ressources prédit que lorsque l'expression des allèles d'un locus est l'objet d'un conflit sur l'allocation, l'allèle sous-exprimé devrait être complètement éteint (Mochizuki *et al.* 1996). Cela n'est pas toujours le cas (voir par exemple Matsuoka *et al.* 1996). Par ailleurs, dans certains cas, le locus est ou non à empreinte selon l'allèle présent (voir Table 1.1), ce qui n'est pas prédit par la théorie du conflit sur l'allocation. Dans ces cas de figure, on pourrait supposer que l'empreinte est le résultat net de pressions de sélection opposées : d'une part le conflit sur l'allocation des ressources sélectionne un déséquilibre d'expression et ce quels que soient les allèles, et d'autre part les pressions liées à l'hétérozygotie rééquilibrent l'expression des deux copies, et ce d'autant plus que les allèles sont récessifs (d'où un résultat allèle-dépendant). C'est probablement lorsque la pression de sélection liée au conflit est faible que l'hétérozygotie conduit à ces résultats intermédiaires.

D'après nos modèles, l'une des conséquences de l'augmentation du nombre de contributions parentales dans le tissu nourricier est la diminution de la purge des allèles délétères. Or, lorsque les allèles délétères s'accumulent dans une population, on peut s'attendre à une dépression de consanguinité plus importante. Nous avons montré qu'il existe bien un effet du type de tissu nourricier sur la dépression de consanguinité (chapitre 3). Cet effet pourrait même être plus déterminant que l'effet du régime de reproduction.

Globalement, on peut considérer que la théorie de l'hétérozygotie doit au minimum venir compléter, dans ses prédictions, la théorie du conflit sur l'allocation des ressources. C'est une théorie qui n'a pas suffisamment été explorée, et mérite de l'être davantage.

6.1.3 Théories alternatives

Avant de conclure sur le pouvoir explicatif des théories du conflit sur l'allocation et de l'hétérosis, il est nécessaire d'évaluer le pouvoir explicatif des autres théories alternatives. Nous avons déjà effectué cette tâche chapitre 5 pour l'empreinte parentale, nous l'effectuons ici pour les autres traits.

6.1.3.1 DF, MCD, polysporie : théorie de l'effet dose

Stebbins (1974, 1976) suggère que l'augmentation de ploïdie qui résulte de la double fécondation, du doublement des contributions maternelle, et de la polysporie devrait augmenter le taux de transcription dans la cellule, ce qui, comme l'hétérozygotie, pourrait être avantageux dans un contexte d'accélération du développement (cf. chapitre d'introduction).

Si l'augmentation de la taille de la cellule qui résulte généralement de l'augmentation de ploïdie pourrait effectivement augmenter les capacités de stockage de celle-ci (comme le montre l'étude de cellules polyploïdes), cette théorie n'est pas compatible avec plusieurs observations. Premièrement l'endo-reduplication est un mécanisme beaucoup plus simple que la double fécondation ou la polysporie pour augmenter le taux de transcription. Par ailleurs, le taux d'expression n'est pas forcément proportionnel au nombre de copie fonctionnelles, comme en attestent les travaux d'Auger *et al.* (2005, Fig. 6.1). C'est probablement dû à des effets de régulation en *trans* et de la compensation de dosage (Baroux *et al.* 2002; Birchler 1993).

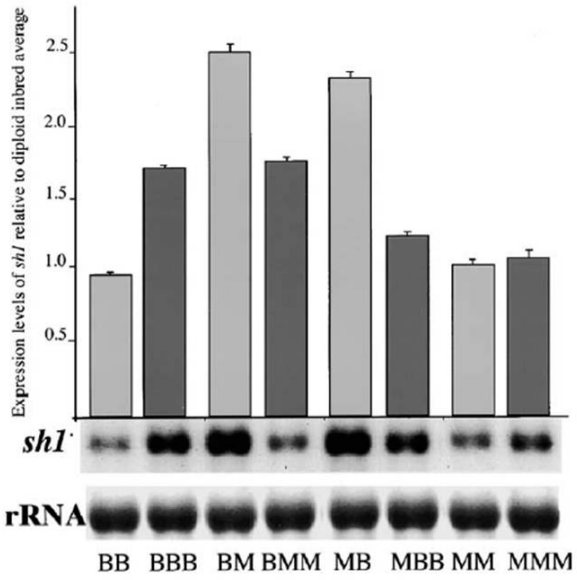

Figure 6.1 – Effets non additifs dans l'albumen triploïde. Les gènes de l'albumen, tels que *sh1* semblent avoir un taux d'expression moins important, notamment à l'état hétérozygote, dans un albumen triploïde que dans un albumen diploïde. Tiré de Auger *et al.* (2005).

6.1.3.2 DF : Optimisation de l'allocation des ressources

Pour, Westoby & Rice (1982) et Willson & Burley (1983) les théories de l'hétérozygotie et le conflit sur l'allocation pourraient être complémentaires. En effet, la double fécondation augmente l'exposition des allèles paternels à la sélection. De ce fait, la double fécondation pourrait être avantageuse d'un point de vue maternel s'il elle permet la sélection des embryons sur la base de leurs qualités génétiques.

Cette théorie requiert qu'au moins une partie des gènes exprimés dans l'albumen soient exprimés dans l'embryon, ce qui est probable.

6.1.3.3 Evolution des tissus exalbuminés et dormance

Forbis *et al.* (2002) montrent que l'évolution du ratio embryon : graine est corrélé à l'évolution de la dormance : plus le ratio est élevé (i.e. pour les embryons prennent de place à maturité dans la graine), plus la dormance est courte. D'après ces auteurs, plusieurs données indiquent que lorsque l'embryon occupe une place plus importante dans la graine, il est à un stade plus avancé de maturité, ce qui permettrait de réduire la dormance. Ces auteurs interprètent donc l'évolution de la tendance vers une graine exalbuminée comme étant le résultat d'une pression de sélection vers une dormance plus courte. Cette hypothèse parait plus plausible que l'hypothèse selon laquelle cette transition pourrait être une étape du conflit mâle-femelle.

6.1.4 Quelle théorie pour quel trait ?

Nous avons inventorié l'ensemble des théories et l'ensemble des traits relatifs à l'évolution des tissus nourriciers de la graine. Nous avons confronté les données aux théories pour chaque trait, et sommes désormais à même d'avoir une vision globale du pouvoir explicatif de chaque théorie.

Si l'on est peu parcimonieux, il existe une ou plusieurs théorie(s) satisfaisante(s) pour chacun des traits. L'empreinte parentale chez les mammifères trouve la meilleure explication dans la théorie du conflit sur l'allocation des ressources. L'empreinte parentale chez les angiospermes trouve les meilleures de ses explications dans les théories du contrôle de l'initiation du développement de l'albumen, et de différentiation cellulaire. La double fécondation trouve la meilleure de ses explications dans la théorie de l'hétérosis ou dans une théorie qui associe hétérosis et conflit (cf. discussion chapitre 2). Le doublement des contributions maternelles peut trouver une explication dans la théorie du conflit sur l'allocation ou dans la théorie de l'hétérosis. La polysporie trouve la meilleure de ses explications dans la théorie de l'hétérosis. Les patrons d'expression hormonaux, les haustoria et hypostases peuvent trouver explication dans la théorie du conflit, ou dans le fait que les plantes utilisent des modes d'allocation compétition-ressemblant. L'évolution de la graine exalbuminée trouve la meilleure de ses explications dans l'évolution de la dormance.

Ainsi, la théorie du conflit sur l'allocation des ressources, qui est actuellement la théorie qui prédomine dans la littérature, n'est la meilleure explication disponible qu'en ce qui concerne l'empreinte parentale chez les mammifères. La théorie de l'hétérosis revient aussi souvent que la théorie du conflit sur l'allocation des ressources, ce qui permet à nouveau d'affirmer que c'est une théorie qui mérite d'être davantage explorée.

6.2 Discussion et perspectives

L'objectif de cette thèse était d'identifier et de discuter les pressions de sélection auxquelles sont soumis les tissus nourriciers de la graine des angiospermes, en particulier celles proposées par la théorie du conflit sur l'allocation des ressources, et la théorie de l'hétérozygotie. L'étude de la littérature nous a permis d'identifier les principaux besoins. En ce qui concerne la théorie du conflit, il s'agissait de confronter les prédictions à d'avantage de données que ce qui avait déjà été fait. En ce qui concerne la théorie de l'hétérozygotie, il s'agissait en premier lieu d'établir des prédictions, mais la confrontation aux données apparaissait également nécessaire.

Notre travail s'ouvre sur plusieurs perspectives, à plusieurs niveaux. Des **perspectives théoriques** tout d'abord. Nous avons utilisé les outils de la génétique des populations, plus précisément la théorie des modificateurs, pour développer des prédictions théoriques pour la théorie de l'hétérozygotie. Ces outils supposent que des locus à effet majeur contrôlent les traits auxquels on s'intéresse et sont exprimés au cours du cycle de vie, et présentent l'avantage de reposer sur des mécanismes réalistes de la transmission des caractères (puisque la transmission de l'information génétique repose sur les lois de Mendel dans ces modèles), contrairement aux modèles phénotypiques. De ce fait, ils se prêtent très bien à l'étude des conflits génétiques intralocus : il nous a suffit de modéliser l'évolution des contributions parentales avec un contrôle par l'allèle mâle d'une part et un contrôle par l'allèle femelle d'autre part pour mettre en évidence un conflit. Par ailleurs, ces outils permettent d'étudier l'évolution des traits à long terme, contrairement aux modèles de génétique quantitative. La limite des modèles deux-locus est qu'ils ne permettent pas de prédire l'effet de nombreux locus, en particulier si les paramètres utilisés par le modèle (sélection, dominances, recombinaison) sont variables dans un génome. Une des perspectives, en termes de modélisation, et d'adopter une approche multi-locus, ce qui permettrait d'évaluer l'effet de cette variabilité sur les prédictions.

Notre travail ouvre également des **perspectives conceptuelles**. Du fait des changements dans les ratios d'expression des contributions maternelles et paternelles, l'expression génétique du tissu nourricier est rarement diploïde. Or, l'embryon, qui transmet les gènes qui sont exprimés dans la graine à la génération suivante, est, lui, diploïde, ce qui a le potentiel de générer un nouveau conflit : un conflit sur l'exposition des allèles délétères à la sélection. L'identification de ce conflit ouvre de nombreuse perspectives pour comprendre l'évolution des modes d'expression génétiques (tels que l'évolution de la régulation en *cis* ou en *trans*). Par ailleurs, si un tel conflit a lieu, on s'attend à ce que les changements d'expression des contributions parentales soient l'objet d'une coévolution antagoniste, en particulier lorsqu'ils sont liés à l'empreinte parentale : si plusieurs locus contrôlent une empreinte, mais sont à des distances différentes du gène à empreinte, le taux de recombinaison avec le locus à empreinte, donc les conditions d'évolution de l'empreinte, diffèrent pour ces locus modificateurs. La théorie du conflit sur l'allocation des ressources prédit également une coévolution antagoniste, mais elle n'est pas nécessairement identique à celle prédite par le conflit sur l'exposition des allèles. Une perspective est d'étudier les différences que l'on attend entre les deux types de coévolutions antagonistes.

En ce qui concerne les concepts, notre travail ouvre également des perspectives quant la question de l'évolution de la dépression de consanguinité. La plupart des travaux théoriques s'attachent à étudier l'effet du régime de reproduction sur la dépression de consanguinité. Or, les tissus nourriciers de la graine jouent un rôle déterminant dans le développement de l'embryon, avec lequel ils partagent des gènes. La prise en compte de la constitution génétique des tissus nourriciers devrait donc enrichir les réflexions portant

sur l'évolution des dépressions de consanguinité.

Ensuite, la quantité de ressources dans la graine est sujette à d'autres pressions de sélection que celles liées à l'accumulation des ressources nécessaires à l'embryon pour atteindre un stade de développement autotrophe : il s'agit des pressions liées à la dispersion, à la dormance, ou aux prédateurs par exemple. En ce qui concerne la dispersion, il est admis qu'elle est souvent l'objet d'un conflit mâle-femelle. La double fécondation, le doublement des contributions maternelles, la polysporie et l'empreinte pourraient donc également être le résultat de ce conflit là, plutôt que le résultat d'un conflit sur l'allocation des ressources.

Enfin, notre travail ouvre des **perspectives empiriques**. L'ensemble des prédictions de nos modèles reposent sur des paramètres de dominance des allèles délétères. Bien qu'il existe des données sur la dominance des allèles délétères dans un tissu diploïde, il n'existe pas de données sur la dominance des allèles délétères dans les tissus non diploïde. Lorsque ce type de données sera disponible, il sera possible d'évaluer plus concrètement les conditions de l'évolution des différents tissus nourriciers.

Par ailleurs, il serait intéressant d'avoir des mesures de dépression de consanguinité propres au tissu nourricier pour évaluer dans quelle mesure il est important de les prendre en compte dans l'étude théorique de la dépression de consanguinité. Il n'existe pas de mesure de la dépression de consanguinité qui soit strictement indépendante de la dépression de consanguinité des autres parties de la graine. Pour estimer la part de la dépression dans la graine liée à l'albumen, il faudrait cultiver *in vitro* des embryons avec et sans albumen.

En confrontant des théories alternatives aux théories du conflit sur l'allocation des ressources et de l'hétérozygotie aux données, nous avons montré que les données relatives à l'empreinte parentale, chez les angiospermes, pourraient être mieux expliqué par deux théories alternatives : la théorie du contrôle de l'initiation du développement de l'albumen, et la théorie de la différentiation cellulaire. L'exploration approfondie de ces théories devrait réellement permettre de comprendre l'évolution de l'empreinte parentale chez les angiospermes.

Bibliographie

Andersson, S. 1990. Paternal Effects on Seed Size in a Population of Crepis-Tectorum (Asteraceae). *Oikos*, **5**9(1), 3–8. 71

Antonovics, J., & Schmitt, J. 1986. Paternal and Maternal Effects on Propagule Size in Anthoxanthum-Odoratum. *Oecologia*, **6**9(2), 277–282. 71

Arathi, H. S., Ganeshaiah, K. N., Uma Shaanker, R., & Hegde, S. G. 1996. Factors affecting embryo abortion in Syzygium cuminii (L) Skeels (Myrtaceae). *International Journal of Plant Sciences*, **1**57(1), 49–52. 60

Arathi, H. S., Ganeshaiah, K. N., Uma Shaanker, R., & Hedge, S. G. 1999. Seed abortion in Pongamia pinnata (Fabaceae). *American Journal of Botany*, **8**6(5), 659–662. 60

Asher Jr, J.H. 1970. Parthenogenesis and genetic variability. II. One-locus models for various diploid populations. *Genetics*, **6**6(2), 369. 95

Auger, D. L., Gray, A. D., Ream, T. S., Kato, A., Coe, E. H., & Birchler, J. A. 2005. Nonadditive gene expression in diploid and triploid hybrids of maize. *Genetics*, **1**69(1), 389–397. 110

Banuelos, M. J., & Obeso, J. R. 2003. Maternal provisioning, sibling rivalry and seed mass variability in the dioecious shrub Rhamnus alpinus. *Evolutionary Ecology*, **1**7(1), 19–31. 71

Baroux, C., Spillane, C., & Grossniklaus, U. 2002. Genomic imprinting during seed development. *Advances in Genetics*, **4**6, 165–214. 97, 110

Baroux, U., Gagliardini, V., Page, D. R., & Grossniklaus, U. 2006. Dynamic regulatory interactions of Polycomb group genes : MEDEA autoregulation is required for imprinted gene expression in Arabidopsis. *Genes & Development*, **2**0(9), 1081–1086. 88, 92

Bawa, K. S., Hedge, S. G., Ganeshaiah, K. N., & Uma Shaanker, R. 1989. Embryo and Seed Abortion in Plants. *Nature*, **3**42(6250), 625–625. 60

Berger, F. 2003. Endosperm : the crossroad of seed development. *Current Opinion in Plant Biology*, **6**(1), 42–50. 63

Berger, F. 2010 (jully). *Parental genomic imprinting in flowering plants*. unpublish. 97

Berger, F., & Chaudhury, A. 2009. Parental memories shape seeds. *Trends in Plant Science*, **1**4(10), 550–556. 93, 97

Berger, F., Hamamura, Y., Ingouff, M., & Higashiyama, T. 2008. Double fertilization - caught in the act. *Trends in Plant Science*, **1**3(8), 437–443. 98

Bhatnagar, SP, & Kallarackal, J. 1980. Cytochemical studies on the endosperm of Linaria bipartita (Vent.) Willd. with a note on the role of endosperm haustoria. *Cytologia*, **4**5, 247–256. 60, 104

Biere, A. 1991. Parental Effects in Lychnis-Flos-Cuculi .1. Seed Size, Germination and Seedling Performance in a Controlled Environment. *Journal of Evolutionary Biology*, **4**(3), 447–465. 26, 71

Birchler, J. A. 1993. Dosage Analysis of Maize Endosperm Development. *Annual Review of Genetics*, **2**7, 181–204. 110

Birchler, James A., Auger, Donald L., & Riddle, Nicole C. 2003. In Search of the Molecular Basis of Heterosis. *Plant Cell*, **1**5(10), 2236–2239. 94, 100

Bix, M., & Locksley, R.M. 1998. Independent and epigenetic regulation of the interleukin-4 alleles in CD4+ T cells. *Science*, **2**81(5381), 1352. 9

Boisnard-Lorig, C, Colon-Carmona, A, Bauch, W, Hodge, S, Doerner, P, Bancharel, E, Dumas, C, Haseloff, J, & Berger, F. 2001. Dynamic analyses of the expression of the HISTONE : : YFP fusion protein in arabidopsis show that syncytial endosperm is divided in mitotic domains. *PLANT CELL*, **1**3(3), 495–509. 14

Borras, L., Slafer, G. A., & Otegui, M. E. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean : a quantitative reappraisal. *Field Crops Research*, **8**6(2-3), 131–146. 82

Brink, R. A. 1956. A Genetic Change Associated with the R Locus in Maize Which Is Directed and Potentially Reversible. *Genetics*, **4**1(6), 872–889. 10

Brink, R. A., & Cooper, D. C. 1940. Double fertilization and development of the seed in angiosperms. *Bot. Gaz.*, **1**02, 1–25. 32, 35, 108

Brink, RA, & Cooper, DC. 1947. The endosperm in seed development. *Botanical review*, **1**32, 423–541. 32, 108

Brocklehurst, P. A. 1977. Factors Controlling Grain Weight in Wheat. *Nature*, **2**66(5600), 348–349. 63, 91

Brown, R. C., Lemmon, B. E., Nguyen, H., & Olsen, O. A. 1999. Development of endosperm in Arabidopsis thaliana. *Sexual Plant Reproduction*, **1**2(1), 32–42. 14

Bulmer, MG. 1986. Genetic models of endosperm evolution in higher plants. *Pages 743–763 of :* Karlin, S., & Nevo, E (eds), *Evolutionary process and theory.* New York : Academic press. 30, 31

Burt, A., & Trivers, R. 2006. *Genes in conflict.* The Belknap press of Harvard university press. 26

Byers, D. L., Platenkamp, G. A. J., & Shaw, R. G. 1997. Variation in seed characters in Nemophila menziesii : Evidence of a genetic basis for maternal effect. *Evolution*, **5**1(5), 1445–1456. 71

Carman, J. G. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, **6**1(1), 51–94. 56, 59, 62, 107

Carmichael, J. S., & Friedman, W. E. 1995. Double fertilization in Gnetum gnemon : The relationship between the cell cycle and sexual reproduction. *Plant Cell*, **7**(12), 1975–1988. 7

Carmichael, J. S., & Friedman, W. E. 1996. Double fertilization in Gnetum gnemon (Gnetaceae) : Its bearing on the evolution of sexual reproduction within the Gnetales and the anthophyte clade. *American Journal of Botany*, **8**3(6), 767–780. 7

Chandra, H. S., & Nanjundiah, V. 1990. The Evolution of Genomic Imprinting. *Development*, 47–53. 94, 95

Charalambous, M., Ward, A., & Hurst, L. D. 2003. Evidence for a priming effect on maternal resource allocation : implications for interbrood competition. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **2**70, S100–S103. 86, 105

Charnov, E. L. 1979. Simultaneous Hermaphroditism and Sexual Selection. *Proceedings of the National Academy of Sciences of the United States of America*, **7**6(5), 2480–2484. 30

Chaudhuri, S., & Messing, J. 1994. Allele-Specific Parental Imprinting of Dzr1, a Post-transcriptional Regulator of Zein Accumulation. *Proceedings of the National Academy of Sciences of the United States of America*, **9**1(11), 4867–4871. 10, 88

Chaudhury, Abdul M., & Berger, Frederic. 2001. Maternal control of seed development. *Seminars in Cell & Developmental Biology*, **1**2(5), 381–386. 54, 91

Chaumeil, J., Le Baccon, P., Wutz, A., & Heard, E. 2006. A novel role for Xist RNA in the formation of a repressive nuclear compartment into which genes are recruited when silenced. *Genes & development*, **2**0(16), 2223. 55

Chess, A., Simon, I., Cedar, H., & Axel, R. 1994. Allelic inactivation regulates olfactory receptor gene expression. *Cell*, **7**8(5), 823–834. 9

Chin, HF. 1995. Germination and storage of banana seeds. *Pages 2–5 of : Proceedings of the workshop on New Frontiers in Resistance Breeding for Nematodes, Fusarium and Sigatoka*. 20

Choi, Y. H., Gehring, M., Johnson, L., Hannon, M., Harada, J. J., Goldberg, R. B., Jacobsen, S. E., & Fischer, R. L. 2002. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. *Cell*, **1**10(1), 33–42. 87, 88

Cooper, DC. 1951. Caryopsis development following matings between diploid and tetraploid strains of Zea mays. *American Journal of Botany*, **3**8(9), 702–708. 17

Cooper, DW, VandeBerg, JL, Sharman, GB, & Poole, WE. 1971. Phosphoglycerate kinase polymorphism in kangaroos provides further evidence for paternal X inactivation. *Nature : New biology*, **2**30(13), 155. 9

Coulter, J.M. 1911. The endosperm of angiosperms. *Bot. Gaz.*, **5**1, 329–348. 14

Crouse, H. V. 1960. The Controlling Element in Sex Chromosome Behavior in Sciara. *Genetics*, **4**5(10), 1429–1443. 9

Curtis, M. D., & Grossniklaus, U. 2008. Molecular control of autonomous embryo and endosperm development. *Sexual Plant Reproduction*, **2**1(1), 79–88. 91

Danilevskaya, O. N. 2003. Duplicated fie genes in maize : Expression pattern and imprinting suggest distinct functions. *Plant Cell*, **1**5(2), 425–438. 10, 88, 91

Davis, G. L. 1966. *Systematic embryology of angiosperms*. Wiley, J. 59, 62, 66

De Jong, T. J., Van Dijk, H., & Klinkhamer, P. G. L. 2005. Hamilton's rule, imprinting and parent-offspring conflict over seed mass in partially selfing plants. *Journal of Evolutionary Biology*, **1**8(3), 676–682. 30, 106

De los Santos, T., Schweizer, J., Rees, C.A., & Francke, U. 2000. Small evolutionarily conserved RNA, resembling C/D box small nucleolar RNA, is transcribed from PWCR1, a novel imprinted gene in the Prader-Willi deletion region, which is highly expressed in brain. *The American Journal of Human Genetics*, **6**7(5), 1067–1082. 106

Dickinson, Hugh, & Scott, Rod. 2002. DEMETER, Goddess of the Harvest, Activates Maternal MEDEA to Produce the Perfect Seed. *Molecular Cell*, **1**0(1), 5–7. 88

Dilkes, B. P., & Comai, L. 2004. A differential dosage hypothesis for parental effects in seed development. *Plant Cell*, **1**6(12), 3174–3180. 100

Donoghue, M. J., & Doyle, J. A. 2000. Seed plant phylogeny : Demise of the anthophyte hypothesis ? *Current Biology*, **1**0(3), R106–R109. 3

Donoghue, M.J., & Scheiner, S.M. 1992. The evolution of endosperm : a phylogenetic account. *Pages 356–389 of :* Wyatt, R. (ed), *In Ecology and Evolution of Plant Reproduction : New Approaches*. New York : Chapman and Hall. 32

Efstratiadis, A. 1994. Parental imprinting of autosomal mammalian genes. *Current Opinion in Genetics & Development*, **4**(2), 265–280. 9

Feil, R., & Berger, F. 2007. Convergent evolution of genomic imprinting in plants and mammals. *Trends in Genetics*, **2**3(4), 192–199. 10, 19

Fenster, C.B. 1991. Effect of male pollen donor and female seed parent on allocation of resources to developing seeds and fruit in Chamaecrista fasciculata (Leguminosae). *American journal of botany*, 13–23. 71

Fitzgerald, J., Luo, M., Chaudhury, A., & Berger, F. 2008. DNA Methylation Causes Predominant Maternal Controls of Plant Embryo Growth. *Plos One*, **3**(5). 54, 87

Fitzgerald, J. N. F., Hui, P. S., & Berger, F. 2009. Polycomb group-dependent imprinting of the actin regulator AtFH5 regulates morphogenesis in Arabidopsis thaliana. *Development*, **1**36(20), 3399–3404. 10, 88, 92

Floyd, S. K., & Friedman, W. E. 2000. Evolution of endosperm developmental patterns among basal flowering plants. *International Journal of Plant Sciences*, **1**61(6), S57–S81. 14

Forbis, T. A., Floyd, S. K., & de Queiroz, A. 2002. The evolution of embryo size in angiosperms and other seed plants : Implications for the evolution of seed dormancy. *Evolution*, **5**6(11), 2112–2125. 20, 21, 23, 59, 111

Friedman, W. E. 1990a. Double Fertilization in Ephedra, a Nonflowering Seed Plant - Its Bearing on the Origin of Angiosperms. *Science*, **2**47(4945), 951–954. 7

Friedman, W. E. 1990b. Sexual Reproduction in Ephedra-Nevadensis (Ephedraceae) - Further Evidence of Double Fertilization in a Nonflowering Seed Plant. *American Journal of Botany*, **7**7(12), 1582–1598. 7

Friedman, W. E. 1991. Double Fertilization in Ephedra-Trifurca, a Nonflowering Seed Plant - the Relationship between Fertilization Events and the Cell-Cycle. *Protoplasma*, **1**65(1-3), 106–120. 7

Friedman, W. E. 1992. Double Fertilization in Nonflowering Seed Plants and Its Relevance to the Origin of Flowering Plants. *International Review of Cytology-a Survey of Cell Biology*, **1**40, 319–355. 7

Friedman, W. E. 1995. Organismal Duplication, Inclusive Fitness Theory, and Altruism - Understanding the Evolution of Endosperm and the Angiosperm Reproductive Syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **9**2(9), 3913–3917. 30

Friedman, W. E. 2001a. Comparative embryology of basal angiosperms. *Current Opinion in Plant Biology*, **4**(1), 14–20. 14

Friedman, W. E. 2001b. Developmental and evolutionary hypotheses for the origin of double fertilization and endosperm. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences*, **3**24(6), 559–567. 14, 15

Friedman, W. E., & Ryerson, K. C. 2009. Reconstructing the Ancestral Female Gametophyte of Angiosperms : Insights from Amborella and Other Ancient Lineages of Flowering Plants. *American Journal of Botany*, **9**6(1), 129–143. 16, 17

Friedman, W. E., & Williams, J. H. 2003. Modularity of the angiosperm female gametophyte and its bearing on the early evolution of endosperm in flowering plants. *Evolution*, **5**7(2), 216–230. 16

Friedman, W. E., Madrid, E. N., & Williams, J. H. 2008. Origin of the fittest and survival of the fittest : Relating female gametophyte development to endosperm genetics. *International Journal of Plant Sciences*, **1**69(1), 79–92. 6, 32, 95, 108

Friedman, WE. 2008. Hydatellaceae are water lilies with gymnospermous tendencies. *NATURE*, **4**53(7191), 94–97. 20

Gachet, Sophie, Véla, Errol, & Tatoni, Thierry. 2005. BASECO : a floristic and ecological database of Mediterranean French flora. *Biodiversity and Conservation*, **1**4(4), 1023–1034. 59, 62

Ganeshaiah, K. N., & Uma Shaanker, R. 1988. Seed Abortion in Wind-Dispersed Pods of Dalbergia-Sissoo - Maternal Regulation or Sibling Rivalry. *Oecologia*, **77**(1), 135–139. 60

Ganeshaiah, K. N., & Uma Shaanker, R. 1992. Frequency-Distribution of Seed Number Per Fruit in Plants - a Consequence of the Self-Organizing Process. *Current Science*, **62**(4), 359–365. 27

Garcia, D., Fitz Gerald, J.N., & Berger, F. 2005. Maternal control of integument cell elongation and zygotic control of endosperm growth are coordinated to determine seed size in Arabidopsis. *The Plant Cell Online*, **17**(1), 52. 54

Garcia-Dorado, A., & Caballero, A. 2000. On the Average Coefficient of Dominance of Deleterious Spontaneous Mutations. *Genetics*, **155**, 1991–2001. 32, 40, 109

Gehring, M., Huh, J. H., Hsieh, T. F., Penterman, J., Choi, Y., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 2006. DEMETER DNA glycosylase establishes MEDEA polycomb gene self-imprinting by allele-specific demethylation. *Cell*, **124**(3), 495–506. 88, 92

Gehring, M., Bubb, K. L., & Henikoff, S. 2009. Extensive Demethylation of Repetitive Elements During Seed Development Underlies Gene Imprinting. *Science*, **324**(5933), 1447–1451. 10, 11

Gimelbrant, A., Hutchinson, J.N., Thompson, B.R., & Chess, A. 2007. Widespread monoallelic expression on human autosome. *Science*, **318**(5853), 1086–1089. 9, 11

Gimelbrant, A.A., Ensminger, A.W., Qi, P., Zucker, J., & Chess, A. 2005. Monoallelic expression and asynchronous replication of p120 catenin in mouse and human cells. *Journal of Biological Chemistry*, **280**(2), 1354. 9

Gould, SJ. 1977. *Ontogeny and phylogeny.* Cambridge : Belknap press. 25

Grimanelli, D., Perotti, E., Ramirez, J., & Leblanc, O. 2005. Timing of the maternal-to-zygotic transition during early seed development in maize. *Plant Cell*, **17**(4), 1061–1072. 54

Grossniklaus, U., & Schneitz, K. 1998. The molecular and genetic basis of ovule and megagametophyte development. *Seminars in Cell & Developmental Biology*, **9**(2), 227–238. 91

Grossniklaus, U., Vielle-Calzada, J. P., Hoeppner, M. A., & Gagliano, W. B. 1998. Maternal control of embryogenesis by medea, a Polycomb group gene in Arabidopsis. *Science*, **280**(5362), 446–450. 8, 10

Guignard, L. 1899. Sur les anthérozoides et la double copulation sexuelle chez les végétaux angiospermes. *C. R. Acad. Sci. Paris*, **128**, 864–871. 6

Guitton, A. E., Page, D. R., Chambrier, P., Lionnet, C., Faure, J. E., Grossniklaus, U., & Berger, F. 2004. Identification of new members of Fertilisation Independent Seed Polycomb Group pathway involved in the control of seed development in Arabidopsis thaliana. *Development*, **131**(12), 2971–2981. 92, 93

Guo, M., Rupe, M. A., Danilevskaya, O. N., Yang, X. F., & Hut, Z. H. 2003. Genome-wide mRNA profiling reveals heterochronic allelic variation and a new imprinted gene in hybrid maize endosperm. *Plant Journal*, **3**6(1), 30–44. 10, 11, 88

Gutierrez-Marcos, J. 2009. Imprinting in plants. *Chinese Science Bulletin*, **5**4(14), 2383–2389. 11, 97

Gutierrez-Marcos, J. F., Costa, L. M., Biderre-Petit, C., Khbaya, B., O'Sullivan, D. M., Wormald, M., Perez, P., & Dickinson, H. G. 2004. maternally expressed gene1 is a novel maize endosperm transfer cell-specific gene with a maternal parent-of-origin pattern of expression. *Plant Cell*, **1**6(5), 1288–1301. 10, 88, 92

Haig, D. 1986. Conflicts among Megaspores. *Journal of Theoretical Biology*, **1**23(4), 471–480. 28, 60, 104

Haig, D. 1990. Brood Reduction and Optimal Parental Investment When Offspring Differ in Quality. *American Naturalist*, **1**36(4), 550–556. 27, 83

Haig, D. 1997. Parental antagonism, relatedness asymmetries, and genomic imprinting. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **2**64(1388), 1657–1662. 31

Haig, D. 1999. Multiple paternity and genomic imprinting. *Genetics*, **1**51(3), 1229–1231. 31

Haig, D., & Westoby, M. 1989. Parent-Specific Gene-Expression and the Triploid Endosperm. *American Naturalist*, **1**34(1), 147–155. 30, 31

Haig, D., & Wilkins, J. F. 2000. Genomic imprinting, sibling solidarity and the logic of collective action. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences*, **3**55(1403), 1593–1597. 31

Hall, J. G. 1990. Genomic Imprinting - Review and Relevance to Human-Diseases. *American Journal of Human Genetics*, **4**6(5), 857–873. 94, 96

Hamilton, W. D. 1964. The genetic evolution of social behavior. *J Theoretical Biology*, **7**, 17–18. 28, 29

Hardling, R., & Nilsson, P. 1999. Parent-offspring and sexual conflicts in the evolution of angiosperm seeds. *Oikos*, **8**4(1), 27–34. 30

Hardling, R., & Nilsson, P. 2001. A model of triploid endosperm evolution driven by parent-offspring conflict. *Oikos*, **9**2(3), 417–423. 30, 31

Haun, W. J., & Springer, N. M. 2008. Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes. *Plant Journal*, **5**6(6), 903–912. 88

Haun, W. J., Laoueillé-Duprat, Sylvia, O'Connell, Mary J., Spillane, Charles, Grossniklaus, Ueli, Phillips, Allison R., Kaeppler, Shawn M., & Springer, Nathan M. 2007. Genomic imprinting, methylation and molecular evolution of maize Enhancer of zeste (Mez) homologs. *Plant Journal*, **4**9(2), 325–337. 10, 19, 20, 88

Haun, W. J., Danilevskaya, O. N., Meeley, R. B., & Springer, N. M. 2009. Disruption of Imprinting by Mutator Transposon Insertions in the 5 ' Proximal Regions of the Zea mays Mez1 Locus. *Genetics*, **1**81(4), 1229–1237. 88

Henikoff, S., Henikoff, J.G., Sakai, A., Loeb, G.B., & Ahmad, K. 2009. Genome-wide profiling of salt fractions maps physical properties of chromatin. *Genome research*, **1**9(3), 460. 11

Hermon, P., Srilunchang, K. O., Zou, J., Dresselhaus, T., & Danilevskaya, O. N. 2007. Activation of the imprinted polycomb group Fie1 gene in maize endosperm requires demethylation of the maternal allele. *Plant Molecular Biology*, **6**4(4), 387–395. 88

Holland, J. N., Chamberlain, S. A., Waguespack, A. M., & Kinyo, A. S. 2009. Effects of Pollen Load and Donor Diversity on Seed and Fruit Mass in the Columnar Cactus, Pachycereus Schottii (Cactaceae). *International Journal of Plant Sciences*, **1**70(4), 467–475. 71

Holländer, G.A., Zuklys, S., Morel, C., Mizoguchi, E., Mobisson, K., Simpson, S., Terhorst, C., Wishart, W., Golan, D.E., Bhan, A.K., *et al.* 1998. Monoallelic expression of the interleukin-2 locus. *Science*, **2**79(5359), 2118. 9

Holliday, R. 1990. Genomic Imprinting and Allelic Exclusion. *Development*, 125–129. 94

House, C., Roth, C., Hunt, J., & Kover, P.X. 2010. Paternal effects in Arabidopsis indicate that offspring can influence their own size. *Proceedings of the Royal Society B.* 70, 71, 73

Hurst, L. D. 1999. Is multiple paternity necessary for the evolution of genomic imprinting ? *Genetics*, **1**53(1), 509–512. 31

Hurst, L. D., & McVean, G. T. 1997. Growth effects of uniparental disomies and the conflict theory of genomic imprinting. *Trends in Genetics*, **1**3(11), 436–443. 85, 106

Hurst, L. D., & McVean, G. T. 1998. Do we understand the evolution of genomic imprinting ? *Current Opinion in Genetics & Development*, **8**(6), 701–708. 85, 86, 105, 106

Hurst, L.D. 1997. Evolutionary theories of genomic imprinting. *Pages 211–237 of :* Reik, W., & Surani, A. (eds), *Genomic Imprinting.* Frontiers in molecular biology. Oxford : Oxford University Press. 9, 85, 86, 93, 94, 95, 96, 100, 101, 106

Ikeda, Y., Kobayashi, Y., Yamaguchi, A., Abe, M., & Araki, T. 2007. Molecular basis of late-flowering phenotype caused by dominant epi-alleles of the FWA locus in Arabidopsis. *Plant and Cell Physiology*, **4**8(2), 205–220. 88

Ingouff, M., Gerald, J. N. F., Guerin, C., Robert, H., Sorensen, M. B., Van Damme, D., Geelen, D., Blanchoin, L., & Berger, F. 2005. Plant formin AtFH5 is an evolutionarily conserved actin nucleator involved in cytokinesis. *Nature Cell Biology*, **7**(4), 374–U34. 10, 88, 92

Ingouff, M., Sakata, T., Li, J., Sprunck, S., Dresselhaus, T., & Berger, F. 2009. The two male gametes share equal ability to fertilize the egg cell in Arabidopsis thaliana. *Current Biology*, **1**9(1), R19–R20. 97

Ingouff, Mathieu, Hamamura, Yuki, GourgueS, Mathieu, Higashiyama, Tetsuya, & Berger, Frederic. 2007. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. *CURRENT BIOLOGY*, **1**7(12), 1032–1037. 98

Jahnke, S., & Scholten, S. 2009. Epigenetic Resetting of a Gene Imprinted in Plant Embryos. *Current Biology*, **1**9(19), 1677–1681. 10, 11, 88

Jones, R. J., Schreiber, B. M. N., & Roessler, J. A. 1996. Kernel sink capacity in maize : Genotypic and maternal regulation. *Crop Science*, **3**6(2), 301–306. 63, 91

Judd, W.S., Campbell, C.S., Bouharmont, J., Kellogg, E.A., Évrard, C.M., & Stevens, P. 2001. *Botanique systématique : une perspective phylogénétique*. De Boeck Université. 63

Jullien, P. E., & Berger, F. 2010. Parental Genome Dosage Imbalance Deregulates Imprinting in Arabidopsis. *Plos Genetics*, **6**(3). 90

Jullien, P. E., Kinoshita, T., Ohad, N., & Berger, F. 2006a. Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. *Plant Cell*, **1**8(6), 1360–1372. 10, 88

Jullien, P. E., Katz, A., Oliva, M., Ohad, N., & Berger, F. 2006b. Polycomb group complexes self-regulate imprinting of the polycomb group gene MEDEA in Arabidopsis. *Current Biology*, **1**6(5), 486–492. 92

Jullien, P. E., Mosquna, A., Ingouff, M., Sakata, T., Ohad, N., & Berger, F. 2008. Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis. *Plos Biology*, **6**(8), 1693–1705. 87

Jullien, Pauline Emilie, & Berger, Frederic. 2009. Gamete-specific epigenetic mechanisms shape genomic imprinting. *Current Opinion in Plant Biology*, **1**2(5), 637–642. 20, 94, 100

Kawabe, A., Fujimoto, R., & Charlesworth, D. 2007. High diversity due to balancing selection in the promoter region of the medea gene in Arabidopsis lyrata. *Current Biology*, **1**7(21), 1885–1889. 86

Kermicle, J. L. 1970. Dependence of the R-mottled aleurone phenotype in maize on mode of sexual transmission. *Genetics*, **6**6, 69–85. 10, 88

Keverne, E. B., Fundele, R., Narasimha, M., Barton, S. C., & Surani, M. A. 1996. Genomic imprinting and the differential roles of parental genomes in brain development. *Developmental Brain Research*, **9**2(1), 91–100. 106

Kinoshita, T., Yadegari, R., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 1999. Imprinting of the MEDEA polycomb gene in the Arabidopsis endosperm. *Plant Cell*, **1**1(10), 1945–1952. 8, 10

Kinoshita, T., Miura, A., Kinoshita, Y., & Kakutani, T. 2003. Parent-of-origin-dependent FWA gene expression in the endosperm of Arabidopsis. *Plant and Cell Physiology*, **4**4, S28–S28. 10, 88

Kinoshita, T., Miura, A., Kinoshita, Y., & Kakutani, T. 2004a. Mechanism of FWA imprinting in Arabidopsis endosperm. *Plant and Cell Physiology*, **4**5, S72–S72. 88

Kinoshita, T., Miura, A., Choi, Y. H., Kinoshita, Y., Cao, X. F., Jacobsen, S. E., Fischer, R. L., & Kakutani, T. 2004b. One-way control of FWA imprinting in Arabidopsis endosperm by DNA methylation. *Science*, **3**03(5657), 521–523. 88, 92

Kinoshita, Y., Saze, H., Kinoshita, T., Miura, A., Soppe, W. J. J., Koornneef, M., & Kakutani, T. 2007. Control of FWA gene silencing in Arabidopsis thaliana by SINE-related direct repeats. *Plant Journal*, **4**9(1), 38–45. 88

Kiyosue, T., Ohad, N., Yadegari, R., Hannon, M., Dinneny, J., Wells, D., Katz, A., Margossian, L., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 1999. Control of fertilization-independent endosperm development by the MEDEA polycomb gene Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **9**6(7), 4186–4191. 92

Klotz, S., KÃ$\frac{1}{4}$hn, I., & Durka, W. 2002. BiolFlor - A database on biological and ecological traits of vascular plants in Germany. *Schriftenreihe fÃ$\frac{1}{4}$r Vegetationskunde 38*, Pp. 334. 59, 62

Köhler, C, & Grossniklaus, U. 2005. Seed development and genomic imprinting in plants. *Progress in Molecular and Subcellular Biology*, **3**8, 237–62. 95

Köhler, C., Hennig, L., Spillane, C., Pien, S., Gruissem, W., & Grossniklaus, U. 2003. The Polycomb-group protein MEDEA regulates seed development by controlling expression of the MADS-box gene PHERES1. *Genes & Development*, **1**7(12), 1540–1553. 10, 88, 92, 97

Köhler, C., Page, D. R., Gagliardini, V., & Grossniklaus, U. 2005. The Arabidopsis thaliana MEDEA Polycomb group protein controls expression of PHERES1 by parental imprinting. *Nature Genetics*, **3**7(1), 28–30. 10

Korbecka, G., Klinkhamer, P. G. L., & Vrieling, K. 2002. Selective embryo abortion hypothesis revisited - A molecular approach. *Plant Biology*, **4**(3), 298–310. 25

Krannitz, P. G., Aarssen, L. W., & Dow, J. M. 1991. The Effect of Genetically Based Differences in Seed Size on Seedling Survival in Arabidopsis-Thaliana (Brassicaceae). *American Journal of Botany*, **7**8(3), 446–450. 26

Kranz, E, von Wiegen, P, Quader, H, & Lorz, H. 1998. Endosperm development after fusion of isolated, single maize sperm and central cells in vitro. *PLANT CELL*, **1**0(4), 511–524. 14

Krishnamurthy, K. S., Uma Shaanker, R., & Ganeshaiah, K. N. 1997. Seed abortion in an animal dispersed species, Syzygium cuminii (L.) Skeels (Myrtaceae) : The chemical basis. *Current Science*, **7**3(10), 869–873. 60

Lande, R., & Schemske, D. W. 1985. The Evolution of Self-Fertilization and Inbreeding Depression in Plants .1. Genetic Models. *Evolution*, **3**9(1), 24–40. 61

Law, R., & Cannings, C. 1984. Genetic Analysis of Conflicts Arising during development of Seeds in the Angiospermophyta. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **2**21(1222), 53–70. 30

Leblanc, O., Pointe, C., & Hernandez, M. 2002. Cell cycle progression during endosperm development in Zea mays depends on parental dosage effects. *Plant Journal*, **3**2(6), 1057–1066. 100

Lee, TD, & Bazzaz, FA. 1986. Maternal regulation of fecundity : non-random ovule abortion inCassia fasciculata Michx. *Oecologia*, **6**8(3), 459–465. 27

Leishman, M.R., Wright, I.J., Moles, A.T., & Westoby, M. 2000. The evolutionary ecology of seed size. *Pages 31–57 of :* Fenner, M. (ed), *Seeds. The Ecology of Regeneration in Plant Communities.* Wallingford, UK : CABI Publishing. 26, 72

Li, Q. J., Xu, B., Chen, X. Y., & Wang, L. J. 2007. The effects of increased expression of an Arabidopsis HD-ZIP gene on leaf morphogenesis and anther dehiscence. *Plant Science*, **1**73, 567–576. 93

Lin, B.Y. 1984. Ploidy barrier to endosperm development in maize. *Genetics*, **1**07(1), 103. 17

Lipow, S. R., & Wyatt, R. 1999. Diallel crosses reveal patterns of variation in fruit-set, seed mass, and seed number in Asclepias incarnata. *Heredity*, **8**3, 310–318. 71

Lloyd, D.G. 1980. Sexual strategies in plants. I. An hypothesis of serial adjustment of maternal investment during one reproductive session. *New Phytologist*, **8**6(1), 69–79. 25

Lund, G., Messing, J., & Viotti, A. 1995. Endosperm-Specific Demethylation and Activation of Specific Alleles of Alpha-Tubulin Genes of Zea-Mays L. *Molecular & General Genetics*, **2**46(6), 716–722. 10

Luo, M., Bilodeau, P., Dennis, E. S., Peacock, W. J., & Chaudhury, A. 2000. Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proceedings of the National Academy of Sciences of the United States of America*, **9**7(19), 10637–10642. 10, 88

Luo, M., Platten, D., Chaudhury, A., Peacock, WJ, & Dennis, E.S. 2009. Expression, imprinting, and evolution of rice homologs of the polycomb group genes. *Molecular plant.* 10, 88, 91

Lyon, M.F. 1972. X-chromosome inactivation and developmental patterns in mammals. *Biological Reviews*, **4**7(1), 1–35. 9

Maheshwari, P. 1948. The angiosperm embryo sac. *The Botanical review*, **1**4(1), 1–56. 4, 5

Maheshwari, P. 1963. *Recent advances in the embryology of angiosperms.* Delhi, India : Ed. International Society of Plant Morphologists, University of Delhi. 4, 8, 18, 59, 62, 66, 107

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., & Kohler, C. 2006. Different Polycomb group complexes regulate common target genes in Arabidopsis. *Embo Reports*, **7**(9), 947–952. 88

Makarevich, G., Villar, C. B. R., Erilova, A., & Kohler, C. 2008. Mechanism of PHERES1 imprinting in Arabidopsis. *Journal of Cell Science*, **1**21(6), 906–912. 88

Matsuoka, S., Thompson, J.S., Edwards, M.C., Bartletta, JM, Grundy, P., Kalikin, L.M., Harper, J.W., Elledge, S.J., & Feinberg, A.P. 1996. Imprinting of the gene encoding a human cyclin-dependent kinase inhibitor, p57KIP2, on chromosome 11p15. *Proceedings of the National Academy of Sciences of the United States of America*, **9**3(7), 3026. 94, 109

Mazer, S. J. 1987. Maternal Investment and Male Reproductive Success in Angiosperms - Parent Offspring Conflict or Sexual Selection. *Biological Journal of the Linnean Society*, **3**0(2), 115–133. 28, 30

Mazer, S. J., Snow, A. A., & Stanton, M. L. 1986. Fertilization Dynamics and Parental Effects Upon Fruit-Development in Raphanus-Raphanistrum - Consequences for Seed Size Variation. *American Journal of Botany*, **7**3(4), 500–511. 71

McGowan, RA, & Martin, CC. 1997. DNA methylation and genome imprinting in the zebrafish, Danio rerio : some evolutionary ramifications. *Biochemistry and Cell Biology*, **7**5(5), 499–506. 9

McGrath, J., & Solter, D. 1984. Completion of mouse embryogenesis requires both the maternal and paternal genomes. *Cell*, **3**7(1), 179–183. 9

McVean, G. T., & Hurst, L. D. 1997. Molecular evolution of imprinted genes : No evidence for antagonistic coevolution. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **2**64(1382), 739–746. 85, 86, 105

Messing, J., & Grossniklaus, U. 1999. *Genomic imprinting in plants.* Results and Problems in Cell Differentiation, vol. 25. Springer, New York. 97

Miyake, T., Takebayashi, N., & Wolf, D. E. 2009. Possible Diversifying Selection in the Imprinted Gene, MEDEA, in Arabidopsis. *Molecular Biology and Evolution*, **2**6(4), 843–857. 86

Mochizuki, A., Takeda, Y., & Iwasa, Y. 1996. The evolution of genomic imprinting. *Genetics*, **1**44(3), 1283–1295. 31, 43, 105, 109

Mohana, G. S., Uma Shaanker, R., Ganeshaiah, K. N., & Dayanandan, S. 2001. Genetic relatedness among developing seeds and intra fruit seed abortion in Dalbergia sissoo (Fabaceae). *American Journal of Botany*, **8**8(7), 1181–1188. 60

Moles, A. T., & Westoby, M. 2004. Seedling survival and seed size : a synthesis of the literature. *Journal of Ecology*, **9**2(3), 372–383. 26

Moles, A.T., Falster, D.S., Leishman, M.R., & Westoby, M. 2004. Small-seeded species produce more seeds per square metre of canopy per year, but not per individual per lifetime. *Journal of Ecology*, **9**2(3), 384–396. 72

Moles, A.T., Ackerly, D.D., Webb, C.O., Tweddle, J.C., Dickie, J.B., Pitman, A.J., & Westoby, M. 2005. Factors that shape seed mass evolution. *Proceedings of the National Academy of Sciences of the United States of America*, **1**02(30), 10540. 26, 72

Moore, T., & Haig, D. 1991. Genomic Imprinting in Mammalian Development - a Parental Tug-of-War. *Trends in Genetics*, **7**(2), 45–49. 9

Moore, T., & Mills, W. 2008. Evolutionary theories of imprinting - Enough already ! *Pages 116–122 of :* Wilkins, J. F. (ed), *Genomic Imprinting.* Advances in Experimental Medicine and Biology, vol. 626. Berlin : SPRINGER-VERLAG BERLIN. 85

Nakamura, M., Katsumata, H., Abe, M., Yabe, N., Komeda, Y., Yamamoto, K. T., & Takahashi, T. 2006. Characterization of the class IV homeodomain-leucine zipper gene family in Arabidopsis. *Plant Physiology*, **1**41(4), 1363–1375. 10, 88, 92

Nakamura, R. R., & Stanton, M. L. 1989. Embryo Growth and Seed Size in Raphanus-Sativus - Maternal and Paternal Effects Invivo and Invitro. *Evolution*, **4**3(7), 1435–1443. 71

Nawaschin, S.G. 1898. Resultate einer Revision der Befruchtungsvorg'ange bei Lilium Martagon und Fritillaria tenella. *Bull. Acad. Sci. St Petersburg*, **9**, 377–382. 6, 14

Ng, K., Pullirsch, D., Leeb, M., & Wutz, A. 2007. Xist and the order of silencing. *Embo reports*, **8**(1), 34–39. 55

Niesenbaum, R. A. 1999. The effects of pollen load size and donor diversity on pollen performance, selective abortion, and progeny vigor in Mirabilis jalapa (Nyctaginaceae). *American Journal of Botany*, **8**6(2), 261–268. 71

Nowack, M. K., Grini, P. E., Jakoby, M. J., Lafos, M., Koncz, C., & Schnittger, A. 2006. A positive signal from the fertilization of the egg cell sets off endosperm proliferation in angiosperm embryogenesis. *Nature Genetics*, **3**8(1), 63–67. 96, 97

Ohad, N, Margossian, L, Hsu, Y C, Williams, C, Repetti, P, & Fischer, R L. 1996. A mutation that allows endosperm development without fertilization. *PNAS*, **9**3(11), 5319–5324. 10, 91

Ohad, N., Yadegari, R., Margossian, L., Hannon, M., Michaeli, D., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 1999. Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell*, **1**1(3), 407–415. 10

Otto, S.P. 1993. The role of Deleterious and Beneficial Mutations in the Evolution of Ploidy level. *Lectures on Mathematics in the life Science*, **2**5. 35

Otto, S.P., & Goldstein, D.B. 1992. Recombination and the evolution of diploidy. *Genetics*, **1**31, 745–751. 35

Otto, S.P., & Marks, J.C. 1996. Mating systems and the evolutionary transition between haploidy and diploidy. *Biological Journal of the Linnean Society*, **5**7, 197–218. 36, 44

Parisi, M., Nuttall, R., Naiman, D., Bouffard, G., Malley, J., Andrews, J., Eastman, S., & Oliver, B. 2003. Paucity of genes on the Drosophila X chromosome showing male-biased expression. *Science*, **2**99(5607), 697. 9

Paschke, M., Bernasconi, G., & Schmid, B. 2005. Effects of inbreeding and pollen donor provenance and diversity on offspring performance under environmental stress in the rare plant Cochlearia bavarica. *Basic and Applied Ecology*, **6**(4), 325–338. 71

Pearce, G. P., & Spencer, H. G. 1992. Population Genetic Models of Genomic Imprinting. *Genetics*, **1**30(4), 899–907. 32, 35

Pernis, B., Chiappino, G., Kelus, A.S., & Gell, P.G.H. 1965. Cellular localization of immunoglobulins with different allotypic specificities in rabbit lymphoid tissues. *Journal of Experimental Medicine*, **1**22(5), 853. 9

Perrot, V., Richerd, S., & Valero, M. 1991. Transition from Haploidy to Diploidy. *Nature*, **3**51(6324), 315–317. 35

Pillot, M., Baroux, C., Vazquez, M. A., Autran, D., Leblanc, O., Vielle-Calzada, J. P., Grossniklaus, U., & Grimanelli, D. 2010. Embryo and Endosperm Inherit Distinct Chromatin and Transcriptional States from the Female Gametes in Arabidopsis. *Plant Cell*, **2**2(2), 307–320. 54, 98

Portereiko, M. F., Lloyd, A., Steffen, J. G., Punwani, J. A., Otsuga, D., & Drews, G. N. 2006. AGL80 is required for central cell and endosperm development in Arabidopsis. *Plant Cell*, **1**8(8), 1862–1872. 10, 88, 92

Queller, D. C. 1983. Kin Selection and Conflict in Seed Maturation. *Journal of Theoretical Biology*, **1**00(1), 153–172. 28, 30, 60, 104

Queller, D. C. 1984. Models of Kin Selection on Seed Provisioning. *Heredity*, **5**3(AUG), 151–165. 30

Raghavan, V. 2003. Some reflections on double fertilization, from its discovery to the present. *New Phytologist*, **1**59(3), 565–583. 6, 7

Raghavan, V. 2006. *Double fertilization : embryo and endosperm development in flowering plants*. Springer. 6, 95, 97

Rajewsky, K. 1996. Clonal selection and learning in the antibody system. *Nature*, **3**81(6585), 751–758. 9

Ravishankar, K. V., Uma Shaanker, R., & Ganeshaiah, K. N. 1995. War of Hormones over Resource-Allocation to Seeds - Strategies and Counterstrategies of Offspring and Maternal Parent. *Journal of Biosciences*, **2**0(1), 89–103. 104, 106

Rhoades, K.L., Singh, N., Simon, I., Glidden, B., Cedar, H., & Chess, A. 2000. Allele-specific expression patterns of interleukin-2 and Pax-5 revealed by a sensitive single-cell RT-PCR analysis. *Current Biology*, **1**0(13), 789–792. 9

Richardson, T. E., & Stephenson, A. G. 1991. Effects of Parentage, Prior Fruit-Set and Pollen Load on Fruit and Seed Production in Campanula Americana L. *Oecologia*, **8**7(1), 80–85. 71

Roze, D., & Lenormand, T. 2005. Self fertilization and the evolution of recombination. *Genetics*, **1**70, 841–857. 45

Russell, S.D. 1985. Preferential fertilization in Plumbago : ultrastructural evidence for gamete-level recognition in an angiosperm. *Proceedings of the National Academy of Sciences of the United States of America*, **8**2(18), 6129. 97

Sadras, V. O., & Denison, R. F. 2009. Do plant parts compete for resources ? An evolutionary viewpoint. *New Phytologist*, **1**83(3), 565–574. 107

Sakai, S. 2007. A new hypothesis for the evolution of overproduction of ovules : An advantage of selective abortion for females not associated with variation in genetic quality of the resulting seeds. *Evolution*, **6**1(4), 984–993. 104

Sargant, E. 1900. Recent work on the results of fertilization in angiosperms. *Ann. Bot.*, **1**4, 689–712. 14

Schemske, D. W., & Pautler, L. P. 1984. The Effects of Pollen Composition on Fitness Components in a Neotropical Herb. *Oecologia*, **6**2(1), 31–36. 71

Schier, A.F. 2007. The maternal-zygotic transition : death and birth of RNAs. *Science*, **3**16(5823), 406. 54

Schwaegerle, K. E., & Levin, D. A. 1990. Quantitative Genetics of Seed Size Variation in Phlox. *Evolutionary Ecology*, **4**(2), 143–148. 71

Sharman, GB. 1971. Late DNA replication in the paternally derived X chromosome of female kangaroos. *Nature*, **2**30, 231–232. 9

Sleutels, F., Barlow, D.P., & Lyle, R. 2000. The uniqueness of the imprinting mechanism. *Current Opinion in Genetics & Development*, **1**0(2), 229–233. 8

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijó, J.A., & Martienssen, R.A. 2009. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. *Cell*, **1**36(3), 461–472. 98

Smith, Christopher C., & Fretwell, Stephen D. 1974. The Optimal Balance between Size and Number of Offspring. *American Naturalist*, **1**08, 499. 27

Solter, D. 1987. Inertia of the Embryonic Genome in Mammals. *Trends in Genetics*, **3**(1), 23–27. 94, 95

Solter, D. 1988. Differential Imprinting and Expression of Maternal and Paternal Genomes. *Annual Review of Genetics*, **2**2, 127–146. 94, 95, 100

Sorensen, F. C. 1982. The Roles of Polyembryony and Embryo Viability in the Genetic System of Conifers. *Evolution*, **3**6(4), 725–733. 25

Sorensen, M. B., Chaudhury, A. M., Robert, H., Bancharel, E., & Berger, F. 2001. Polycomb group genes control pattern formation in plant seed. *Current Biology*, **1**1(4), 277–281. 92, 97

Spencer, H. G. 2000. Population genetics and evolution of genomic imprinting. *Annual Review of Genetics*, **3**4, 457–477. 95

Spencer, H. G., & Williams, J. M. 1997. The evolution of genomic imprinting : Two modifier-locus models. *Theoretical Population Biology*, **5**1(1), 23–35. 32, 35, 43, 56

Spencer, H. G., Feldman, M. W., & Clark, A. G. 1998. Genetic conflicts, multiple paternity and the evolution of genomic imprinting. *Genetics*, **1**48(2), 893–904. 31

Spencer, H. G., Clark, A. G., & Feldman, M. W. 1999. Genetic conflicts and the evolutionary origin of genomic imprinting. *Trends in Ecology & Evolution*, **1**4(5), 197–201. 31

Spichiger, RE, Savolainen, VV, & Figeat, M. 2000. *Botanique systématique des plantes à fleurs*. Presses Polytechniques et Universitaires Romandes. 59, 62

Spillane, C., MacDougall, C., Stock, C., Kohler, C., Vielle-Calzada, J. P., Nunes, S. M., Grossniklaus, U., & Goodrich, J. 2000. Interaction of the Arabidopsis Polycomb group proteins FIE and MEA mediates their common phenotypes. *Current Biology*, **1**0(23), 1535–1538. 97

Spillane, C., Schmid, K. J., Laoueille-Duprat, S., Pien, S., Escobar-Restrepo, J. M., Baroux, C., Gagliardini, V., Page, D. R., Wolfe, K. H., & Grossniklaus, U. 2007. Positive darwinian selection at the imprinted MEDEA locus in plants. *Nature*, **4**48(7151), 349–U8. 10, 14, 19, 20

Springer, N. M., Danilevskaya, O. N., Hermon, P., Helentjaris, T. G., Phillips, R. L., Kaeppler, H. F., & Kaeppler, S. M. 2002. Sequence relationships, conserved domains and expression patterns for maize homologs of the polycomb group genes E(z), esc, and E(Pc). *Plant Physiology*, **1**28(4), 1332–1345. 88

Stebbins, G. L. 1976. Seeds, seedlings, and the origin of angiosperms. *Pages 300–311 of :* Beck, C.B. (ed), *Origin and early evolution of angiosperms*. New York : Columbia Univ. Press. 25, 32, 110

Stebbins, GL. 1974. *Flowering plants : Evolution above the species level*. Cambridge : Belknap press. 25, 110

Stephenson, A. G. 1981. Flower and Fruit Abortion - Proximate Causes and Ultimate Functions. *Annual Review of Ecology and Systematics*, **1**2, 253–279. 27

Stewart-Cox, J. A., Britton, N. F., & Mogie, M. 2004. Endosperm triploidy has a selective advantage during ongoing parental conflict by imprinting. *Proceedings of the Royal Society of London Series B-Biological Sciences*, **2**71(1549), 1737–1743. 31

Strasburger. 1900. Einige bemerkungen zur frage nach der "doppelten befruchtung" bei angiospermen. *Bot. Zeit.*, **5**8, 293–316. 14

Stupar, R.M., Gardiner, J.M., Oldre, A.G., Haun, W. J., Chandler, V.L., & Springer, N.M. 2008. Gene expression analyses in maize inbreds and hybrids with varying levels of heterosis. *BMC Plant Biology*, **8**(33), 1–32. 11

Surani, M.A. 1991. Genomic imprinting : developmental significance and molecular mechanism. *Current Opinion in Genetics & Development*, **1**(2), 241–246. 1

Surani, M.A. 2001. Reprogramming of genome function through epigenetic inheritance. *Nature*, **4**14(6859), 122–128. 97

Szafraniec, K., Wloch, D. M., Sliwa, P., Borts, R. H., & Korona, R. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast Saccharomyces cerevisiae. *Genetical Research*, **8**2(1), 19–31. 32, 40, 109

Takagi, N., & Sasaki, M. 1975. Preferential inactivation of the paternally derived X chromosome in the extraembryonic membranes of the mouse. *Nature*, **2**56, 640 – 642. 9

Takhtajan, A. 1976. Neoteny and the origin of flowering plants. *Pages 207–219 of :* Beck, C.B. (ed), *Origin and early evolution of angiosperms.* New York : Columbia university press. 25

Takhtajan, A. 1991. *Evolutionary Trends in Flowering Plants.* Columbia : Columbia University Press. 32

Tiffney, BH. 1981. Diversity and major events in the evolution of land plants. *Pages 193–230 of :* Niklas, K, J (ed), *Paleobotany, paleoecology and evolution*, vol. 2. New York : Praeger. 32

Tiwari, S., Schulz, R., Ikeda, Y., Dytham, L., Bravo, J., Mathers, L., Spielman, M., Guzman, P., Oakey, R. J., Kinoshita, T., & Scott, R. J. 2008. MATERNALLY EXPRESSED PAB C-TERMINAL, a Novel Imprinted Gene in Arabidopsis, Encodes the Conserved C-Terminal Domain of Polyadenylate Binding Proteins. *Plant Cell*, **2**0(9), 2387–2398. 10, 88, 92

Turnbull, L. A., Rees, M., & Crawley, M. J. 1999. Seed mass and the competition/colonization trade-off : a sowing experiment. *Journal of Ecology*, **8**7(5), 899–912. 26, 72

Uma Shaanker, R., & Ganeshaiah, K. N. 1996. Polyembryony in plants : A weapon in the war over offspring numbers ? *Trends in Ecology & Evolution*, **1**1(1), 26–27. 104

Uma Shaanker, R., & Ganeshaiah, K. N. 1997. Conflict between parent and offspring in plants : Predictions, processes and evolutionary consequences. *Current Science*, **7**2(12), 932–939. 30, 71, 107

Uma Shaanker, R., Ganeshaiah, K. N., & Bawa, K. S. 1988. Parent-Offspring Conflict, Sibling Rivalry, and Brood Size Patterns in Plants. *Annual Review of Ecology and Systematics*, **1**9, 177–205. 59, 62, 104

Uma Shaanker, R., Ravishankar, K. V., Hegde, S. G., & Ganeshaiah, K. N. 1996. Does endosperm reduce intra-fruit competition among developing seeds ? *Plant Systematics and Evolution*, **2**01(1-4), 263–270. 59, 60

Vanderkloet, S. P., & Tosh, D. 1984. Effects of Pollen Donors on Seed Production, Seed Weight, Germination and Seedling Vigor in Vaccinium-Corymbosum L. *American Midland Naturalist*, **1**12(2), 392–396. 71

Varmuza, S., & Mann, M. 1994. Genomic Imprinting - Defusing the Ovarian Time Bomb. *Trends in Genetics*, **1**0(4), 118–123. 96

Vassilieva, L. L., Hook, A. M., & Lynch, M. 2000. The fitness effects of spontaneous mutations in Caenorhabditis elegans. *Evolution*, **5**4(4), 1234–1246. 32, 40, 109

Vielle-Calzada, J. P., Thomas, J., Spillane, C., Coluccio, A., Hoeppner, M. A., & Grossniklaus, U. 1999. Maintenance of genomic imprinting at the Arabidopsis medea locus requires zygotic DDM1 activity. *Genes & Development*, **1**3(22), 2971–2982. 10, 87, 88

Vielle-Calzada, J. P., Baskar, R., & Grossniklaus, U. 2000. Delayed activation of the paternal genome during seed development. *Nature*, **4**04(6773), 91–94. 54, 98

Vinkenoog, R. 2000. Hypomethylation promotes autonomous endosperm development and rescues postfertilization lethality in fie mutants. *Plant Cell*, **1**2(11), 2271–2282. 95

Vinkenoog, R., Bushell, C., Spielman, M., Adams, S., Dickinson, H. G., & Scott, R. J. 2003. Genomic imprinting and endosperm development in flowering plants. *Molecular Biotechnology*, **2**5(2), 149–184. 17, 91

Wang, D. F., Tyson, M. D., Jackson, S. S., & Yadegari, R. 2006. Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, **1**03(35), 13244–13249. 10, 88

Wang, Y., Joh, K., Masuko, S., Yatsuki, H., Soejima, H., Nabetani, A., Beechey, C.V., Okinami, S., & Mukai, T. 2004. The mouse Murr1 gene is imprinted in the adult brain, presumably due to transcriptional interference by the antisense-oriented U2af1-rs1 gene. *Molecular and cellular biology*, **2**4(1), 270. 106

Waser, N. M., Shaw, R. G., & Price, M. V. 1995. Seed Set and Seed Mass in Ipomopsis-Aggregata - Variance Partitioning and Inferences About Postpollination Selection. *Evolution*, **4**9(1), 80–88. 71

Webb, C. O., Ackerly, D. D., & Kembel, S. W. 2008. Phylocom : software for the analysis of phylogenetic community structure and trait evolution. *Bioinformatics*, **2**4, 2098–2100. 18

West, J.D., Frels, W.I., Chapman, V.M., & Papaioannou, V.E. 1977. Preferential expression of the maternally derived X chromosome in the mouse yolk sac. *Cell*, **1**2(4), 873–882. 9

Westoby, M., & Rice, B. 1982. Evolution of the Seed Plants and Inclusive Fitness of Plant-Tissues. *Evolution*, **3**6(4), 713–724. 25, 27, 28, 30, 60, 83, 104, 111

Wiens, D. 1984. Ovule Survivorship, Brood Size, Life-History, Breeding Systems, and Reproductive Success in Plants. *Oecologia*, **6**4(1), 47–53. 66

Williams, J. H., & Friedman, W. E. 2004. The four-celled female gametophyte of Illicium (Illiciaceae ; Austrobaileyales) : Implications for understanding the origin and early evolution of monocots, eumagnoliids, and eudicots. *American Journal of Botany*, **9**1(3), 332–351. 14

Willison, K. 1991. Opposite imprinting of the mouse Igf2 and Igf2r genes. *Trends in genetics : TIG*, **7**(4), 107. 31

Willson, MF, & Burley, N. 1983. *Mate Choice in Plants : Tactics, Mechanisms, and Consequences*. Princeton, New Jersey : Princeton university press. 14, 25, 27, 28, 30, 32, 35, 56, 83, 105, 111

Wolfe, L. M. 1995. The Genetics and Ecology of Seed Size Variation in a Biennial Plant, Hydrophyllum-Appendiculatum (Hydrophyllaceae). *Oecologia*, **1**01(3), 343–352. 71

Xiao, W. Y., Brown, R. C., Lemmon, B. E., Harada, J. J., Goldberg, R. B., & Fischer, R. L. 2006. Regulation of seed size by hypomethylation of maternal and paternal genomes. *Plant Physiology*, **1**42(3), 1160–1168. 87

Zhang, D. Y. 1998. Evolutionarily stable reproductive strategies in sexual organisms : 4. Parent-offspring conflict and selection of seed size in perennial plants. *Journal of Theoretical Biology*, **1**92(2), 143–153. 31

Zhang, D. Y., & Jiang, X. H. 1998. Evolutionarily stable reproductive strategies in sexual organisms. 5. Joint effects of parent-offspring conflict and sibling conflict in perennial plants. *Journal of Theoretical Biology*, **1**92(3), 275–281. 31

Zhang, D. Y., & Jiang, X. H. 2000. Costly solicitation, timing of offspring conflict, and resource allocation in plants. *Annals of Botany*, **8**6(1), 123–131. 30

Annexe **A**

Manuscrit : Ploidy and the Evolution of Endosperm of Flowering Plants

Copyright © 2010 by the Genetics Society of America
DOI: 10.1534/genetics.109.110833

Ploidy and the Evolution of Endosperm of Flowering Plants

Aurélie Cailleau,[1] Pierre-Olivier Cheptou and Thomas Lenormand

Centre d'Ecologie Fonctionnelle et Evolutive, 34293 Montpellier, France

Manuscript received October 9, 2009
Accepted for publication November 16, 2009

ABSTRACT

In angiosperms, spermatozoa go by pair in each pollen grain and fertilize, in addition to the egg cell, one of its sister cells, called the central cell. This "double fertilization" leads to the embryo on the one hand and to its nutritive tissue, the endosperm, on the other hand. In addition, in most flowering plants, the endosperm is triploid because of a doubled maternal genetic contribution in the central cell. Most of the hypotheses trying to explain these eccentricities rest on the assumption of a male/female conflict over seed resource allocation. We investigate an alternative hypothesis on the basis of the masking of deleterious alleles. Using analytical methods, we show that a doubled maternal contribution and double fertilization tend to be favored in a wide range of conditions when deleterious mutations alter the function of the endosperm. Furthermore, we show that these conditions vary depending on whether these traits are under male or female control, which allows us to describe a new type of male/female conflict.

FEMALES provision their offspring with resources. These resources can be accumulated before fertilization. When some of the gametes are not fertilized, this strategy is extremely wasteful. In a number of species, females have evolved conditional strategies avoiding this problem—in particular by providing most resources after fertilization [*i.e.*, to diploid offspring instead of haploid gametes (Westoby and Rice 1982; Baroux *et al.* 2002)]. This postfertilization provisioning occurs, for instance, in viviparous animals, such as mammals, some reptiles (De Fraipont *et al.* 1996), some sharks (Wourms 1993), some scorpions (Brown and Formanowicz 1996), and velvet worms (Tutt *et al.* 2002). In plants, resource provisioning to offspring can also occur before or after fertilization.

The life cycle of plants typically alternates a sporophytic stage ($2n$) and a gametophytic stage (n) that are both multicellular. A newly formed sporophyte (the diploid zygote) will typically use the resources accumulated in tissues derived from the female haploid gametophyte and not in tissues derived from the diploid sporophyte. For instance, in ferns and mosses, the new embryo develops "parasitically" on the free gametophyte. In gymnosperms, the embryo acquires its resource from a nonfree female gametophyte (it is encapsulated within sporophytic tissues). In some of them (cycads and ginkgoes) the resource is fully accumulated before fertilization, while in some others (Pinaceae) the resource accumulates both before and after fertilization (Friedman 2001b). The former situation is comparable to what happens in many fish while the latter is comparable to what happens in many bird species (egg formation before fertilization and incubation after fertilization). In angiosperms, resource provisioning to the seed really starts after fertilization (Friedman 2001b), a situation comparable to placental mammals (Figure 1).

Compared to other seed plants, the case of angiosperms is complicated by the occurrence of a "double fertilization" (Nawaschin 1898; Guignard 1899): two identical (mitotically derived) spermatozoa from a single pollen grain fertilize in parallel two cells of the female gametophyte. The fertilization of the egg cell produces the embryo. The fertilization of the central cell produces the endosperm, a specific organ that serves as an interface for resource transfer between the diploid mother and its offspring (like the placenta in mammals). This tissue can either be transitory or serve as a storage tissue in the seed. In this article, we study the evolution of this nutritive tissue. More specifically, we focus on the evolution of its genetic constitution.

The diversity of endosperms reflects directly the diversity of female gametophytes (Friedman *et al.* 2008). In some angiosperms [Austrobaileyales (Friedman *et al.* 2003), Oenothera (Maheshwari 1963), and Nymphealeas (Williams and Friedman 2002)], the female gametophyte is made of four haploid cells that are mitotically derived from a single spore. One of them is the central cell and contains one maternal contribution (1m). This type of gametophyte is referred to as "Oenothera." After double fertilization (which adds one paternal contribution, 1p), the central cell becomes a diploid endosperm (1m:1p). This endosperm is genetically identical to the embryo. In other angiosperms, the female gametophyte is eight-nucleated, but seven-celled

[1]*Corresponding author:* CEFE-CNRS UMR 5175, 1919 Route de Mende, 34293 Montpellier Cedex 05, France.
E-mail: aurelie.cailleau@cefe.cnrs.fr

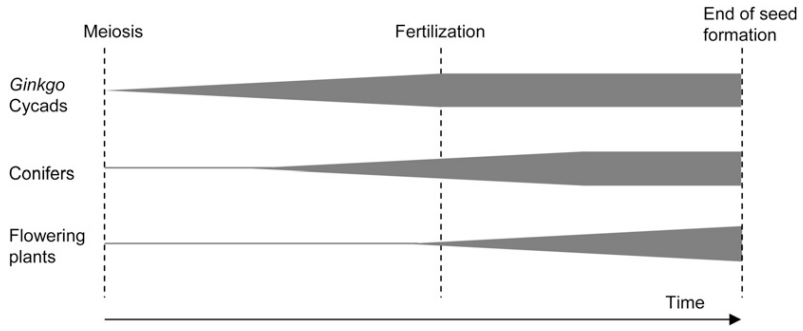

FIGURE 1.—Timing of resource accumulation in the seed. All the resources are allocated before fertilization in cycads and ginkgoes. Resources are allocated both before and after fertilization in conifers and all the resources are allocated after fertilization in angiosperms.

because the central cell remains binucleate (2m). This type of gametophyte is referred to as "Polygonum." Double fertilization again carries one paternal contribution (1p). Thus, the central cell becomes a triploid endosperm (2m:1p). Other types of female gametophytes can be found sporadically among angiosperms. They are characterized by an increased number of maternal contributions in the central cell and/or by polyspory. These types are derived from the Polygonum type and we do not consider them further. As already mentioned, in gymnosperms and ferns the embryo feeds directly on the haploid female gametophyte (1m:0p), which can be considered as the ancestral state of the nutritive tissue.

In this article, we focus on the evolution of double fertilization (DF) (0p → 1p) and maternal contribution doubling (MCD) (1m → 2m) from this 1m:0p ancestral state. In principle, two scenarios can be considered (Figure 2). In the first scenario, DF evolves before MCD (1m:0p → 1m:1p → 2m:1p). In the second, MCD evolves first and DF second (1m:0p → 2m:0p → 2m:1p). The latter scenario involves an additional step (secondary loss of the second maternal contribution, 2m:1p → 1m:1p) to account for the occurrence of 1m:1p endosperms in Nymphaeales (Nuphar, Nymphea, Cabomba) and the Illiciales-Trimeniaceae-Austrobaileyaceae clade (Schisandra, Illicium). It is therefore less parsimonious. For this reason, we consider only the first scenario (1m:0p → 1m:1p → 2m:1p) in this article.

Three main theories have been proposed to explain the evolution of DF and MCD. They all rely on the fact that some resources are accumulated to the seed after fertilization. The first theory proposes that postfertilization resource allocation triggers a genetic conflict over resource allocation among parents and/or offspring. Indeed, an antagonistic coevolution can occur between the male, the female, and/or the embryos when their optimal seed sizes differ, to take control over resource allocation (TRIVERS 1974; HAIG 2000). Because the endosperm plays a pivotal role in the control of resource allocation, DF and MCD are considered in this theory to be adaptations allowing the father and the mother to take and take back the control over resource allocation, respectively (CHARNOV 1979; WESTOBY and RICE 1982; QUELLER 1983; WILLSON and BURLEY 1983; LAW and CANNINGS 1984; BULMER 1986; SHAANKER *et al.* 1988; HAIG and WESTOBY 1989; FRIEDMAN 1995; SHAANKER and GANESHAIAH 1997; HARDLING and NILSSON 1999, 2001). These "allocation conflict theories" include the kinship theory of genomic imprinting (HAIG 2000, 2004; DE JONG *et al.* 2005) and have received the most extensive theoretical development.

A second theory (hereafter metabolic theory) proposes that DF and MCD evolve because of the metabolic properties of polyploid tissues (STEBBINS 1976; DONOGHUE and SCHEINER 1992). More specifically, the idea is that DF and MCD increase the endosperm ploidy level, which enhances its postfertilization metabolic efficiency (increased transcription rate, increased cell size, and resource storage). The problem with this view is that endomitosis would achieve the same effect and is very common at the beginning of gametophyte/endosperm development in both angiosperm and gymnosperm (FRIEDMAN 2001a,b): DF and MCD seem unnecessary

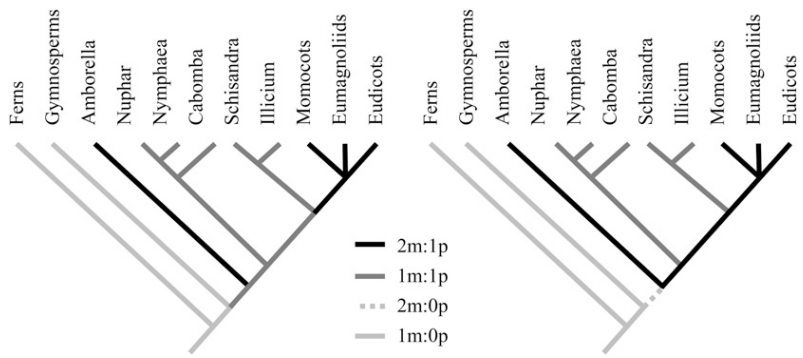

FIGURE 2.—Possible scenarios for the evolution of the genetic makeup of the embryo nutritive tissue, modified from FRIEDMAN and RYERSON (2009). In the left scenario, there are one 1m:0p → 1m:1p transition and two 1m:1p → 2m:1p transitions and thus three steps. In the right scenario, there are one 1m:0p → 2m:0p transition, one 2m:0p → 2m:1p transition, and two 2m:1p → 1m:1p transitions and thus four steps. In addition, the latter scenario involves a 2m:0p state that has never been observed in any taxa.

to increase ploidy in the endosperm in the first place. Moreover, this theory does not account for the presence of a male nucleus specifically (Willson and Burley 1983).

A third theory proposes that DF and MCD evolve because of the genetic properties of polyploid tissues (Brink and Cooper 1940). More specifically, the idea is that DF combines two different genomes (maternal and paternal) in the nutritive tissue, which masks the effect of recessive deleterious mutations and thus enhances the postfertilization metabolic efficiency of the nutritive tissue. Following Friedman *et al.* (2008), we refer to this theory as the "heterozygosity theory." The heterozygosity theory explanation for endosperm's genetic evolution seems close to explanations for ploidy evolution. Several models have been worked out to understand the effect of the masking of deleterious mutations in the evolution of ploidy levels (Perrot *et al.* 1991; Bengtsson 1992; Otto and Goldstein 1992). These models have shown that masking protects diploid individuals from the effect of recessive deleterious mutations, but that it also reduces the elimination of these mutations from the population, creating genetic association favorable to haploidy (Otto and Goldstein 1992). The problem of the evolution of the genetic constitution of the endosperm is, however, not directly comparable to the problem of life cycles evolution. First, the phenotype under selection (*e.g.*, resource in the seed) is at least partly determined by the genotype of the nutritive tissue (the gametophyte or the endosperm), which can differ from the genotype of the diploid embryo. Second, double fertilization is not symmetrical with respect to sexes: it masks maternal alleles but reveals paternal alleles in the endosperm. Third, double fertilization can be under pollen and/or female gametophyte control, while the timing of meiosis in life cycle evolution is controlled by the diploid stage. In fact, no formal model has been created to explore specifically the heterozygosity theory. The aim of this article is to develop such models.

We build two models to study the evolution of DF and MCD. More explicitly, we study the evolution of a modifier locus that changes their probability of occurrence. We show that masking deleterious mutations can favor both DF and MCD. However, unlike models for the evolution of ploidy, we find conditions where DF evolves in opposite directions depending on the sex where the modifier is expressed. In different conditions, the same scheme occurs for MCD. This situation reveals that male/female conflict extends well beyond the problem of resource allocation to embryos: a conflict can also occur over the genetic constitution of embryos' nutritive tissues.

MODEL

We analyze the evolution of DF and MCD in two distinct deterministic models. Both involve two linked diallelic loci. The first one (the viability locus) is subject to deleterious mutations and determines the survival of the embryo while the other one (the modifier locus) controls the probability of DF or MCD.

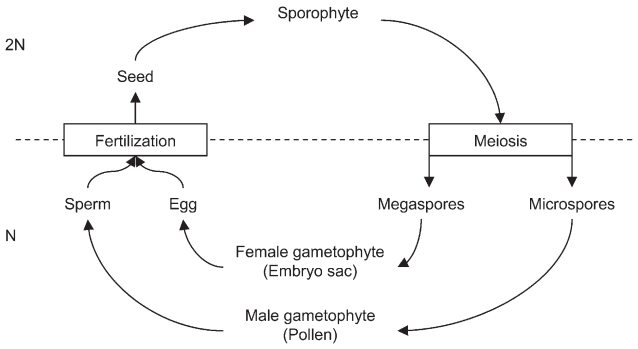

Figure 3.—Life cycle. The sporophyte is born from the fertilization of a male gamete and a female gamete. These gametes are produced by the male and female gametophytes (the pollen and the embryo sac, respectively). In the sporophyte's sexual organs, meiosis produces spores that develop in new gametophytes.

Life cycle: To build each model, we first write exact recursions of the frequency of the different haploid genotypes. We follow the different events in the life cycle over one generation, which alternates a gametophytic haploid with a sporophytic diploid stage (Figure 3). We start at the gametophytic stage. The male gametophyte (pollen) contains two identical sperm cells (they are mitotically derived). The female gametophyte (embryo sac) contains also a given number of genetically identical cells, among which are the egg cell and the central cell. First, fertilization occurs in a panmictic way (pairing one pollen grain with one embryo sac). The genetic constitution of the endosperm is determined at this stage, but differently in the different models (see below). Second, the nutritive tissue and the embryo develop (seed formation). Selection occurs among seeds in the population, on the basis of the genotype of their nutritive tissue. Third, the surviving embryos become hermaphroditic sporophytes and produce haploid spores through meiosis (where the two loci recombine at a rate R). Finally, the spores undergo mutation and each produces a new gametophyte through mitosis. To describe the different events in this life cycle, we introduce several parameters, specific to each locus. We present each locus in turn.

Viability locus: The first locus is expressed in the embryo's nutritive tissue (we refer to this tissue as the endosperm) where it potentially affects any function (*e.g.*, the accumulation or the transfer of resources to the embryo or the protection of the embryo). This function, in turn, determines the survival of the embryo. We model this directly by assigning different viabilities to embryos associated to different endosperm genotypes. We consider that recurrent deleterious mutations occur at a rate μ at the viability locus. We denote P_a the frequency of a (the deleterious allele) and P_A the

TABLE 1

Genotypic control of selection

Endosperm ploidy level	Endosperm genotype (m:p)	Embryo's W	Model
n (1m:0p)	A: $\emptyset$	1	DF
n (1m:0p)	a: $\emptyset$	$1 - s$	DF
$2n$ (1m:1p)	A: A	1	DF, MCD
$2n$ (1m:1p)	A: a	$1 - hs$	DF, MCD
$2n$ (1m:1p)	a: a	$1 - s$	DF, MCD
$3n$ (2m:1p)	AA: A	1	MCD
$3n$ (2m:1p)	AA: a	$1 - h_2 s$	MCD
$3n$ (2m:1p)	aa: A	$1 - h_1 s$	MCD
$3n$ (2m:1p)	aa: a	$1 - s$	MCD

We assume $h_2 < h$ and $h_2 < h_1$, $0 \leq h \leq 1$, $0 \leq s \leq 1$.

frequency of A (the wild-type allele). The embryos associated to a nutritive tissue that carries only the wild-type allele (A, AA, AAA) have a survival of 1. Those associated to a nutritive tissue that carries only the deleterious allele (a, aa, aaa) have a survival of $1 - s$. The embryos associated to a heterozygous nutritive tissue have survival $1 - hs$, $1 - h_1 s$, and $1 - h_2 s$ for Aa, aaA, and AAa genotypes, respectively. The parameters h, h_1, and h_2 are the coefficients of dominance of a on A. We consider $h_1 \geq h_2$ (a alleles are more exposed in aaA endosperms than in aAA endosperms). Because we assume that $W_a = W_{aa} = W_{aaa}$ and $W_A = W_{AA} = W_{AAA}$, our model excludes potential effects of gene copy number (*i.e.*, dosage effects in "metabolic theory"). Because we do not introduce any competition between embryos or endosperms produced by the same sporophyte (the viability of a given embryo depends only on the genotype at the viability locus of its own endosperm), our model excludes any effect involved in the allocation conflict theories. In fact, our model can be understood as if there was a single female gametophyte produced by each sporophyte. Thus, by construction, our model focuses exclusively on the heterozygosity theory.

Modifier locus: The second locus controls the genetic constitution of the nutritive tissue. More specifically, it controls the rate of double fertilization (*i.e.*, 1m:0p *vs.* 1m:1p, in the DF model) or the rate of doubling of maternal contribution (*i.e.*, 1m:1p *vs.* 2m:1p, in the MCD model). We consider two alleles M and m with frequency P_M and P_m, respectively. In each model, the modifier locus is expressed in the pollen (paternal control), the embryo sac (maternal control), or both (biparental control). In the first two cases, the expression of the modifier locus is haploid. In these cases, we denote ν_M or ν_m the probability of double fertilization (or maternal contribution doubling) when the gametophyte controlling the trait carries the M or the m allele, respectively. In the biparental case, we assume that the modifier effect is additive (*i.e.*, the probabilities of DF and MCD are ν_M, $(\nu_M + \nu_m)/2$, and ν_m when the embryo is MM, Mm, and mm, respectively). To give a concrete example, in a DF model with a paternal control of the modifier locus, the mating between AM male gametophytes (pollen) and am female gametophytes will produce a proportion ν_M of embryos whose nutritive tissue is 1m:1p (Aa, with fitness $1 - hs$) and a proportion $1 - \nu_M$ embryos whose nutritive tissue is 1m:0p (a, with fitness $1 - s$). Tables 1 and 2 summarize all these notations. Without loss of generality, we discuss the fate of modifier alleles that increase the probabilities of DF and MCD ($\nu_m > \nu_M$).

Recursion equations: The recursion equations describing the dynamics of the model are given in APPENDIX A for the case of the evolution of DF with a paternal control. These recursions can be written for the frequency of the four haplotypes (p_{aM}, p_{Am}, p_{AM}, and p_{am}). Equivalently, they can be expressed for the allele frequencies (p_m and p_a) and linkage disequilibrium ($C_{am} = p_{AM}p_{am} - p_{aM}p_{Am} = p_{am} - p_a p_m$). In this case, we define C_{am} such that it is positive when the modifier allele (m) is positively associated to the deleterious allele at the viability locus (a). To fully understand the model, we also describe the population during the diploid phase. In diploids, several genetic associations $C_{U,V}$ analogous to C_{am} can be defined (as in, *e.g.*, BARTON and TURELLI 1991), where U denotes a set of alleles on the maternally inherited chromosome and V a set of alleles in the paternally inherited chromosome. In our case, we have two biallelic loci and we have therefore 11 associations: $C_{a,m}$, $C_{m,a}$, $C_{am,\emptyset}$, $C_{\emptyset,am}$, $C_{a,a}$, $C_{m,m}$,

TABLE 2

Genotypic control of the genetic makeup of the endosperm

Control	Modifier locus genotype	Probability of: DF: double fertilization (= proportion of 1m:1p endosperms)	Probability of: MCD: doubling (= proportion of 1m:2p endosperms)
Biparental	MM	$\nu_M = \nu$	$\nu_M = \nu$
	Mm	$(\nu_M + \nu_m)/2 = \nu + \delta_\nu/2$	$(\nu_M + \nu_m)/2 = \nu + \delta_\nu/2$
	mm	$\nu_m = \nu + \delta_\nu$	$\nu_m = \nu + \delta_\nu$
Uniparental	M	$\nu_M = \nu$	$\nu_M = \nu$
	m	$\nu_m = \nu + \delta_\nu$	$\nu_M = \nu + \delta_\nu$

$C_{a,am}$, $C_{m,am}$, $C_{am,a}$, $C_{am,m}$, and $C_{am,am}$. For instance, the $C_{a,m}$ association measures the fact that the presence of the a allele on the maternally inherited chromosome covaries with the presence of the m allele on the paternally inherited chromosome.

ANALYSIS

We analyze DF and MCD models with two complementary methods: quasi-linkage equilibrium, (QLE) and stability analysis (see, *e.g.*, Otto and Day 2007). In these two methods, we assume that the frequency of deleterious mutations p_a is close to its equilibrium value (noted p_a^*) and small (which entails that C_{am} is also small). This assumption requires that $\mu \ll s$. This is the only common assumption to the two approaches. In the QLE approach, we also assume that s is small, R is large, and the modifier has a weak effect. In the stability analysis, the only additional assumption to $\mu \ll s$ is that the modifier frequency is small. To easily define the orders of magnitude, we use a small parameter ξ as a scaling throughout.

The QLE analysis assumes that linkage disequilibrium equilibrates much faster than allele frequencies, so that it can be considered to always be at a "quasi-equilibrium." This approach works best if the modifier effect is small and the recombination rate large (that is why we consider $\nu_M = \nu$ and $\nu_m = \nu + \delta_\nu$, where δ_ν is proportional to ξ; in the case of a modifier with a specifically maternal or paternal effect, we write $\delta_{\nu,\emptyset}$ and $\delta_{\emptyset,\nu}$ instead of δ_ν, respectively). With this approach, we focus on convergence stability. We perform the stability analysis to assess the robustness of this QLE analysis and to investigate what is going on for low recombination rate. This method considers the fate of the modifier introduced at a low frequency in a population at equilibrium at the viability locus. We first start by studying the DF model.

Mutation–selection at the viability locus in the DF model: To start, it is important to understand what is going on at the viability locus alone. According to the assumptions given above for the QLE approach, the change over one generation (noted with Δ) of the frequency of the deleterious allele at the viability locus is

$$\Delta p_a = \mu - p_a\left(\frac{1}{2}(1-\nu) + h\nu\right)s + O(\xi)^3, \qquad (1)$$

and thus, at equilibrium ($\Delta p_a = 0$), the frequency of the a allele (noted with an asterisk) depends on the rate of double fertilization (ν):

$$p_a^* = \frac{\mu}{((1/2)(1-\nu) + h\nu)s} + O(\xi)^2. \qquad (2)$$

The same equations are obtained under the assumption of the stability analysis, except that the orders of the two equations are $O(\xi)^2$ and $O(\xi)$, respectively. With complete double fertilization ($\nu = 1$), the endosperm is diploid and we retrieve the equilibrium frequency of a deleterious mutation under diploid selection (μ/hs). However, with no double fertilization ($\nu = 0$), $p_a^* = 2\mu/s$: the efficiency of selection is halved compared to the classic haploid selection situation (μ/s). Indeed, half the alleles (the paternal ones) are not exposed to selection since they are in the embryo, but not in the endosperm. Interestingly, this result points out that having a haploid endosperm is not equivalent to haploidy. Note that the equilibrium described above is a good approximation only if $h > s$ when ν is close to 1. A more complicated approximation can be found for small values of h and $1 - \nu$ (see Chasnov 2000 for a similar treatment), but, as explained below, a better approximation is not necessary to accurately predict how DF and MCD evolve in our models.

QLE analysis of the DF model: We compute the overall frequency change at the modifier locus Δp_m as a sum of two terms. The first, $\Delta p_m^{\text{direct}}$, is proportional to $p_m(1 - p_m)$ and corresponds to direct selection. The second, $\Delta p_m^{\text{indirect}}$, is proportional to C_{am} and corresponds to indirect selection. Under the QLE assumptions given above, we obtain at leading orders

$$\begin{aligned}
\Delta p_m &= \Delta p_m^{\text{direct}} + \Delta p_m^{\text{indirect}} \\
\Delta p_m^{\text{direct}} &= s\delta_\nu\left(\frac{1}{2} - h\right)p_m(1-p_m)p_a(1-p_a) + O(\xi)^5 \\
\Delta p_m^{\text{indirect}} &= -sC_{am}\left(\frac{1}{2}(1-\nu) + h\nu\right) + O(\xi^2 C_{am}). \qquad (3)
\end{aligned}$$

We study these terms separately for two reasons. First they represent qualitatively different processes. Second, their relative magnitude varies in different ways throughout the parameter range (*e.g.*, we expect indirect selection to become increasingly important as R decreases and direct selection to vanish when h becomes close to $\frac{1}{2}$), and it is therefore quantitatively useful to combine their leading orders.

Direct selection: The sign of $\Delta p_m^{\text{direct}}$ is given by the sign of $\frac{1}{2} - h$. When mutations are recessive ($h < \frac{1}{2}$), direct selection favors the evolution of DF. In fact, h can be understood as the "exposition" (to selection) of the a allele in a 1m:1p seed, while $\frac{1}{2}$ can be understood as its exposition in a 1m:0p seed. Indeed, as said before, only half the a alleles (the maternally inherited ones) are exposed to selection in 1m:0p endosperms. Direct selection simply tends to minimize the exposition of the deleterious alleles. This effect is quite different from the "masking effect" described in models of ploidy evolution (Otto and Goldstein 1992). In these models, exposition of the deleterious alleles is always greater with haploidy (Aa individuals have a better survival than haploid a individuals as long as $h < 1$), but this is compensated by the fact that a diploid individual has twice the chance to carry a deleterious mutation (it has two alleles). As a consequence, diploidy is favored by direct selection when exposition (h) is at least halved to compensate for this doubled risk of inheriting a

deleterious mutation (*i.e.*, when $h < \frac{1}{2}$). In contrast, in our model, all individuals have the same risk to carry a deleterious mutation (all embryos are diploid), but the exposition of the deleterious alleles is greater with 1m:0p only if $h < \frac{1}{2}$ (because only half the alleles, the maternal ones, are exposed to selection without DF).

Indirect selection: The sign of $\Delta p_m^{\text{indirect}}$ directly depends on the sign of the linkage disequilibrium C_{am}. To evaluate indirect selection more precisely, it is useful to express the value of C_{am} in terms of the model's parameters. Under the QLE assumptions, C_{am} reaches a pseudoequilibrium value (C_{am}^{QLE}) that is obtained by solving $\Delta C_{am} = 0$, yielding

$$C_{am}^{\text{QLE}} = \delta_v \frac{s(K-h)}{2R} p_m(1-p_m)p_a(1-p_a) + O(\xi)^4, \quad (4)$$

where $K = 1 - R$, R, and $\frac{1}{2}$ for maternal, paternal, or biparental expression of the modifier locus, respectively. Thus, the sign of C_{am}^{QLE} depends entirely on the sign of $K - h$, that is, of $1 - R - h$, $R - h$, and $\frac{1}{2} - h$, for maternal, paternal, and biparental expression of the modifier locus, respectively. This result can be interpreted by following step by step how the linkage disequilibrium is built through selection and meiosis.

Building of genetic associations by selection: A lower exposition of deleterious alleles is immediately beneficial to an individual but leads to their accumulation. More specifically, in diploid embryos, deleterious mutations tend to become associated to modifier alleles that minimize exposition. A paternally acting modifier exposes the deleterious allele carried on the paternal chromosome (exposition is increased from 0 to h in the case of DF), which enhances the purge of these paternally inherited chromosomes. As a consequence, a paternally acting modifier tends to become associated to less loaded paternal chromosomes ($C'_{\emptyset,am} - C_{\emptyset,am} < 0$, where the prime indicates that the association is measured after selection):

$$C'_{\emptyset,am} = C_{\emptyset,am} + \delta_{\emptyset,v}(0-h)s\,p_m(1-p_m)p_a(1-p_a) + O(\xi)^4. \quad (5)$$

On the contrary, a paternally acting modifier always reduces the exposition of the deleterious alleles carried on the maternal chromosome (from 1 to h when there is DF), which reduces purging on these chromosomes and generates more loaded maternal chromosomes ($C'_{a,m} - C_{a,m} > 0$):

$$C'_{a,m} = C_{a,m} + \delta_{\emptyset,v}(1-h)s\,p_m(1-p_m)p_a(1-p_a) + O(\xi)^4. \quad (6)$$

The exact opposite occurs for a maternally acting modifier ($C'_{am,\emptyset} - C_{am,\emptyset} > 0$ and $C'_{m,a} - C_{m,a} < 0$):

$$C'_{am,\emptyset} = C_{am,\emptyset} + \delta_{v,\emptyset}(1-h)s\,p_m(1-p_m)p_a(1-p_a) + O(\xi)^4 \quad (7)$$

$$C'_{m,a} = C_{m,a} + \delta_{v,\emptyset}(0-h)s\,p_m(1-p_m)p_a(1-p_a) + O(\xi)^4. \quad (8)$$

All of these effects occur when the modifier has an effect in both male and female gametophytes (*i.e.*, in the biparental case).

Effect of meiosis: The haploid linkage disequilibrium after meiosis (C'_{am}) results from the diploid associations as

$$C'_{am} = \frac{C'_{am,\emptyset} + C'_{\emptyset,am}}{2}(1-R) + \frac{C'_{a,m} + C'_{m,a}}{2}R \quad (9)$$

(Barton and Turelli 1991); *i.e.*, the diploid *cis*-associations are carried over at a rate $1 - R$ while the diploid *trans*-associations are transformed into C_{am} at a rate R. In particular, as the recombination rate becomes lower, *cis*-associations contribute more to generating C_{am}. Combining Equations 5–9, one easily retrieves Equation 4. For example, with a paternally acting modifier, the *cis*-association $C_{\emptyset,am}$ is proportional to $-h$, while the *trans*-association $C_{a,m}$ is proportional to $1 - h$. Because recombination breaks up *cis*-associations and transforms *trans*-associations into haploid *cis*-associations (Equation 9), the modifier stays overall associated to purged genomes (and favored by indirect selection) when $(1-h)R - h(1-R) < 0$, *i.e.*, when $R < h$ (Equation 4, Figure 4A). In the case of a maternally acting modifier, the roles of *cis*- and *trans*-associations (and thus the effect of recombination) are reversed. The modifier stays overall associated to purged genomes when $(1-h)(1-R) - hR < 0$, *i.e.*, when $1 - R < h$ (Equation 4, Figure 4B). In the biparental case, the four associations $C_{am,\emptyset}$, $C_{\emptyset,am}$, $C_{a,m}$, $C_{m,a}$ play a role and the effect of recombination cancels out. The modifier stays overall associated to purged genomes when $(1-R)(1-h) - hR + R(1-h) - h(1-R) < 0$, *i.e.*, when $h > \frac{1}{2}$ (Equation 4).

Overall selection on the modifier: To obtain the selection coefficient on the modifier [defined as $s_m \equiv \Delta p_m / p_m(1-p_m)$], we replace C_{am} by C_{am}^{QLE} in Equation 3, yielding

$$s_m = s\delta_v p_a(1-p_a) \times \left[\left(\frac{1}{2} - h\right) - \frac{s(K-h)}{2R}\left(\frac{1}{2}(1-v) + hv\right)\right] + O(\xi)^5. \quad (10)$$

The sign of s_m depends on the term within the brackets. In Figure 4, we illustrate the conditions where s_m is positive. When $R = \frac{1}{2}$, $|C_{am}|$ is low: indirect selection is negligible and the sign of s_m is given by $\frac{1}{2} - h$. In this case two equilibria ($v = 0$ and $v = 1$) exist, and DF is favored when $h < \frac{1}{2}$. When R is lower, indirect selection can counterbalance direct selection. With a paternally acting modifier, DF is favored from $h = 0$ up to $h_{\text{pat}}^{*\text{DF}} = \frac{1}{2} + H_{\text{crit}}^{\text{DF}}$, where

$$H_{\text{crit}}^{\text{DF}} = \frac{1}{2}\frac{s(1-2R)}{4R} + O\left(\frac{s}{R}\right)^2. \quad (11)$$

On the contrary, with a maternally acting modifier, DF is favored from $h = 0$ up to $h_{\text{mat}}^{*\text{DF}} = \frac{1}{2} - H_{\text{crit}}^{\text{DF}}$. In other words, DF is favored for $h < \frac{1}{2}$ when $R = \frac{1}{2}$. Otherwise ($R < \frac{1}{2}$), DF is favored for a range of h values wider than

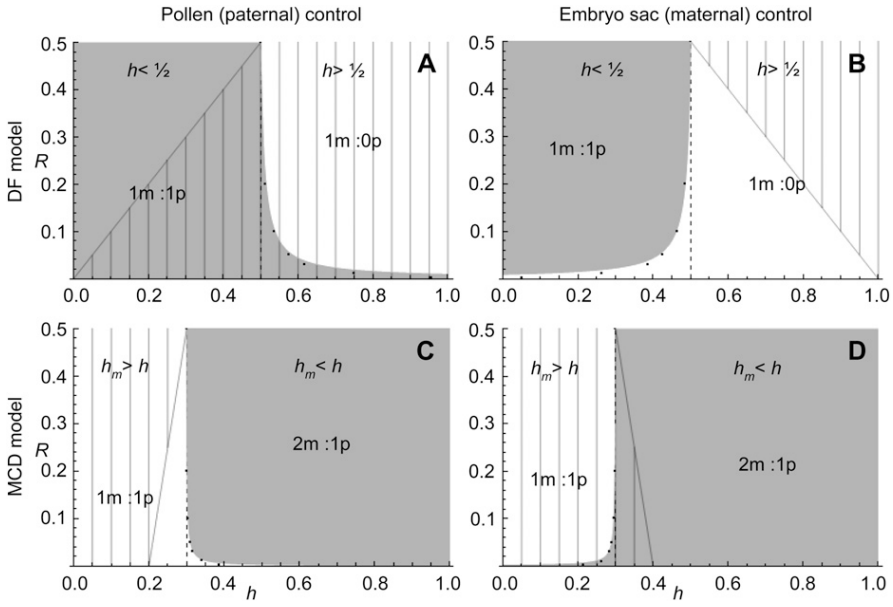

FIGURE 4.—Illustration of the QLE analysis of the DF and MCD models. The direction of selection at the modifier locus is indicated for different h (x-axis) and R (y-axis) values. In A and B (DF model), 1m:1p is favored in the shaded area whereas 1m:0p is favored in the open area. In C and D (MCD model), 2m:1p is favored in the shaded area whereas 1m:1p is favored in the open area. A, C and B, D illustrate paternal and maternal expression of the modifier locus, respectively. In all panels, the dashed area corresponds to a negative linkage disequilibrium (the modifier is associated to the purged chromosome). Direct selection favors DF on the left-hand side of the dashed line (which is $h = \frac{1}{2}$, A and B), whereas it favors MCD on the right-hand side of the dashed line (which is $h = h_m = 0.3$, C and D). The QLE approximation becomes inaccurate when R is very low (the points represent critical h values obtained directly from exact simulations). Parameter values are $s = 5 \times 10^{-2}$, $\mu = 10^{-3}$, $\upsilon = 0.1$, $\delta_\upsilon = 0.01$, $h_1 = 0.2$, and $h_2 = 0.4$.

$0 < h < \frac{1}{2}$ with a paternal expression of the modifier locus and for a range of h values narrower than $0 < h < \frac{1}{2}$ with a maternal expression of the modifier locus. Note that in both the maternal and the paternal case, an intermediate unstable equilibrium (ν different from zero or one) exists in a very small portion of the parameter space when h is between h^* and $h^* + \Delta H$, where

$$\Delta H = \frac{s^2(1-2R)^2}{32R^2} + O\left(\frac{s}{R}\right)^3. \qquad (12)$$

In this region, the evolutionary outcome depends on initial conditions. Above $h^* + \Delta H$, DF is disfavored. DF is selected in different conditions depending on whether the modifier locus is paternally or maternally expressed ($h^{*\mathrm{DF}}_{\mathrm{mat}} \neq h^{*\mathrm{DF}}_{\mathrm{pat}}$); it is a case of male/female conflict. With a biparentally expressed modifier, direct and indirect selection has always opposite signs, but indirect selection is never strong enough to counterbalance direct selection. Thus, the critical value of h below which DF is favored is $\frac{1}{2}$ regardless of other parameter values.

Stability analysis of the DF model: To check the QLE analysis for small R values, we perform a stability analysis. To do so, we consider a population where M is fixed ($p_{Am} = p_{am} = 0$) and at mutation–selection equilibrium at the viability locus ($p_{aM} = p^*_a$, Equation 2). We then compute the frequency change of a rare modifier allele introduced in this population. Technically, we linearize the exact recursions on the Am and am haplotypes, assuming they are rare. We then find the dominant eigenvalue of the transition matrix corresponding to these two equations, yielding

$$\lambda = 1 + \frac{\mu(\nu_M - \nu_m)(4R(h-(1/2)) + sK_1K_2)}{(1-\nu_M + 2h\nu_M)(2R + s(1-R)K_1)} + O(\xi)^2, \qquad (13)$$

where K_1 and K_2 depend on whether the modifier is maternally, paternally, or biparentally expressed. Their values are given in Table 3. Assuming that mutations are rare, the rate of DF will evolve through successive fixations in the direction predicted by the sign of $\lambda - 1$. The results of the stability analysis are qualitatively comparable to those of QLE. The critical value of h below which DF is favored predicted by this analysis is close to the critical value predicted under QLE assumptions (h^*_{pat}, h^*_{mat}). In fact, the absolute difference between the critical h values predicted by the two approaches is $s\delta_\nu(1/8 - 1/16R) + O(s)^2$, which is negligible when R is large, as expected. In the stability analysis, as in the QLE, an intermediate unstable equilibrium exists in a small portion of the parameter space. In this region, both full and no DF are locally convergent stable states (the outcome depends on initial conditions). The area of this region is very small, $\sim s/32$ (if one considers that the parameter space is $h \in [0, 1]$ and $R \in [0, 0.5]$). This small region is illustrated in Figure 5. In a slightly larger region of the parameter space around the latter, a single convergent stable state is present, but is not stable faced with mutations of large effects. In all cases, the results between QLE and stability analysis are very comparable and qualitatively identical. We now switch to the MCD model.

MCD model analysis: We analyze this model in the same way as the DF model, using the same assumptions.

Mutation–selection at the viability locus: The mutation–selection equilibrium at the viability locus is

TABLE 3

Values of K_1 and K_2 in the stability analysis for each model

	Maternally acting modifier	Paternally acting modifier	Biparentally acting modifier
DF			
K_1	$1-(1-h)\nu_m+h\nu_M$	$1-(1-h)\nu_M+h\nu_m$	$1-(1-h)\bar{\nu}+h\bar{\nu}$
K_2	h	$(h-1)$	$(h-1/2)(1-2R)$
MCD			
K_1	$-\frac{1}{2}((h_1-h_2)/2)(\nu_M-\nu_m)$	$\frac{1}{2}((h_1-h_2)/2)(\nu_M-\nu_m)$	0
K_2	$-((h_1-h_2)/2)$	$-((h_1-h_2)/2)$	0

$\bar{\nu}=(\nu_M+\nu_m)/2$

$$p_a^*=\frac{\mu}{(h(1-\nu)+h_m\nu)s}+O(\xi)^2, \tag{14}$$

where $h_m=(h_1+h_2)/2$. This value is comparable to the equilibrium frequency in the DF model (Equation 2). Indeed, in the DF model, the exposition of the *a* allele ranges from h (full DF) to $\frac{1}{2}$ (no DF). Similarly, in the MCD model, the exposition of the *a* allele ranges from h_m (full MCD) to h (no MCD).

Frequency change at the modifer locus: The frequency change at the modifier locus is determined by direct and indirect selection, as in Equation 3:

$$\begin{aligned}\Delta p_m &= \Delta p_m^{\text{direct}}+\Delta p_m^{\text{indirect}}\\ \Delta p_m^{\text{direct}} &= s\delta_\nu(h-h_m)p_m(1-p_m)p_a(1-p_a)+O(\xi)^5\\ \Delta p_m^{\text{indirect}} &= -sC_{am}(h(1-\nu)+h_m\nu)+O(\xi^2C_{am}).\end{aligned} \tag{15}$$

The sign of direct selection is given by the sign of $h-h_m$; *i.e.*, direct selection favors the state where the average exposition of deleterious alleles is maximal. This term favors MCD when $h_m<h$. The sign of indirect selection depends directly on the sign of C_{am}, whose QLE value has the same form as in Equation 4,

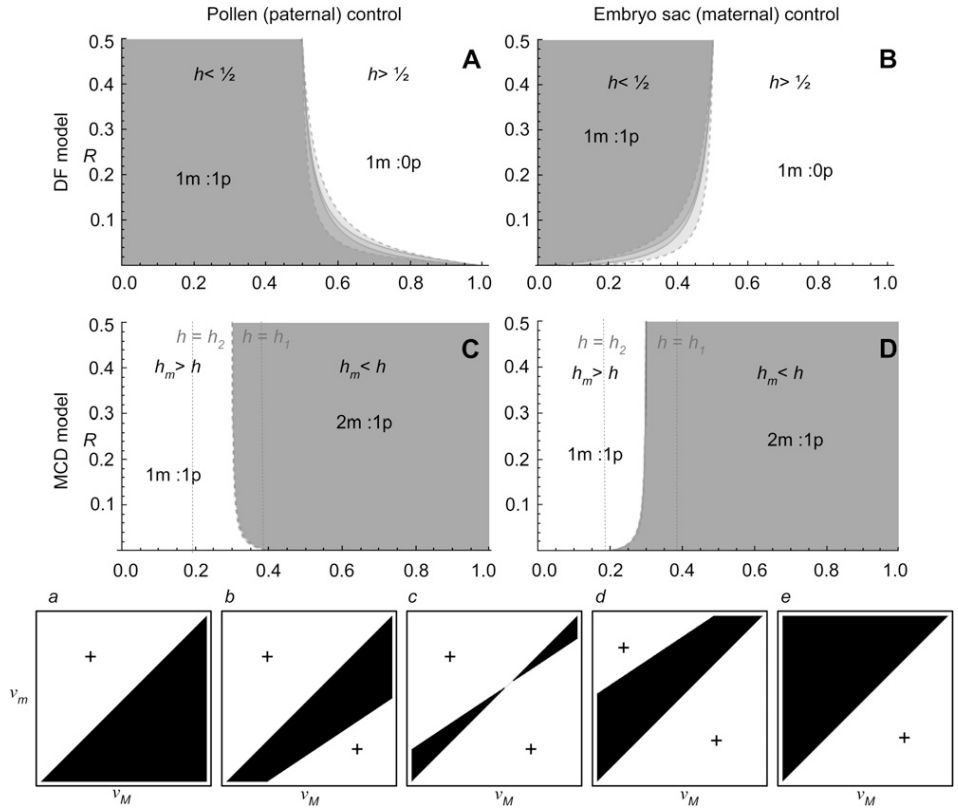

FIGURE 5.—Illustration of the stability analysis of DF and MCD models. The direction of selection at the modifier locus is indicated for different dominance (h, x-axis) and recombination values (R, y-axis). In A and B (DF model), 1m:1p is favored in the area with dark shading [that corresponds to the pairwise invasibility plot (PIP) a], whereas 1m:0p is favored in the open area (that corresponds to PIP e). The areas with intermediate shading correspond to b, c, and d, that is, to cases where (b) DF is only convergent stable, (c) an intermediary, unstable equilibrium exists, and (d) DF is not convergent stable. In C and D (MCD model), 2m:1p is favored in the shaded area whereas 1m:1p is favored in the open area. We can see that the intermediate cases are indistinguishable for the MCD model. The parameters values are $h_1=0.2$, $h_2=0.4$, $\mu=10^{-3}$, and $s=10^{-1}$.

$$C_{am}^{\mathrm{QLE}} = \delta_{\nu}\frac{s(K+h-h_m)}{2R}p_m(1-p_m)p_a(1-p_a) + O(\xi)^4, \tag{16}$$

where $K = 0$, $(1-2R)((h_2-h_1)/2)$, and $(1-2R)((h_1-h_2)/2)$ for biparental, paternal, and maternal expression of the modifier locus, respectively. In the very likely situation where $h_1 > h_2$, K is positive for a maternally expressed modifier and negative for a paternally expressed one. Because the sign of C_{am} depends entirely on the sign of $K + h - h_m$, it is therefore more easily positive with a maternal expression of the modifier locus. The way *cis*- and *trans*-associations are built by selection is qualitatively the same as in the DF model:

$$C'_{\emptyset,am} = C_{\emptyset,am} + \delta_{\emptyset,\nu}(h-h_1)s\,p_a p_m(1-p_m)(1-p_a) + O(\xi)^4 \tag{17}$$

$$C'_{a,m} = C_{a,m} + \delta_{\emptyset,\nu}(h-h_2)s\,p_a p_m(1-p_m)(1-p_a) + O(\xi)^4 \tag{18}$$

$$C'_{am,\emptyset} = C_{am,\emptyset} + \delta_{\nu,\emptyset}(h-h_1)s\,p_a p_m(1-p_m)(1-p_a) + O(\xi)^4 \tag{19}$$

$$C'_{m,a} = C_{m,a} + \delta_{\nu,\emptyset}(h-h_2)s\,p_a p_m(1-p_m)(1-p_a) + O(\xi)^4. \tag{20}$$

An important difference between DF and MCD is that MCD can increase *or* decrease the exposition of maternally inherited deleterious alleles (depending on the value of h_1, the exposition of the *a* allele in *aaA* endosperms) and can increase *or* decrease the exposition of paternally inherited deleterious alleles (depending on the value of h_2, the exposition of the *a* allele in *AAa* endosperms). For instance, a paternally acting modifier increasing MCD decreases the exposition of the deleterious allele carried on the paternal chromosome from h to h_1, if $h > h_1$, but increases it if $h < h_1$. In the first case, the paternally acting modifier tends to become associated to more loaded paternal chromosomes ($C'_{\emptyset,am} - C_{\emptyset,am} > 0$), while in the second case it tends to become associated to less loaded chromosomes ($C'_{\emptyset,am} - C_{\emptyset,am} < 0$). Combining Equations 17–20 with Equation 9, one easily retrieves the value of C_{am}^{QLE}. With uniparental expression of the modifier locus, three configurations can occur. In the case where $h_2 < h < h_1$, the situation is "reversed" compared to what happens in the DF model, because in these conditions MCD increases the maternal allele's exposition and decreases the paternal allele's one. For instance, for a paternally acting modifier, the *cis*-association $C_{\emptyset,am}$ is proportional to $h - h_2$ (positive), while the *trans*-association $C_{a,m}$ is proportional to $h - h_1$ (negative). The modifier stays overall associated to purged genomes when $(h-h_2)(1-R) - (h-h_1)R < 0$, *i.e.*, when $R < (h-h_2)/(2h-h_1-h_2)$ (Figure 4C). For a maternally acting modifier, the modifier stays overall associated to purged genomes when $(h-h_1)(1-R) - (h-h_2)R < 0$, *i.e.*, when $R < (h-h_1)/(2h-h_1-h_2)$ (Figure 4D). In the case of biparental expression of the modifier locus, the effect of recombination cancels out as in the DF model, and the modifier stays overall associated to purged genomes when $h_m > h$. In the second case, where $h_2 < h_1 < h$, both maternal and paternal alleles have a better exposition at the 2m:1p state. The haploid linkage disequilibrium C_{am} is always negative. On the contrary, in the third case, where $h < h_2 < h_1$, both maternal and paternal alleles have a better exposition at the 1m:1p state and C_{am} is always positive.

Overall selection on the modifier: The selection coefficient at the modifier locus at QLE is of the same form as Equation 10:

$$s_m = s\delta_{\nu}p_a(1-p_a)\times\left[(h-h_m) - s\frac{(K+h-h_m)}{2R}(h(1-\nu)+h_m\nu)\right] + O(\xi)^5. \tag{21}$$

In Figure 4, we illustrate the conditions where s_m is positive. As in the DF model, there is in most cases only one convergent stable state (either full or no MCD). However, in a region of the parameter space full and no MCD may be both convergent stable states and their domain of attraction is determined by an unstable intermediate equilibrium. As in the DF model, this region is, however, very small. When $R = \frac{1}{2}$, the sign of s_m simply depends on the sign of $h - h_m$. When R is smaller, indirect selection can counterbalance direct selection: with a paternally acting modifier, MCD is favored from $h = 1$ down to $h_{\mathrm{pat}}^{*\mathrm{MCD}} = h_m + H_{\mathrm{crit}}^{\mathrm{MCD}}$, where

$$H_{\mathrm{crit}}^{\mathrm{MCD}} = h_m\frac{s(1-2R)}{4R}(h_1-h_2) + O\left(\frac{s}{R}\right)^2, \tag{22}$$

which is narrower than from $h = 1$ down to $h = h_m$. On the contrary, with a maternally acting modifier, MCD is favored from $h = 1$ to $h_{\mathrm{mat}}^{*\mathrm{MCD}} = h_m - H_{\mathrm{crit}}^{\mathrm{MCD}}$, which is wider than from $h = 1$ to $h = h_m$. Like DF, MCD may be selected or counterselected depending on whether the modifier locus is paternally or maternally expressed. When the recombination rate is low, there is a male/female conflict regarding MCD evolution. With a biparentally expressed modifier, indirect selection is never strong enough to oppose direct selection: MCD is favored when $h_m < h$.

Stability analysis: We make a stability analysis as for the DF model. The dominant eigenvalue of the matrix is

$$\lambda = 1 + \frac{\nu_M - \nu_m}{(1-\nu_M)h + \nu_M h_m} \times \frac{R(h_m-h) + s(1/2-R)(h_m(1-\bar{\nu}) + h\bar{\nu} + K_1)(h_m - h + K_2)}{R + s(1-R)(h_m(1-\bar{\nu}) + h\bar{\nu} + K_1)}\mu + o(\xi), \tag{23}$$

where $\bar{\nu} = (\nu_M + \nu_m)/2$ and where K_1 and K_2 depend on whether the modifier is maternally, paternally, or biparentally expressed. Their values are given in Table 3.

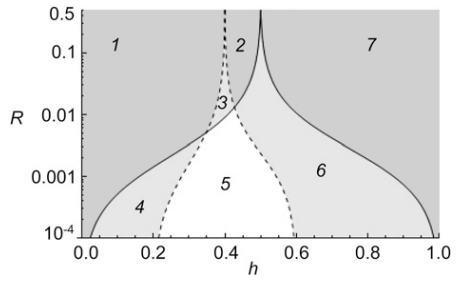

FIGURE 6.—Combined evolution of DF and MCD. The different possible scenarios for the combined evolution of DF and MCD are shown when $h_m < \frac{1}{2}$. ($h_1 = 0.6$ and $h_2 = 0.2$ in this example, yielding $h_m = 0.4$). The solid lines indicate the critical maternal and paternal h values for DF model. The dashed lines indicate the critical maternal and paternal h values for MCD model. In area 1, DF is favored, and MCD is not (1m:1p is thus expected to evolve). In area 2, both DF and MCD are favored (2m:1p evolves). In area 3, DF is favored and MCD is favored only if under maternal control (thus either 1m:1p or 2m:1p may evolve depending on which sex predominantly controls MCD). In area 4, DF is favored only if under paternal control, and MCD is not favored (thus either 1m:0p or 1m:1p may evolve). In area 5 DF is favored only if under paternal control, while MCD is favored only if under maternal control, and thus 1m:0p, 1m:1p, or 2m:1p may be favored. In area 6 DF is favored only if under paternal control and MCD is always favored (either 1m:1p or 2m:1p may evolve). Eventually, in area 7 neither DF nor MCD is favored, and 1m:0p evolves. The figure is drawn using the results of the stability analysis and with $\mu = 10^{-3}$, $s = 10^{-1}$, $\nu_M = 0.1$, and $\nu_m = 0.1$.

The direction of selection is illustrated in Figure 4. As in the DF model, the QLE and the stability analysis yield comparable results. The absolute difference between the critical h values predicted by the two approaches is $s\delta_\nu(h_1 - h_2)^2(1/8 - 1/16R) + O(s)^2$, which is negligible when R is large, as expected. As in the DF model, there is also a small region of the parameter space with an intermediate unstable equilibrium, but it is so small that it can hardly be seen in Figure 5.

Combining the evolution of DF and MCD: So far, we have described the conditions favoring DF and MCD independently. We now consider how these conditions combine to give an overall picture of the evolution of the different types of endosperms (1m:0p, 1m:1p, and 2m:1p). There are a large number of possibilities depending on the relative values of h_1, h_2, and h. Considering that $h_2 < h < h_1$ [as one would expect following the classical metabolic theory of dominance (KACSER and BURNS 1981)] and that $h_m < \frac{1}{2}$, we can delineate seven cases. They are illustrated in Figure 6. However, all seven cases are not equally plausible. In particular, h is very likely to be $< \frac{1}{2}$ since the deleterious mutations have always been found to be recessive on average in various diploid taxa (see, *e.g.*, GARCÍA-DORADO and CABALLERO 2000; VASSILIEVA *et al.* 2000; SZAFRANIEC *et al.* 2003). With this additional condition, only six cases are possible (all but case 7 in Figure 6). With a high recombination rate (*i.e.*, from the conditions where a male/female conflict occurs), only two evolutionary outcomes are possible: 1m:1p if $h < h_m$ and 2m:1p if $h > h_m$ (cases 1 and 2 in Figure 6). For lower recombination rates, four other possibilities can also occur, including the 1m:0p case (see cases 3–6 in Figure 6).

Robustness of the results: These two-locus models are simplified ways to investigate the effect of masking on DF and MCD evolutions. Integrating the overall selection coefficient over a full genome with many viability loci might be achieved by summing the frequency change of the modifier caused independently by each of them (see, *e.g.*, OTTO 2003). This approach is, however, complicated by three factors. First, one would have to integrate over an explicit genetic map (to account for different R values). Second, different loci may exhibit different parameter values (h, h_1, h_2, s, μ) and the distributions of these parameter values are not well known (including possible correlations between those parameters, for instance, s and h). Third, selective interactions may occur among viability loci and our models do not explore these effects. Finally, we do not incorporate departure from random mating that might affect the conclusions. Future work is required to account for these complications.

DISCUSSION

In this article we modeled the masking of deleterious mutations in the endosperm and its influence on the evolution of DF and MCD. More specifically, we followed the fate of alleles modifying the probability of occurrence of DF or MCD at a modifier locus with a sex-limited expression. In our models, a second locus is expressed in the endosperm and controls the amount of resources available to the embryo. At this locus, deleterious alleles diminish the ability to accumulate resources, and the seeds inheriting this allele have therefore a lower survival rate. In these models, there is no resource reallocation between the different embryos produced by the same mother plant. Thus, all of our results are unrelated to theories based on competition between embryos or kin selection and are attributable only to masking and its possible side effects.

Our work is the first formal demonstration that the masking of deleterious mutations is sufficient for DF and MCD to evolve. In fact, two different effects combine in this theory: first, a direct effect of masking, which depends only on the dominance level of deleterious mutations, and second, an indirect and side effect of masking, proportional to linkage disequilibrium, the value of which depends on whether the expression of the modifier locus is paternal or maternal. The latter effect depends mainly on the values of dominance (h_1, h_2, and h) and recombination rate. When the recombination rate is high, the indirect effect is always negligible and evolution favors the state with highest masking. Contrary to what is suggested in the literature (WILLSON and BURLEY 1983), the 1m:1p state is not necessarily the

situation with the strongest masking: for instance, masking can be stronger in 2m:1p than in 1m:1p, when $h_m < h$. When recombination is low, the indirect effect can become predominant and the outcome depends on whether DF or MCD are paternally, maternally, or biparentally controlled.

A new type of male/female conflict: Our results indicate that for low recombination rates, there is a male/female conflict for the genetic constitution of the endosperm. This male/female conflict results from the fact that increasing the exposition of alleles inherited from a given parent diminishes the exposition of alleles inherited from the other one. In the case of paternal expression of the modifier locus, indirect selection increases the range of h values where DF evolves, while it reduces the range of h_m values where MCD evolves. Exactly the opposite occurs with maternal expression of the modifier locus. With biparental expression of the modifier locus, the indirect effect is always negligible even when recombination is low. For instance, with $R < \frac{1}{2}$ and $h = 0.5$, a paternal but not a maternal DF allele would spread (see Figures 4 and 5). Similarly, with $R < \frac{1}{2}$ and $h = h_m$, a maternal but not a paternal MCD allele would spread (see Figures 4 and 5).

Evolutions of DF and MCD: Both 2m:1p and 1m:1p endosperms can be found in angiosperms. How can such a diversity be explained? Deleterious mutations tend to be recessive in diploids, with average h ranging between 0.1 and 0.4 depending on the studies (see, *e.g.*, García-Dorado and Caballero 2000; Vassilieva *et al.* 2000; Szafraniec *et al.* 2003). As a consequence, we would expect that paternally expressed genes favoring DF should always spread or be maintained. This is at least the case for the genes that determine the occurrence of two mitotically derived spermatozoa, since this characteristic, which is a necessary condition for DF evolution, is maintained in all angiosperms (but intriguingly also in gymnosperms). We would also expect that maternally expressed genes favoring DF should spread provided that the average recombination rate is not too low. In other words, for low recombination rates, we do not expect DF to be systematically maintained (depending on which sex controls the trait, see Figure 6). Among angiosperms with postfertilization resource allocation by the female sporophyte, few have lost DF or at least a functional endosperm, the only ones being Podostemaceae and Trapaceae (Chopra and Sachar 1963). These species exhibit high selfing rates (Haig 1988). Including selfing effects in our models will be the purpose of a future study. The Orchidaceae is the other important group where DF occurs but no functional endosperm is present (Chopra and Sachar 1963). In this group, however, there are virtually no resources allocated by the female sporophyte to the seed: the seedlings acquire resources through mycorrhiza. Thus, the degeneration of endosperm in orchids may be unrelated to DF or MCD evolution.

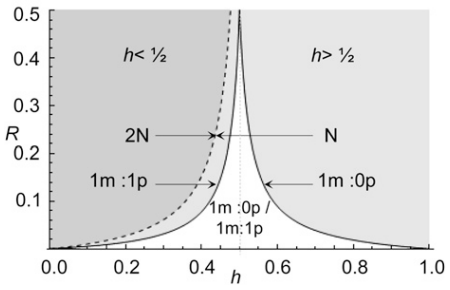

Figure 7.—Comparison of the conditions for the evolution of diploidy and DF. The overlap of the conditions of h and R is shown, where diploidy ($2N$) is favored by the conditions of h and R where DF is favored. Indeed, $2N$ is favored on the left of the dashed line, while DF is favored on the left of the solid lines (the left solid line corresponds to the case of a maternal expression of the modifier locus and the right solid line corresponds to the case of a paternal expression of the modifier locus). The figure is drawn using the results of the stability analysis and with $s = 10^{-1}$, $\mu = 10^{-3}$, $\nu_M = 0.1$, and $\nu_m = 0.1$.

Compared to diploids, little is known about the typical dominance of deleterious mutations in triploids (*i.e.*, the dominance h_1 of the *aaA* genotype and the dominance h_2 of the *AAa* genotype). It seems, however, likely the average dominance in triploids is close to that in diploids [*i.e.*, $(h_1 + h_2)/2 = h_m \approx h$], so that we would expect to be often close to the situation where there is a male/female conflict over the evolution of MCD. In these conditions, we would expect to find both types of endosperm (1m:1p and 2m:1p, see Figure 6). Consistent with this prediction, some angiosperms [in Nymphaeaceae, Hydatellaceae, Cabombaceae, Illiciaceae, Schisandraceae, Austrobaileyaceae, and Oenothera (Maheshwari 1963; Friedman 2001a; Williams and Friedman 2002; Friedman and Ryerson 2009)] have a 1m:1p endosperm (DF, no MCD) and others a 2m:1p endosperm (other groups, both DF and MCD evolved).

DF and diploidy: It is worth comparing our models of DF evolution with models for the evolution of diploid life cycles (in particular, the model of Otto and Goldstein 1992, which is referred to hereafter as OG92 and is reinvestigated in appendix b). Following the OG92 model, the evolution toward diploid life cycles ($2N$) requires deleterious mutations to be recessive and recombination to be "sufficiently large." In fact, these conditions are close to the conditions favoring DF when the expression of the modifier locus is maternal. We can compare these conditions in Figure 7: all other parameter values being equal, the range of R and h values in which DF evolves is slightly wider, but comparable with the range of R and h values favoring $2N$. In the two models (DF and OG92), the selection coefficient acting on the modifier can be decomposed into two terms, a direct and an indirect effect that are, superficially at least, comparable. In both models, direct selection favors diploidy when $h < \frac{1}{2}$, but for different reasons. In the OG92 model, the exposition of the deleterious alleles is always higher with $1N$ (full exposition) than with $2N$

(exposition at a level h). In the DF model, only maternal alleles are exposed. Thus, the average exposition is greater with "haploidy" (1m:0p) only if $h < \frac{1}{2}$. As we have already mentioned, this is compensated by the fact that, unlike in the DF model, $2N$ individuals have twice the chance to carry a deleterious mutation compared to $1N$ individuals. A second important difference is that, in the DF but not in the OG92 model, the direction and strength of indirect selection differ for paternally or maternally expressed modifiers. In the OG92 model, *cis*-associations between $2N$ modifier and deleterious mutations are always positive (in other words, modifiers of haploidy are found more often on purged chromosomes). In contrast, because of the asymmetry inherent to the DF model, where haploidy necessarily concerns maternally inherited alleles at the viability locus, purging of deleterious alleles causes positive *cis*-association only with maternally expressed modifiers. In other words, maternally acting modifiers of haploidy are found more often on purged chromosomes, while paternally acting modifiers of haploidy are found more often on nonpurged chromosomes. Overall, $2N$ and OG92 models are in fact quite different, in terms of both direct and indirect selection. In terms of parameter range, DF is necessarily favored if $2N$ is favored (the reverse is not true). This is consistent with the view that angiosperms have both evolved DF and the shortest haploid phase.

DF and imprinting: Our model of DF evolution can also be closely compared to some models of imprinting evolution (Spencer and Williams 1997). Indeed, the evolution toward imprinting in animals can be viewed, at the level of a locus, as a transition from a diploid (1m:1p) to a functionally haploid state (0m:1p for maternal imprinting or 1m:0p for paternal imprinting). In the absence of double fertilization, only the maternally inherited haploid genome is expressed in the endosperm. This situation is therefore similar to a genome-wide paternal imprint. More specifically, Spencer and Williams (1997) consider a two-locus model in which the primary locus determines embryo survival and is exposed to deleterious mutations and the second, a modifier locus, is maternally expressed (*e.g.*, during female meiosis) and determines whether the maternal copy of the primary locus is imprinted or not in the embryo. This model would be comparable to a DF model with a paternal expression of the modifier locus. Indeed, in both cases the modifier is *cis*-acting, meaning that the allele causing silencing is on the same chromosome as the silenced allele (we define the *trans*-acting modifier as the reverse situation, where the allele causing silencing is not on the same chromosome as the silenced allele). In the DF model with paternal expression, the silenced allele is paternal (as in all DF models) and the control is also paternal. In the Spencer and Williams (1997) model, the silenced allele is maternal and the modifier is also controlled maternally. The comparison of our model with this model of imprinting highlights an important and possibly general feature of models for the evolution of gene expression. More than a maternal *vs.* paternal effect, the crucial issue is whether the modifier is *cis*- or *trans*-acting. It is much more difficult to evolve a *cis*- than a *trans*-acting silencing allele because indirect selection always benefits a modifier associated to purged chromosomes. In fact, for any given locus controlling the expression of genes exposed to deleterious mutations, there is a potential for conflict between its alleles. This conflict can occur if the modifier alleles do not have an additive effect (such as when there is a sex-limited expression of the modifier, but the phenomenon may be much more general).

Kin conflict *vs.* heterozygosity theory: A widespread view is that DF and MCD evolved in response to a kin conflict over resource allocation (Queller 1984; Dominguez 1995; Friedman 1995; Shaanker and Ganeshaiah 1997; Hardling and Nilsson 1999). This theory relies strongly on the idea that embryos compete for maternal resources. In fact, this theory is appealing because it claims to explain quite generally the evolution of traits involved in postfertilization resource allocation in both plants (DF, MCD, and imprinting) and animals (imprinting). It is, however, mostly supported from observations related to imprinting (see, *e.g.*, Dominguez 1995). The central argument is that the male/female conflict theory predicts a male–female coevolution, which can be detected by distinctive adaptations and counteradaptations. There is, for instance, the well-known case of imprinting of the fetal growth factor *Igf2* and its inhibitor *Igf2r*. Similarly, DF and MCD are viewed as steps in this male/female conflict in plants. There are two problems with this argument. First, what is interpreted here as the signature of a male/female conflict does not necessarily support the kin conflict theory since any theory involving an escalation would be consistent with this interpretation. For instance, our model shows that a male/female conflict can occur over the genetic constitution of the endosperm even in the absence of problems of resource allocation among kin. Second, under a conflict theory, we would expect to see, at least in some cases, an escalation of adaptations and counteradaptations. This is clearly not the case for the evolution of the ploidy of the endosperm: no plants exhibit, for instance, 2m:2p or 3m:2p endosperms, while nothing seems to limit *a priori* such escalation (contrary to the *Igf2* and *Igf2r* system where the escalation is maximal: the expression of these genes is totally suppressed when inherited maternally or paternally, respectively). Our models show that, depending on parameter values, DF and MCD may be stable states (advantageous for both males and females), which might in fact more closely reflect the data. Thus, the available evidence does not critically favor one theory over the other. More specific tests would be required to discriminate between them.

Combining theories: The masking effects generated by DF and MCD could interfere with resource allocation to the seed. Willson and Burley (1983) proposed that DF, by masking deleterious mutations in the endosperm, should increase resource allocation to the associated seed (and favor the transmission of paternally inherited alleles). They viewed MCD as a return to a state closer to 1m:0p, which, by exposing deleterious mutations in the endosperm, should decrease resource allocation to the associated seed (and favor the transmission of maternally inherited alleles). This view attempts to bridge the different theories by linking heterosis in the endosperm with the amount of resource allocation to the seed. It may provide a mechanism explaining how DF and MCD could play a role in a male/female conflict over resource allocation. The first part of this argument holds, but could be stated more precisely. Provided $h < \frac{1}{2}$, DF increases transitorily the mean fitness in a 1m:0p population. (It is, however, not true at equilibrium where the load is higher in a 1m:1p population if $h < \frac{1}{2}$.) Similarly, the second part of the argument can hold, but only under specific conditions. MCD transitorily decreases the mean fitness in a 1m:1p population only if $h < h_m$. This latter condition may, however, not be systematically verified. Overall, combining these two theories would require including them in a proper model, to investigate the conditions where they predict different and distinguishable outcomes. In any case, it seems necessary to include the impact of deleterious mutations because they necessarily occur and because they can produce similar patterns of genetic conflict.

Alternative nutritive tissues: Accumulating resources directly from an embryonic tissue (like cotyledons) or from a maternal or sporophytic tissue (like the seed coat) could be considered as an alternative to double fertilization. These tissues become predominant in the mature seed in some angiosperms, but the endosperm nearly always subsists, at least transitorily or as an interface. We may wonder why the endosperm is nearly always maintained in such cases. We can tentatively suggest two hypotheses. First, the endosperm can develop immediately after fertilization, while the development of a nutritive embryonic tissue would be constrained by the necessary steps of embryo development. Second, a sporophytic maternal tissue (1m1m′:0p) would confer a similar direct effect of masking but would lead to a very inefficient purge of deleterious mutations because it dissociates the genotype of the embryo from the phenotype on which selection occurs. Indeed, a deleterious m' allele would decrease the survival of embryos not carrying the mutation, while a deleterious p allele would not be purged (because it is not expressed in a 1m1m′:0p perisperm) although it is transmitted to the next generation. Thus, one can see that selection should be more efficient in purging deleterious mutations with a 1m:1p than with a 1m1m′:0p nutritive tissue.

In some angiosperms, endosperms derived from polysporic gametophytes are alternatives to 1m:1p or 2m:1p endosperms. In a tetrasporic or a bisporic gametophyte, the endosperm can be made of m, m', and p contributions (Friedman *et al.* 2008). As with the perisperm, polyspory dissociates the genotype of the embryo from the phenotype on which selection occurs, so that purging would be less efficient with polyspory. However, compared with a monosporic endosperm (2m:1p) where the two maternal alleles are mitotically derived, a polysporic endosperm (1m1m′:1p, for instance) is unlikely to carry two deleterious maternal alleles, and *aaA* endosperms would be comparatively less frequent with polyspory. For this reason, masking of deleterious alleles would be more efficient with polyspory. A further development of our theory should therefore include a model to evaluate how these antagonistic direct (masking) and indirect (purging) effects combine and to predict conditions when polyspory should evolve.

Conclusion: Double fertilization and maternal contribution doubling can both evolve because of the masking/exposition of deleterious mutations. This theory is consistent with the pattern seen in seed plants as well as with the theory of kin conflict. Interestingly, our model also predicts that conflicting selection pressures for DF and MCD may occur in male and female gametophytes. The effect of deleterious mutations cannot therefore be neglected in a comprehensive theory of DF and MCD evolution.

We thank E. Imbert for stimulating discussions and Dan Schoen, an anonymous reviewer, and Evolutionary Genetics and Ecology Group students, in particular, F. Débarre, for useful comments on the manuscript. This work was supported by a starting grant from the European Research Council to T.L. A.C. benefited from a fellowship from the French Ministry of Research.

LITERATURE CITED

Baroux, C., C. Spillane and U. Grossniklaus, 2002 Evolutionary origins of the endosperm in flowering plants. Genome Biol. **3:** 1026.1021–1026.1025.

Barton, N. H., and M. Turelli, 1991 Natural and sexual selection on many loci. Genetics **127:** 229–255.

Bengtsson, B. O., 1992 Deleterious mutations and the origin of the meiotic ploidy cycle. Genetics **131:** 741–744.

Brink, R. A., and D. C. Cooper, 1940 Double fertilization and development of the seed in angiosperms. Bot. Gaz. **102:** 1–25.

Brown, C. A., and D. R. Formanowicz, 1996 Reproductive investment in two species of scorpion, Vaejovis waueri (Vaejovidae) and Diplocentrus linda (Diplocentridae), from West Texas. Ann. Entomol. Soc. Am. **89:** 41–46.

Bulmer, M., 1986 Genetic models of endosperm evolution in higher plants, pp. 743–763 in *Evolutionary Process and Theory*, edited by S. Karlin and E. Nevo. Academic Press, New York.

Charnov, E. L., 1979 Simultaneous hermaphroditism and sexual selection. Proc. Natl. Acad. Sci. USA **76:** 2480–2484.

Chasnov, J. R., 2000 Mutation selection balance, dominance and the maintenance of sex. Genetics **156:** 1419–1425.

Chopra, R., and R. Sachar, 1963 Endosperm, pp. 135–170 in *New Advances in the Embryology of Endosperm*, edited by P. Maheshwari. International Society of Plant Morphologists, Delhi, India.

De Fraipont, M., J. Clobert and R. Barbault, 1996 The evolution of oviparity with egg guarding and viviparity in lizards and snakes: a phylogenetic analysis. Evolution **50:** 391–400.

De Jong, T. J., H. Van Dijk and P. G. L. Klinkhamer, 2005 Hamilton's rule, imprinting and parent-offspring conflict over seed mass in partially selfing plants. J. Evol. Biol. **18:** 676–682.
Dominguez, C. A., 1995 Genetic conflicts-of-interest in plants. Trends Ecol. Evol. **10:** 412–416.
Donoghue, M. J., and S. M. Scheiner, 1992 The evolution of endosperm: a phylogenetic account, pp. 356–389 in *In Ecology and Evolution of Plant Reproduction: New Approaches*, edited by R. Wyatt. Chapman & Hall, New York.
Friedman, W. E., 1995 Organismal duplication, inclusive fitness theory, and altruism—understanding the evolution of endosperm and the angiosperm reproductive syndrome. Proc. Natl. Acad. Sci. USA **92:** 3913–3917.
Friedman, W. E., 2001a Comparative embryology of basal angiosperms. Curr. Opin. Plant Biol. **4:** 14–20.
Friedman, W. E., 2001b Developmental and evolutionary hypotheses for the origin of double fertilization and endosperm. C. R. Acad. Sci. Ser. III Sci. Vie. Life Sci. **324:** 559–567.
Friedman, W. E., and K. C. Ryerson, 2009 Reconstructing the ancestral female gametophyte of angiosperms: insights from Amborella and other ancient lineages of flowering plants. Am. J. Bot. **96:** 129–143.
Friedman, W. E., W. N. Gallup and J. H. Williams, 2003 Female gametophyte development in Kadsura: implications for Schisandraceae, Austrobaileyales, and the early evolution of flowering plants. Int. J. Plant Sci. **164:** S293–S305.
Friedman, W. E., E. N. Madrid and J. H. Williams, 2008 Origin of the fittest and survival of the fittest: relating female gametophyte development to endosperm genetics. Int. J. Plant Sci. **169:** 79–92.
García-Dorado, A., and A. Caballero, 2000 On the average coefficient of dominance of deleterious spontaneous mutations. Genetics **155:** 1991–2001.
Guignard, L., 1899 Sur les anthérozoïdes et la double copulation sexuelle chez les végétaux angiospermes. C. R. Acad. Sci. Paris **128:** 864–871.
Haig, D., 2000 The kinship theory of genomic imprinting. Annu. Rev. Ecol. Syst. **31:** 9–32.
Haig, D., 2004 Genomic imprinting and kinship: How good is the evidence? Annu. Rev. Genet. **38:** 553–585.
Haig, D., and M. Westoby, 1988 Inclusive fitness, seed resource and maternal care, pp. 60–79 in *Plant Reproductive Ecology: Patterns and Strategies*, edited by J. L. Doust and L. L. Doust. Oxford University Press, Oxford.
Haig, D., and M. Westoby, 1989 Parent-specific gene-expression and the triploid endosperm. Am. Nat. **134:** 147–155.
Hardling, R., and P. Nilsson, 1999 Parent-offspring and sexual conflicts in the evolution of angiosperm seeds. Oikos **84:** 27–34.
Hardling, R., and P. Nilsson, 2001 A model of triploid endosperm evolution driven by parent-offspring conflict. Oikos **92:** 417–423.
Kacser, H., and J. A. Burns, 1981 The molecular-basis of dominance. Genetics **97:** 639–666.
Law, R., and C. Cannings, 1984 Genetic analysis of conflicts arising during development of seeds in the Angiospermophyta. Proc. R. Soc. Lond. Ser. B Biol. Sci. **221:** 53–70.
Maheshwari, P., 1963 *Recent Advances in the Embryology of Angiosperms*. Ed. International Society of Plant Morphologists, University of Delhi, Delhi, India.
Nawaschin, S. G., 1898 Resultate einer Revision der Befruchtungsvorgänge bei Lilium Martagon und Fritillaria tenella. Bull. Acad. Sci. St. Petersburg **9:** 377–382.
Otto, S. P., 2003 The advantages of segregation and the evolution of sex. Genetics **164:** 1099–1118.
Otto, S. P., and T. Day, 2007 *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press, Princeton, NJ.
Otto, S. P., and D. B. Goldstein, 1992 Recombination and the evolution of diploidy. Genetics **131:** 745–751.
Perrot, V., S. Richerd and M. Valero, 1991 Transition from haploidy to diploidy. Nature **351:** 315–317.
Queller, D. C., 1983 Kin selection and conflict in seed maturation. J. Theor. Biol. **100:** 153–172.
Queller, D. C., 1984 Models of kin selection on seed provisioning. Heredity **53:** 151–165.
Shaanker, R. U., and K. N. Ganeshaiah, 1997 Conflict between parent and offspring in plants: predictions, processes and evolutionary consequences. Curr. Sci. **72:** 932–939.
Shaanker, R. U., K. N. Ganeshaiah and K. S. Bawa, 1988 Parent-offspring conflict, sibling rivalry, and brood size patterns in plants. Annu. Rev. Ecol. Syst. **19:** 177–205.
Spencer, H. G., and J. M. Williams, 1997 The evolution of genomic imprinting: two modifier-locus models. Theor. Popul. Biol. **51:** 23–35.
Stebbins, G. L., 1976 Seeds, seedlings, and the origin of angiosperms, pp. 300–311 in *Origin and Early Evolution of Angiosperms*, edited by C. B. Beck. Columbia University Press, New York.
Szafraniec, K., D. M. Wloch, P. Sliwa, R. H. Borts and R. Korona, 2003 Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast Saccharomyces cerevisiae. Genet. Res. **82:** 19–31.
Trivers, R. L., 1974 Parent-offspring conflict. Am. Zool. **14:** 249–264.
Tutt, K., C. H. Daugherty and G. W. Gibbs, 2002 Differential life-history characteristics of male and female Peripatoides novaezealandiae (Onychophora: Peripatopsidae). J. Zool. **258:** 257–267.
Vassilieva, L. L., A. M. Hook and M. Lynch, 2000 The fitness effects of spontaneous mutations in Caenorhabditis elegans. Evolution **54:** 1234–1246.
Westoby, M., and B. Rice, 1982 Evolution of the seed plants and inclusive fitness of plant-tissues. Evolution **36:** 713–724.
Williams, J. H., and W. E. Friedman, 2002 Identification of diploid endosperm in an early angiosperm lineage. Nature **415:** 522–526.
Willson, M., and N. Burley, 1983 Sporogenesis and double fertilization, pp. 75–93 in *Mate Choice in Plants: Tactics, Mechanisms, and Consequences*, edited by R. May. Princeton University Press, Princeton, NJ.
Wourms, J. P., 1993 Maximization of evolutionary trends for placental viviparity in the spadenose shark, Scoliodon-laticaudus. Environ. Biol. Fishes **38:** 269–294.

Communicating editor: D. Charlesworth

APPENDIX A: RECURSIONS

The exact recursions of gamete frequencies (respectively, *AM*, *aM*, *AM′*, and *aM′*) are for the example of the DF model with a paternal expression of the modifier locus,

$$p'_{AM} = \left(\frac{1}{2T}\right)\begin{pmatrix} 2p_{AM}^2 + 2p_{AM}p_{Am} + p_{AM}p_{aM}(2-(1-(1-2h)\nu_M)s) \\ \quad + p_{AM}p_{am}(1-R)(2-(1-(1-2h)\nu_M + h(\nu_M-\nu_m))s) \\ \quad + p_{aM}p_{Am}R(2-(1-(1-2h)\nu_M + (h-1)(\nu_M-\nu_m))s) \end{pmatrix}(1-\mu) \tag{A1}$$

$$p'_{aM} = \left(\frac{1}{2T}\right)\begin{pmatrix} 2p_{AM}^2\mu + p_{AM}p_{aM}(2-s+\nu_M s - 2h\nu_M s)(1+\mu) + 2p_{AM}p_{Am}\mu \\ \quad + 2p_{aM}^2(1-s) + p_{AM}p_{am}(2(1(1-2h)\nu_M + h(\nu_M-\nu_m))s)(R(1-\mu)+\mu) \\ \quad + p_{aM}p_{Am}(2-(1+h\nu_M-(1-h)\nu_m)s)(1-R(1-\mu)) + 2p_{aM}p_{am}(1-s) \end{pmatrix} \tag{A2}$$

$$p'_{Am} = \left(\frac{1}{2T}\right)\begin{pmatrix} 2p_{AM}p_{Am} + 2p_{Am}^2 + p_{AM}p_{am}R(2-(1-(1-2h)\nu + h(\nu_M-\nu_m))s) \\ \quad + p_{Am}p_{aM}(1-R)(2-(1+2h\nu_M+\nu_m-2\nu_M + h(\nu_M-\nu_m))s) \\ \quad + p_{Am}p_{am}(2-(1-2\nu_M+\nu_m+2h\nu_m)s) \end{pmatrix}(1-\mu) \tag{A3}$$

$$p'_{am} = \left(\frac{1}{2T}\right)\begin{pmatrix} 2p_{am}^2(1-s) + 2p_{Am}^2\mu - 2p_{AM}p_{Am}\mu + 2p_{aM}p_{am}(1-s) \\ \quad + p_{aM}p_{Am}(2-(1+h\nu_M-(1-h)\nu_m)s)(R(1-\mu)+\mu) \\ \quad + p_{Am}p_{am}(2-s+(1-2h)\nu_m s)(1+\mu) \\ \quad + p_{AM}p_{am}(2-(1-(1-2h)\nu_M + h(\nu_M-\nu_m))s)(1-R(1-\mu)) \end{pmatrix}, \tag{A4}$$

where

$$T = \begin{pmatrix} (p_{aM}+p_{am})((p_{AM}+p_{Am})(1-2h)\nu_M - 1) \\ \quad + (p_{Am}(p_{aM}+p_{am}) - (p_{aM}p_{Am} + (p_{AM}+2p_{Am})p_{am})h)(\nu_M-\nu_m) \end{pmatrix} s. \tag{A5}$$

APPENDIX B: QLE ANALYSIS OF THE OTTO AND GOLDSTEIN (1992) MODEL

Whatever the expression of the modifier locus (uniparental or biparental expression), the selection coefficient of diploidy ($2N$), calculated with QLE assumptions, is

$$s_m = s\delta_\nu p_a(1-p_a)\left[\left(h-\frac{1}{2}\right) + s\frac{(h-1)}{2R}[(1-\nu)(R-1)-\nu h]\right] + o(\xi)^5. \tag{B1}$$

The eigenvalue is given in their original article. [All parameters are used equivalently except $d_{11} = \nu_M$, $d_{12} = (\nu_M + \nu_M)/2$, and $d_{22} = \nu_m$.]

Annexe **B**

Manuscrit : Patterns of resource allocation and male-female conflict : a case study in maize

Patterns of resource allocation and male-female conflict: a case study in maize

Cailleau A., Grimanelli D., Cheptou P.-O., Blanchet E., Lenormand T.

October 29, 2010

1 Introduction

Seed ability to survive and germinate (seed success) varies with the resource accumulated during seed maturation (Biere, 1991; Krannitz *et al.* , 1991; Andersson, 1990). Different factors may underly these variations. The embryo needs resources to initiate germination. There may be a threshold of resources under which the embryo cannot germinate. Above that threshold, seed sucess should increase as the accumulated resource increases. However, dispersion, escape from predators, or landfill may diminish as the seed becomes larger and heavier (Leishman *et al.* , 2000). As a consequence, seed sucess may also be limited if too much resource is accumulated.

The plant fitness faces a trade-off between seed success and seed number: an increase in a given component implies a decrease in the other one, because plant resource for reproduction is limited. In this context, the question of how a given plant adjusts seed number and seed mass as a function of resource availability has attracted much attention. Smith and Fretwell (1974) proposed a model investigating that question. In that model, plant fitness increases with a diminishing return per unit of resource invested in each seed. Plant optima in seed size is the point where the product of the initial number of developing seeds and the rate of surviving seeds is maximized. Seed number is fixed as a function of that optima and resource availability, following $N_Y = N/I_Y$, where I_Y is the amount of resource accumulated per offspring, N the amount of available resource, and N_Y the number of seeds.

Smith and Fretwell's model makes the hypothesis that the mother sporophyte controls resource allocation. In the litterature, maternal effects used to be widely considered to control most of seed traits, including seed weight (Denney, 1992). Before double fertilization (the parallel fertilization of the egg cell and its sister cell, the central cell; the first fertilization product gives rise to the embryo, and the second to the endosperm, Guignard, 1899; Nawaschin, 1898) was discovered, seed physiologists were in fact surprised by the occurrence of paternal effects (or Xénia) on what they considered maternal tissues (Denney, 1992). Even now that double fertilization is established, the idea that maternal effects predominate remains widespread (Sadras, 2007).

However, several arguments indicate that paternal genes in the embryo may at least partly control resource accumulation (De Jong & Scott, 2007). For example, *Arabidopsis thaliana* imprinted PcG genes seem to control the timing of

mitosis and cellularization in endosperm, a paternal excess of these genes causing a delayed timing of cellularization (Alleman & Doctor, 2000; Scott *et al.* , 1998; Leblanc *et al.* , 2002). A delayed timing of cellularization allows more mitosis, and ultimately more endosperm cells. The number of endosperm cells is a major determinant of seed ability to accumulate resources (sink capacity) (Brocklehurst, 1977; Jones *et al.* , 1996). If paternal genes partly control resource allocation, a more comprehensive model is necessary to predict seed size. The optimum in seed size for the embryo depends on its genetic relatedness to other embryos, that itself depends on the breeding system. When the plant is allogamous, the optimum seed size for the embryo is expected to be higher than that of the mother sporophyte (De Jong & Scott, 2007).

When the optima of the mother sporophyte and embryo differ, a Conflict Over Resource Allocation between the mother sporophyte and the embryo (parent-offspring conflit) is expected (hereafter refered to as CORA, Trivers, 1974). In addition, the genes inherited maternally and paternally by the embryo may also differ in their optima. The optima of maternally inherited genes should be intermediate between that of the mother sporophyte and the embryo, while the optima of paternally inherited genes should be above that of the embryo (De Jong & Scott, 2007). In that case, a male-female conflict over resource allocation is also expected.

CORA is hypothesized to drive the evolution of numerous traits associated to the determinism of seed resource allocation (Uma Shaanker & Ganeshaiah, 1997). For example, CORA is hypothesized to drive the evolution of integuments derived specifically from the mother sporophyte in the ancestor of seed plants (these integuments being hypothesized to limit seed growth, Westoby & Rice, 1982). CORA is hypothesized to drive the evolution of the genetic make-up of the endosperm, the nutritive tissue of the embryo in angiosperms (Charnov, 1979; Westoby & Rice, 1982; Willson & Burley, 1983; Queller, 1983, 1984; Law & Cannings, 1984; Bulmer, 1986; Haig & Westoby, 1989; Friedman, 1995; Hardling & Nilsson, 2001, 1999). More precisely, it is hypothesized to have driven the evolution of double fertilization (Guignard, 1899; Nawaschin, 1898), a doubling of maternal genomic contributions in the central cell (Maheshwari, 1948), polyspory (the participation of several meiotic product in the gametophyte, Maheshwari, 1948) and genomic imprinting (differential methylation of paternal and maternal genes, Kermicle & Alleman, 1990). In angiosperms, CORA is also hypothesized to have driven the evolution of septa (which physically separate seeds in fruits, Uma Shaanker *et al.* , 1988), hypostase (Westoby & Rice, 1982), haustoria (embryonic structures that protrude into maternal tissues, Queller, 1983; Bhatnagar & Kallarackal, 1980), strike (tendency of polar nuclei to become non-functional in polysporic gametophytes, Haig, 1986), overproduction of ovules and polyembryony (hypothesized to be adaptations allowing selective abortion, Sakai, 2007; Uma Shaanker & Ganeshaiah, 1996), or hormonal patterns of expression (Queller, 1983; Ravishankar *et al.* , 1995). All these traits are hypothesized to be adaptations and counter-adaptations leading to increases and decreases in embryo's control over resource allocation in the context of the arm race generated by genetic conflicts (Uma Shaanker & Ganeshaiah, 1997).

CORA makes two important assumptions. A first assumption is that offspring has the opportunity to modify resource allocation (De Jong & Scott, 2007; House *et al.* , 2010). Several studies attempted to show that paternal effects occur (Obeso, 2004a,b; Banuelos & Obeso, 2003; House *et al.* , 2010;

Andersson, 1990). Some detected paternal effects but their methods have been questioned (De Jong & Scott, 2007). Additional studies of this kind are required to know whether that first assumption is filled. However, male genotypes could modify the amount of resource they receive whithout competition. Differences between males may be due to different abilities to efficiently transform the resource. A second assumption of CORA is that offspring compete.To show that competition occur, it is necessary to show that when a given genotype acquire a given resource, neighboring genotypes receive less resource, i.e. that realocation occurs. No study adressed this issue specifically, i.e. question whether paternal effects correspond to differential competitive abilities and result in resource realocation.

To test whether realocation occurs, it is necessary that seeds sired by different males interact. Hence, one has to sire different seeds with pollen from different (but variable) genotypes on the same plant, controling for paternity. Ideally, the plant should be allogamous, because the conflict is expected to be stronger in such plants, and subject to natural selection for seed size. House *et al.* (2010) mix males on the same plant and keep information about paternity, but all fathers are always present in all plants, so the dataset could not be used to study interactions between specific couples of males.

Our study is the first to study whether and how realocation occurs. We choose maize that allows to keep information on paternity (with a phenotypic aleurone marker) and is allogamous. As a cultivated plant, it has been subject to artificial selection, mainly for exhibiting a limited number of inflorescence (Sadras, 2007). However, this is quite independant on whether seed mass is maternally or paternally controlled. We mixed pollens of maize lines following a diallel design, and fit seed mass with a model made of competition-dependent and competition-independent terms. If competition-dependent terms are significant, realocation should occur.

2 Methods

2.1 Material

Maize is a monoïc allogamous annual species, without self-incompatibility system. Maize exhibit an albuminate seed that presents an extensive endosperm at maturity. Male inflorescences are formed at the apex of the plant. One or several ears (i.e. female inflorescences) are formed at the base of leaves. Ears have a regular architecture. They are cylinders were flowers are aligned following lines (that make the rounds of the cylinder) and columns (that go from the base to the apex of the cylinder). Each flower contains one ovule so that this alignment is roughy kept by seeds. In maize, seed size is often negatively correlated to seed set in a given ear (Sadras, 2007), which may indicates that resource allocation is limited and that embryos are potentially in competition for it.

Seeds from six lines were obtained from the Maize Genetic Cooperation Stock Center (http://maizecoop.cropsci.uiuc.edu): A619, W23, W22, B73, Mo17, and an unidentified line we refer to as B1-Peru. They were grown and multiplied by selfing in a field in Garlin, France, in 2008. All lines were again grown in a field in Garlin, France, in 2009, in order to control pollination following our cross

design. Flowering dates were recorded in 2008 in order to synchronize them in 2009.

The R-locus determines aleurone color. Each line was represented both by an isoline with a recessive marker at R locus (white or yellow) and by an isoline with a dominant marker at that locus (black or molted), except to B1-Peru, available only with the dominant marker (Table 1 indicate each line with the stock and marker names). Wherever the pollen was mixed, the female carried the recessive marker, one male carried a dominant marker, and the other one a recessive marker. These crosses allowed us to recognize siring fathers for each seed on a ear.

Line	R-locus allele	Phenotype	Stock reference
A619	R1-st	Molted	X18C
A619	c1r1	Yellow	218DB
Mo17	R1-scm2	Black	X26B
Mo17	c1r1	Yellow	218DB
W23	R1-r	Black	M142A
W23	r1-g	Light yellow	M142T
W22	R1-sc	Black	X19A
W22	r1-g	Yellow	X14L
B73	R1-nj	Navajo	X17D
B73	c1r1	Yellow	218DB
NA	B1-Peru	Black	218GA

Table 1: Maize lines

2.2 Cross design

The experimental design is given Table 2. It is a diallel, i.e. each line was used both as male and female. Fertilizations were made following classical methods (pollen was collected in paper bags and placed in the chosen ear). In most cases, two males were mixed to sire a given ear. To mix pollens, one pollen bag was spilt in the other one. We also fertilized some ears without mixing males. Only four lines were used as females for mixed crosses (A619, Mo17, W22, W23). Only the uppermost ear was used, other ones were covered by paper bags to avoid uncontrolled effects.

2.3 Measures

Ovule set was estimated [as number of columns x (number of non-fertilized ovule + number of seeds on a column)] for each ear (except ears were the apex was missing). Missing seeds were counted. Then, all ears were shelled, and seeds were sorted, counted and weighed by genotype.

For each ear, we then estimated the number of seeds from each genotype initially present on the ear. The estimation was the proportion of that genotype (estimated from non-missing seeds) multiplied by the total number of seeds initially present on the ear (i.e. number of non-missing seeds + number of missing seeds). We also estimated the weight of seeds from each genotype. The weight of non-missing seeds was divided by their number, and multiplied by the estimation of the total number of seeds from that genotype.

		Female						
Male 1	Male 2	A619	B1-Peru	B73	Mo17	W22	W23	Sum
A619	A619	16	11	5	5	13	11	61
A619	B1-Peru	13	0	0	0	15	8	36
A619	B73	5	0	0	1	11	8	25
A619	Mo17	8	0	0	8	17	13	46
A619	W22	19	0	0	6	19	15	59
A619	W23	12	0	0	0	21	20	53
B1-Peru	B1-Peru	10	7	6	5	7	8	43
B1-Peru	B73	13	0	0	5	11	7	36
B1-Peru	Mo17	7	0	0	4	12	9	32
B1-Peru	W22	11	0	0	0	15	10	36
B1-Peru	W23	10	0	0	1	13	15	39
B73	B73	7	10	10	7	8	8	50
B73	Mo17	8	0	0	5	11	7	31
B73	W22	9	0	0	2	13	8	32
B73	W23	11	0	0	2	13	10	36
Mo17	Mo17	9	10	4	8	8	9	48
Mo17	W22	10	0	0	8	12	13	43
Mo17	W23	10	0	0	4	20	16	50
W22	W22	13	14	6	4	13	14	64
W22	W23	8	0	0	3	14	15	40
W23	W23	5	16	7	6	14	14	62
Number of crosses		214	68	38	84	280	238	922
Number of measures		368	68	38	132	497	412	1515

Table 2: Sample size (number of ear) for each kind of cross. In the case of mixed crosses, there is two siring males, and each correspond to one measure of seed mass. Hence we have 922 ears and 1515 measures of seed mass.

2.4 Statistical analysis

2.4.1 Model specification

We analyse the mass of ressources acquired by a set of seeds from a given genotype, on a given pod, as a function of four terms. Two of them do not depend on competition, while the two other one depend on competition. These terms are designed to represent four biologically different processes of resource allocation.

Let M_{ijkl} be global mass of a set of seeds from a given genotype, on a given pod (l). The model can be written :

$$M_{ijkl} = I_i f_j + I_j n_j + (D_i + D_j n_j + D_k n_k) f^*_{jki} \tag{1}$$

We detail each of these terms hereafter.

Competition-independent terms We consider two modes of resource allocation that do not involve competition. First, the plant may allocate a fixed amount of resource per ear (i.e. independantly of the number of seed per ear). That resource would be shared equally between all seeds on the ear. Because it

is allocated by the plant, we suppose that it may depend on the female genotype (i). We note I_i (I for Independant of competition) that amount of resource. A given seed acquires a fraction $1/(n_j+n_k)$ (n_j and n_k being the numbers of seeds sired by males j and k, respectively) of I_i. Hence, the set of seeds sired by a focal male j acquires a fraction $f_j = n_j/(n_j + n_k)$ (equal to the frequency of male j) of I_i, i.e. an amount of resource equals to $I_i f_j$.

Another amount of resource may be attracted seeds. This phenomenon is often reffered to as 'sink effect'. Because its strenght should depend on the genotype of seeds, that amount of resource may vary with male genotype (j). We note I_j the amount of resource attracted in a seed sired by male j. The set of seeds sired by male j acquire an amount $I_j n_j$ of resource via this sink-effect.

Competition-dependent terms We also consider two modes of resource allocation that involve competition. The plant of genotype i may allocate a given amount of resource to a pool, which would be divided between genotypes following a function that depends on the competive abilities of males. Let D_i (D for Dependant of competition) be that amount of resource. We note f^*_{jki} the fraction of D_i received by seeds sired by male j when competing with seeds sired by male k on female i :

$$f^*_{jki} = \frac{f_j \alpha_{jki}}{f_j \alpha_{jki} + f_k \alpha_{kji}}; \tag{2}$$

where $f_k = n_k/(n_k + n_j)$, α_{jki} is the competitive ability of male j to acquire the resource face to male k on female i, and α_{kji} is the competitive ability of male k to acquire the resource face to male j on female i. Alternatively, we can assume that α_{jki} and α_{kji} are not very different, in which case we can assume that $\alpha_{jki}/\alpha_{kji}$ is close to one, and linearize f^*_{jki} :

$$f^*_{jki} = f_j + a_{jki} f_j f_k + o\left(\frac{\alpha_{jki}}{\alpha_{kji}}\right)^1 \tag{3}$$

By construction, $a_{jki} = -a_{kji}$, i.e. what is taken on one side is lost on the other side, reflecting directly the competition of resources. The set of seeds sired by male j acquire an amount $D_i f^*_{jki}$ of resource from the pool allocated by the female plant.

Another amount of resource may arrive in the pool subject to competition, due to sink-effect. In that case, each seed on the ear (independently of its genotype) contribute to the sink effect and the total amount of resources attracted in the pool is then divided between seeds depending on their competitive abilities. Let D_j be the amount of resource attracted by one seed sired by male j, and D_k be the amount of resource attracted by one seed sired by male k. The total amount of resource attracted by the males in the pool subject to competition is $(D_j n_j + D_k n_k)$. We assume that the competitive abilities of males are the same than for D_i, i.e. is f^*_{jki}. The set of seeds sired by male j acquire an amount $(D_j n_j + D_k n_k) f^*_{jki}$ of resource via this effect.

Full model We can make the model more complex, by supposing that male-female interactions occur, either in the competition dependent or the competition independant term.

Let I_{ij} be the amount of resource attracted directly in the seed by sink strength, that depends on the interaction between female i and siring father j. Let D_{ij} and D_{ik} be the amounts of resources attracted by sink strength in the pool of resource subject to competition and that depends on male-female interactions. Let D_{jk} be the amount of resources attracted by sink strength in the pool of resource subject to competition and that depends on the interaction between males. Finally, let D_{ijk} be the amount of resources attracted by sink strength in the pool of resource subject to competition and that depends on the triple-interaction between the two males and the female (a summary of these terms is given Table 3). The model that include all terms is :

$$\begin{aligned} M_{ijkl} = {} & I_i f_j + I_j n_j + I_{ij} n_j f_j \\ & + (D_i + D_j n_j + D_k n_k + D_{ij} n_j + D_{ik} n_k + D_{jk} n_j n_k + D_{ijk} n_j n_k) f^*_{jki} \end{aligned} \tag{4}$$

From a pattern of resource allocation, it is impossible to distinguish full maternal control (with equal resource allocation to each seed) from zero maternal control but equally competitive seeds competing for resources. In our model, this ambiguity also occurs : from Equation 3, we see that f^*_{jki} can be expressed as a sum of two terms, $f_j + a_{jki} f_j f_k$, where the former can be interpreted as being competition independant (each genotype receive resources in proportion to its frequency). However, in our model, competition ability is supposed to be related to a global pool of resource (made of male and female contributions). This constraint allows us to discriminate competition-independant and competition-dependant terms.

Estimate	Competition	Vary with	Nb of parameters
I_i	No	Female i	6
I_j	No	Male j	6
I_{ij}	No	Males j and female i	36
D_i	Yes	Female i	6
D_j, D_k	Yes	Male j and male k, respectively	6
D_{ij}, D_{ik}	Yes	Males j or k and female i	36
$D_{jk} = D_{kj}$	Yes	Males j and k	15
D_{ijk}	Yes	Males j, male k and female i	60
a_{jki}	Yes	Males j, male k and female i	60

Table 3: Description of model's terms. Estimates involving a triple interaction (ν) can be estimated for only four females, because there is no replicate of pollen mix for two females. The total number of parameters for the full model is 600.

2.4.2 Model selection

Our model is non-linear and was fitted using least mean square using Mathematica (Wolfram, 1988). We simplified it based on AIC comparisons. The model have to be non-linear to constrain f^*_{jki} to be the same for the all the resources allocated by a competition-dependant mode: on a biological ground, there is no reason to imagine that the resource contributed in the pool subject to competition by the plant and by sink-effect be different.

Additionaly, the male effect on the pool of resource subject to competition (D_j, D_{ij}) it the same for a male of a given genotype, be this male the focal male (j) or the competing male (k).

Simplifications concerning the competitive abilities In the full model, we allow male competitive ability to differ with the competing male. One major simplification we test is transitivity of the competitive abilities. If competitive abilities are transitive, one parameter per male genotype is suficient to predict its ability to compete with any other male. We write a_{ji} the competitive ability of male j on female i. Because competitive abilities are relative, we fix the competitive ability of one male (A619) to zero. To implement this simplication, we replace each a_{jki} by $a_{ji} - a_{ki}$.

In the full model, we estimate different a_{jki} depending on the female (i). Another simplification is that the value of a_{jk} does not depend on the female.

Thus, we also build models where competitive abilities are transitive and independent of the female, i.e. where $a_{ji} = a_j$ and $a_{ki} = a_k$. To implement this simplication, we replace each a_{jk} by $a_j - a_k$.

Simplifications for other parameters We performed two other kinds of simplification. The first consists to remove terms (no effect). The second consists to constrain some estimates to be equal between genotypes, which is equivalent to consider that the effect is not variable.

3 Results

3.1 Model selection

The AIC of the full model is 9917.44. Many combinations of simplifications are tested. A summary of tested models and their AIC is given Table 4. In a first series of model ($M1$ to $M4$), we evaluate whether interactions have to be kept or dropped. I_{ij} and D_{ijk} are not significant, while D_{ij} and D_{jk} are significant. The model where a does not depend on the female is better than the model where it does, which suggest that there is no male-female interaction on males ability to acquire the resource that is subject to competition. The best models were those assuming transitivity. We summarize this information in the model $S1.4$ (Table 4), which provides the best AIC obtained so far, 9705.26.

In a second series of models, we test whether I_i and I_j are significant, and if not, if they can be reduced to simpler terms by constraining some parameters. The best model is $M5.3$ (AIC 9698.46), where I_i and I_j are not significant, even when reduced to simpler terms by constraining estimates. Interistingly, this indicates that there is no competition-independant resource allocation.

3.2 Estimates of the best model

To sum up, the best model (AIC=9698.46, R^2=0.9847) assumes :

1. no competition-independent term ($-I_i - I_j - I_{ij}$);
2. no triple interaction between the female and the two males in the competition-dependent term ($-D_{ijk}$);

Effect	D_j. D_k	D_i	D_{jk}	D_{ji}. D_{ki}	D_{ijk}	a	I_j	I_i	I_{ij}	Adjusted R^2	AIC	R^2	Nb parameters	df
Full mod.	•	•	•	•	•	a_{jki}	•	•	•	0.9837	9917.42	0.9869	291	1224
$M1.1$	•	•	•	•	x	a_{jki}	•	•	•	0.9835	9867.33	0.9857	201	1314
$M1.2$	•	•	•	x	x	a_{jki}	•	•	•	0.9827	9904.7	0.9846	165	1350
$M1.3$	•	•	x	•	x	a_{jki}	•	•	•	0.9832	9884.96	0.9852	186	1329
$M1.4$	•	•	x	x	x	a_{jki}	•	•	•	0.9823	9928.53	0.9841	150	1365
$M2.1$	•	•	•	•	•	a_{jki}	•	•	x	0.9839	9871.13	0.9866	255	1260
$M2.2$	•	•	•	•	x	a_{jki}	•	•	x	0.9834	9841.3	0.9852	165	1350
$M3.1$	•	•	•	•	•	a_{jk}	•	•	•	0.9845	9787.7	0.9867	216	1299
$M3.2$	•	•	•	•	•	$a_j - a_k$	•	•	•	0.9845	9772.79	0.9866	206	1309
$M3.3$	•	•	•	•	•	$a_{ji} - a_{ki}$	•	•	•	0.9842	9829.86	0.9867	237	1278
$M4.1$	•	•	x	x	x	a_{jk}	•	•	x	0.9784	10129.2	0.9789	39	1476
$M4.2$	•	•	x	x	x	$a_j - a_k$	•	•	•	0.9785	10110.4	0.9789	29	1486
$S1.4$	•	•	•	•	x	$a_j - a_k$	•	•	x	0.9841	9705.77	0.9849	80	1435
$M5.1$	•	•	•	•	x	$a_j - a_k$	x	•	x	0.9840	9709.99	0.9847	74	1441
$M5.2$	•	•	•	•	x	$a_j - a_k$	•	x	x	0.9840	9709.73	0.9847	74	1441
$M5.3$	•	•	•	•	x	$a_j - a_k$	x	x	x	0.9840	**9698.46**	0.9847	68	1447
$M5.4$	•	•	•	•	x	$a_j - a_k$	I_J	•	x	0.9840	9711.72	0.9847	75	1440
$M5.5$	•	•	•	•	x	$a_j - a_k$	•	I_I	x	0.9840	9707.52	0.9848	75	1440
$M5.6$	•	•	•	•	x	$a_j - a_k$	x	I_I	x	0.9840	9699.95	0.9847	69	1446

Table 4: AIC for all models. Column names refers to parameters given Table 3. • indicate the term is kept as in the full model, while x indicates it is withdrawn. As in the text, indices are as follows : j indicates the focal male, k the other male, i females. I_J and I_I indicate that estimates are not allowed to vary between genotypes.

3. transitivity and no female effect on the ability of males to compete for the competition-dependant resource ($a_{ijk} \rightarrow a_j - a_k$).

This corresponds to :

$$M_{ijkl} = (D_i + D_j n_j + D_k n_k + D_{ij} n_j + D_{ik} n_k + D_{jk} n_j n_k) f^*_{jki} \qquad (5)$$

where $f^*_{jki} = f_j + (a_j - a_k) f_j f_k$. The estimates of this model are given Tables 6, 7, and 8, in the appendix.

3.2.1 Competitive abilities

All the resources are allocated through a mode that involve competition, and competitive abilities of males are transitive. This corresponds to a very simple mode of resource allocation. Competitive abilities of males are given Table 5. W23 is the least competitive male, while B73 is the more competitive one. If the frequency of these two genotypes is 0.5, this correspond to B73 acquiring 53% of the resource, while W23 acquires 47% of the resource. Hence, in that case, W23 acquires 88% of the resource acquired by B73.

A619	a_1	0.000
B1-Peru	a_2	-0.030
Mo17	a_3	0.058
B73	a_4	0.089
W22	a_5	0.038
W23	a_6	-0.021

Table 5: Competitive abilities of male genotypes.

3.2.2 Origin of resources available in the pool

The amount of resource that is available in the pool and subject to competition is made i) of a basic amount of resource contributed by the plant whatever the number of seeds presents on the ear, and ii) a supplementary amount of resource, due to a sink effect, that increases linearly with seed set.

To illustrate the variations of these two sources, we consider a standard ear with 200 seeds (100 of male j, 100 of male k). We use parameter estimates to compute the amount of resource contributed by the males and the female in that case, for each combination of genotypes {i,j,k}. We represent in the Figure 1 the total amount of resource available in the pool as a function of seed set. Intercepts represent the fixed amount of resource contributed by the female, and slopes represent the increase in the resource contributed to the pool due to sink-effect. We can see that females are not only variable for the initial amount of resource contributed to the pool (variability of intercepts), but also for the sink effect (variability of slopes).

According to this computation, the amount of resources contributed by the plant varies from zero to 50% depending on the female genotype (Figure 2). We can also see Figure 2 that this (relative) amount of resource does not vary a lot with the males presents on the ear. Of note, according to these estimates,

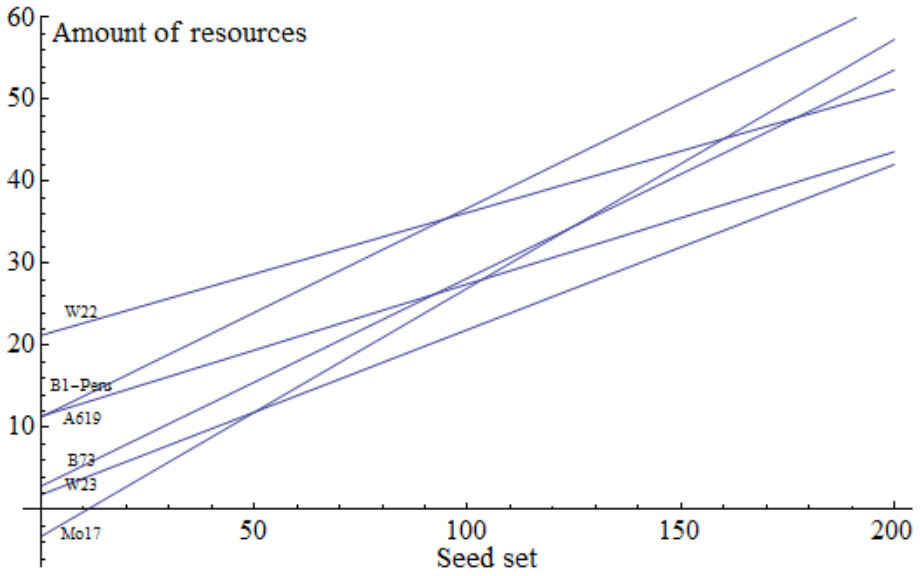

Figure 1: Total amount of resource in the pool as a function of seed set. Slopes were computed by averaging the sink-effects due to all combinations of males j and k.

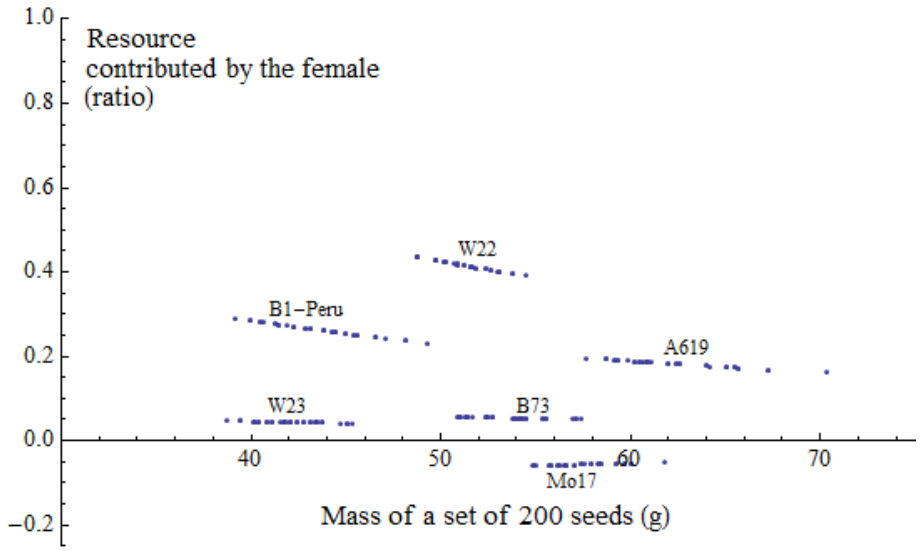

Figure 2: Ratio of the resource carried by the mother plant as a function of seed mass. The points are simulated considering that one male (j) sires 100 seeds and another male (k) sires 100 other seeds on a ear of female i. Each point represents one kind of cross (i.e. one combination {i,j,k}). Each series of point correspond to a different female (whose genotype is indicated).

one female (Mo17) would contribute a small but 'negative' amount of resources. In that case, the estimate is certainly artefactual (there is no constraint in the model to have positive terms) and not significantly different from zero.

Now that we know the amounts of resources contributed by the plant and by sink strength, we can ask if males differ in their sink effect. We represent Figure 3 the differences in the amounts of resource contributed by two males from the same ear. We can see that this difference ranges from zero to 6g. Hence, altough the ratio of the resource contributed by sink effect *versus* plant varies a lot, the sink strenghts of different males on a given ear is not very different.

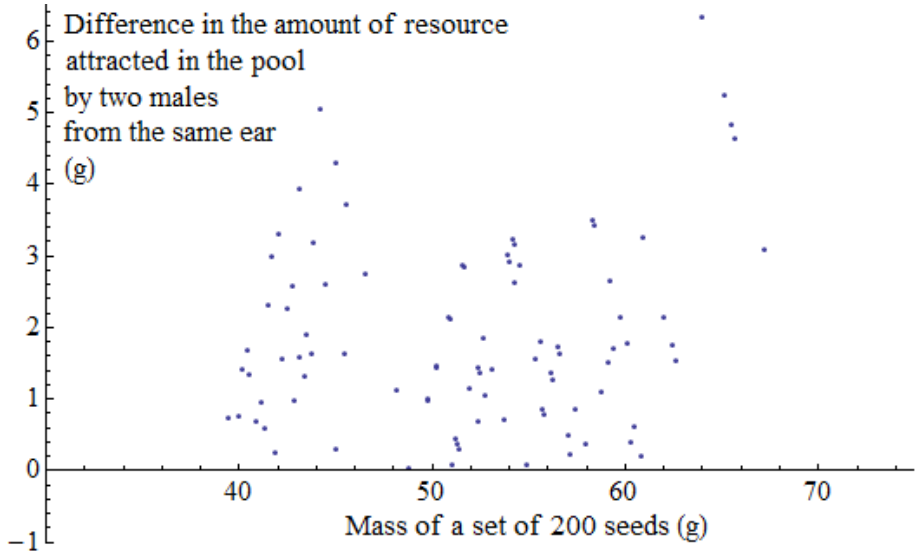

Figure 3: Difference in the amounts of resource carried by two males from the same ear. The points are simulated considering that one male (j) sires 100 seeds and another male (k) sires 100 other seeds on a ear of female i. Each point represents one kind of cross (i.e. one combination {i,j,k}).

4 Discussion

We study the control and patterns of resource allocation in maize. We show that resources are allocated to the ear by two processes : a fixed amount does not depend on seed number (and represent zero to half of the resource depending on the female), while another amount is contributed by a sink effect. Second, we show that all the resources are subject to competition. Third, we show that males differ in their competitive abilities to acquire those resources. Finaly, we show that competitive abilities are transitive.

The pattern we describe with that experiment is consistent with the statement of Borras *et al.* (2004) that maize is sink-limited in most growing conditions, since more than half the resource acquired by seeds are attracted in the ear by sink effect. This pattern of a strong sink effect and the fact that the female allocate the resource in a way that allows competition between males are

also consistent with the observation that maternal vessels never enter the developing seed, so that the extent to which assimilates enter the seed completely depends on the passive barriers that are maternal tissues (Ravishankar *et al.* , 1995).

The theory of parent-offspring conflit over resource allocation assumes that seeds compete for resources. This assumption is validated by our experiment with maize. However, it is now necessary to replicate such an experiment with other species to show that competition between seeds is the general case.

Moreover, two view of the parent offspring conflict over resource allocation coexist in the litterature. In the introduction, we described the view according to which female optima is to allocate resources equally, while male optima is to bias resource allocation. The other view is that maternal optima is not to allocate resources equally between seeds, but rather as a function of their genetic qualities (Westoby & Rice, 1982; Willson & Burley, 1983; Haig, 1990). According to that view, male optima is that the seed is not aborted. In our experiment, resources are not allocated equally among seeds. Two interpretations are possibles. First, maternal optima could be to allocate resources equally, but the expression of paternal genes in the albumen may give some control to males, which allow them to bias resource allocation. Alernatively, maternal optima may be to allocate resources as a function of seeds genetic qualities. It was suggested that the mechanism of this selective allocation could be to leave embryos compete together, the fittest being expected to acquire more resource. At first sight, this may be consistent with our results, since competition occurs in our experiment. However, whether competitive ability is indicative of seed genetic quality may be questioned. Genetic quality should depend on the number of deleterious mutations in the haploid genotypes of gametes, but also on the combining ability of these genotypes (i.e. heterosis). Now, our experiment showed that for maize lines, competitive abilities do not depend on the female. Hence, the competitive abilities we measure are at least not indicative of the variations in genetic qualities that results from the combining ability of male and female genotypes. Overall, the view according to which female optima is to allocate resources equally, while male bias resource allocation is more consistent with our results.

References

Alleman, M., & Doctor, J. 2000. Genomic imprinting in plants: observations and evolutionary implications. *Plant Molecular Biology*, **43**(2), 147–161. 2

Andersson, S. 1990. Paternal Effects on Seed Size in a Population of Crepis-Tectorum (Asteraceae). *Oikos*, **59**(1), 3–8. 1, 3

Banuelos, M. J., & Obeso, J. R. 2003. Maternal provisioning, sibling rivalry and seed mass variability in the dioecious shrub Rhamnus alpinus. *Evolutionary Ecology*, **17**(1), 19–31. 2

Bhatnagar, SP, & Kallarackal, J. 1980. Cytochemical studies on the endosperm of Linaria bipartita (Vent.) Willd. with a note on the role of endosperm haustoria. *Cytologia*, **45**, 247–256. 2

Biere, A. 1991. Parental Effects in Lychnis-Flos-Cuculi .1. Seed Size, Germination and Seedling Performance in a Controlled Environment. *Journal of Evolutionary Biology*, **4**(3), 447–465. 1

Borras, L., Slafer, G. A., & Otegui, M. E. 2004. Seed dry weight response to source-sink manipulations in wheat, maize and soybean: a quantitative reappraisal. *Field Crops Research*, **86**(2-3), 131–146. 12

Brocklehurst, P. A. 1977. Factors Controlling Grain Weight in Wheat. *Nature*, **266**(5600), 348–349. 2

Bulmer, MG. 1986. Genetic models of endosperm evolution in higher plants. *Pages 743–763 of:* Karlin, S., & Nevo, E (eds), *Evolutionary process and theory*. New York: Academic press. 2

Charnov, E. L. 1979. Simultaneous Hermaphroditism and Sexual Selection. *Proceedings of the National Academy of Sciences of the United States of America*, **76**(5), 2480–2484. 2

De Jong, T. J., & Scott, R. J. 2007. Parental conflict does not necessarily lead to the evolution of imprinting. *Trends in Plant Science*, **12**(10), 439–443. 1, 2, 3

Denney, J. O. 1992. Xenia Includes Metaxenia. *Hortscience*, **27**(7), 722–728. 1

Friedman, W. E. 1995. Organismal Duplication, Inclusive Fitness Theory, and Altruism - Understanding the Evolution of Endosperm and the Angiosperm Reproductive Syndrome. *Proceedings of the National Academy of Sciences of the United States of America*, **92**(9), 3913–3917. 2

Guignard, L. 1899. Sur les anthérozoides et la double copulation sexuelle chez les végétaux angiospermes. *C. R. Acad. Sci. Paris*, **128**, 864–871. 1, 2

Haig, D. 1986. Conflicts among Megaspores. *Journal of Theoretical Biology*, **123**(4), 471–480. 2

Haig, D. 1990. Brood Reduction and Optimal Parental Investment When Offspring Differ in Quality. *American Naturalist*, **136**(4), 550–556. 13

Haig, D., & Westoby, M. 1989. Parent-Specific Gene-Expression and the Triploid Endosperm. *American Naturalist*, **134**(1), 147–155. 2

Hardling, R., & Nilsson, P. 1999. Parent-offspring and sexual conflicts in the evolution of angiosperm seeds. *Oikos*, **84**(1), 27–34. 2

Hardling, R., & Nilsson, P. 2001. A model of triploid endosperm evolution driven by parent-offspring conflict. *Oikos*, **92**(3), 417–423. 2

House, C., Roth, C., Hunt, J., & Kover, P.X. 2010. Paternal effects in Arabidopsis indicate that offspring can influence their own size. *Proceedings of the Royal Society B*. 2, 3

Jones, R. J., Schreiber, B. M. N., & Roessler, J. A. 1996. Kernel sink capacity in maize: Genotypic and maternal regulation. *Crop Science*, **36**(2), 301–306. 2

Kermicle, J. L., & Alleman, M. 1990. Gametic Imprinting in Maize in Relation to the Angiosperm Life-Cycle. *Development*, 9–14. 2

Krannitz, P. G., Aarssen, L. W., & Dow, J. M. 1991. The Effect of Genetically Based Differences in Seed Size on Seedling Survival in Arabidopsis-Thaliana (Brassicaceae). *American Journal of Botany*, **78**(3), 446–450. 1

Law, R., & Cannings, C. 1984. Genetic Analysis of Conflicts Arising during development of Seeds in the Angiospermophyta. *Proceedings of the Royal Society of London. Series B, Biological Sciences*, **221**(1222), 53–70. 2

Leblanc, O., Pointe, C., & Hernandez, M. 2002. Cell cycle progression during endosperm development in Zea mays depends on parental dosage effects. *Plant Journal*, **32**(6), 1057–1066. 2

Leishman, M.R., Wright, I.J., Moles, A.T., & Westoby, M. 2000. The evolutionary ecology of seed size. *Pages 31–57 of:* Fenner, M. (ed), *Seeds. The Ecology of Regeneration in Plant Communities.* Wallingford, UK: CABI Publishing. 1

Maheshwari, P. 1948. The angiosperm embryo sac. *The Botanical review*, **14**(1), 1–56. 2

Nawaschin, S.G. 1898. Resultate einer Revision der Befruchtungsvorg´ange bei Lilium Martagon und Fritillaria tenella. *Bull. Acad. Sci. St Petersburg*, **9**, 377–382. 1, 2

Obeso, J. R. 2004a. A hierarchical perspective in allocation to reproduction from whole plant to fruit and seed level. *Perspectives in Plant Ecology Evolution and Systematics*, **6**(4), 217–225. 2

Obeso, J. R. 2004b. Seed provisioning within holly fruits: test of the hierarchical model. *Evolutionary Ecology*, **18**(2), 133–144. 2

Queller, D. C. 1983. Kin Selection and Conflict in Seed Maturation. *Journal of Theoretical Biology*, **100**(1), 153–172. 2

Queller, D. C. 1984. Models of Kin Selection on Seed Provisioning. *Heredity*, **53**(AUG), 151–165. 2

Ravishankar, K. V., Uma Shaanker, R., & Ganeshaiah, K. N. 1995. War of Hormones over Resource-Allocation to Seeds - Strategies and Counterstrategies of Offspring and Maternal Parent. *Journal of Biosciences*, **20**(1), 89–103. 2, 13

Sadras, V. O. 2007. Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crops Research*, **100**(2-3), 125–138. 1, 3

Sakai, S. 2007. A new hypothesis for the evolution of overproduction of ovules: An advantage of selective abortion for females not associated with variation in genetic quality of the resulting seeds. *Evolution*, **61**(4), 984–993. 2

Scott, R. J., Spielman, M., Bailey, J., & Dickinson, H. G. 1998. Parent-of-origin effects on seed development in Arabidopsis thaliana. *Development*, **125**(17), 3329–3341. 2

Smith, Christopher C., & Fretwell, Stephen D. 1974. The Optimal Balance between Size and Number of Offspring. *American Naturalist*, **108**, 499. 1

Trivers, R.L. 1974. Parent-Offspring Conflict. *American Zoologist*, **14**, 249–264. 2

Uma Shaanker, R., & Ganeshaiah, K. N. 1996. Polyembryony in plants: A weapon in the war over offspring numbers? *Trends in Ecology & Evolution*, **11**(1), 26–27. 2

Uma Shaanker, R., & Ganeshaiah, K. N. 1997. Conflict between parent and offspring in plants: Predictions, processes and evolutionary consequences. *Current Science*, **72**(12), 932–939. 2

Uma Shaanker, R., Ganeshaiah, K. N., & Bawa, K. S. 1988. Parent-Offspring Conflict, Sibling Rivalry, and Brood Size Patterns in Plants. *Annual Review of Ecology and Systematics*, **19**, 177–205. 2

Westoby, M., & Rice, B. 1982. Evolution of the Seed Plants and Inclusive Fitness of Plant-Tissues. *Evolution*, **36**(4), 713–724. 2, 13

Willson, MF, & Burley, N. 1983. *Mate Choice in Plants: Tactics, Mechanisms, and Consequences*. Princeton, New Jersey: Princeton university press. 2, 13

Wolfram, S. 1988. *Mathematica: a system for doing mathematics by computer*. Addison-Wesley Longman Publishing Co., Inc. Boston, MA, USA. 7

A Estimates

Parameter	Estimate	Parameter	Estimate	Parameter	Estimate
$D_{j=1}$	-613.524	$D_{i=1}$	11.308	a_1	0.000
$D_{j=2}$	-77439.500	$D_{i=2}$	11.403	a_2	-0.030
$D_{j=3}$	-61418.300	$D_{i=3}$	-3.255	a_3	0.058
$D_{j=4}$	-13679.200	$D_{i=4}$	2.876	a_4	0.089
$D_{j=5}$	42081.300	$D_{i=5}$	21.271	a_5	0.038
$D_{j=6}$	-77602.500	$D_{i=6}$	1.836	a_6	-0.021

Table 6: Simple estimates. 1=A619, 2=B1-Peru, 3=Mo17, 4=B73, 5=W22, 6=W23.

Female \ Male	A619	B1-Peru	Mo17	B73	W22	W23
A619	613.773	613.713	613.814	613.794	613.683	613.725
B1-Peru	77439.7	77439.6	77439.8	77439.7	77439.6	77439.6
Mo17	61418.5	61418.4	61418.6	61418.5	61418.4	61418.5
B73	13679.4	13679.3	13679.5	13679.4	13679.3	13679.4
W22	-42081.	-42081.2	-42081.	-42081.	-42081.2	-42081.1
W23	77602.8	77602.7	77602.8	77602.8	77602.7	77602.7

Table 7: D_{ij} estimates : male-female interactions in the competition-dependant term.

Male	B1-Peru	Mo17	B73	W22	W23
A619	-0.000107099	0.000075322	-0.000297057	0.0000257276	0.0000618033
B1-Peru		-0.000218047	-0.000159386	-0.00014808	-0.0000316529
Mo17			-0.000255826	-0.0000164673	-0.000252854
B73				-0.0000307757	-0.000367332
W22					0.000149686

Table 8: σ estimates : male-male interactions in the competition-dependant term.

Annexe C

Manuscrit : Masking effect, inbreeding and the evolution of double fertilization

Masking effect, inbreeding and the evolution of double fertilization

Cailleau A., Lenormand T.

October 29, 2010

1 Introduction

Double fertilization is in angiosperms the parallel fertilization of two cells of the gametophyte. The egg cell on the one hand, and the central cell on the other hand. The first fertilization results in the embryo, while the second results in its nutritive tissue, the endosperm. As a consequence of double fertilization, endosperm contains paternal genes as well as maternal genes. Moreover, depending on the number of maternal products (polar nuclei) present in the central cell, the endosperm of angiosperms can contain one or several sets of maternal genes. This nutritive tissue thus differs, from a genetic point of view, from the nutritive tissue of gymnosperms, which always contains only one set of maternal chromosomes. We hereafter note m a given set of maternal genes, and p a given set of paternal genes (all sets are exactly identical, i.e. derived from the same product of meiosis via mitosis). Following this notation, the embryo is $1m:1p$, the gymnosperm nutritive tissue is $1m:0p$, and the angiosperm nutritive tissue can be $1m:1p$ or $2m:1p$ depending on the number of polar nuclei in the central cell. The $0m:1p$ nutritive tissue is ancestral to seed plants, and double fertilization is a derived character of angiosperms (Friedman & Ryerson, 2009; Friedman, 2001a,b). The scenario $1m:0p \rightarrow 1m:1p \rightarrow 2m:1p$ is the more likely scenario of endosperm evolution (Cailleau *et al.* , 2010). We focus here in the first step of this scenario, that is the evolution of double fertilization.

We study double fertilization evolution in the light of masking effect. Masking effects were first studied in the case of life cycle evolution (Otto & Goldstein, 1992). These effects were shown to favor a longer diploid phase relative to the haploid phase when deleterious mutations are recessive and recombination is sufficiently high. In these model, selection can be decomposed on two terms. A term of direct selection, or direct masking effect, which favors diploidy over haploidy when deleterious mutations are recessives and inversely, and a term of indirect selection, or hitch-hiking, where the drawback associated to genetic load (deleterious mutations accumulate around genes raising diploidy) counterbalances the beneficial effect of masking if recombination is low.

The evolution of double fertilization, that is the transition from $1m:0p$ to $1m:1p$, could be considered at first sight to be a transition from haploidy to diploidy. In fact, this transition is different, because it results in higher level of paternal gene expression, but lower levels of maternal gene expression (asymmetrical gene expression). Any deleterious effect in the endosperm affects seed

survival. Hence, the embryo undergoes directly the selective pressures corresponding to the genotype of the endosperm. This results in a possible decoupling between the phenotype under selection (that depend on the endosperm genes) and the genotype transmitted ($1m : 1p$ embryo). We shown (Cailleau *et al.* , 2010) that this situation can result in a male-female conflict over the masking of deleterious mutations, because as in the case of life cycle evolution, the lowered expression of a given chromosome increases its genetic load, and can decrease its fitness in the long term. As a result, in a panmictic population, direct selection favor double fertilization when deleterious mutations are recessives, while indirect selection have opposite directions depending whether double fertilization is maternelly or paternally controlled (negative and positive, respectively). The effect of indirect selection is negligeable when recombination is high. As a result, double fertilization evolves when deleterious mutations are recessive or recombination is low in the case of paternal control, while it evolves when deleterious mutations are recessive and recombination is not low in the case of maternal control. This is a male-female masking conflict, that may occur each time the genes transmitted to the next generation are asymmetrically expressed, e.g. for all imprinted genes, or for further changes in the ratio of $m : p$ genes in the endosperm.

Selfing increases homozygosity, and consequently necessarily affects the masking of deleterous mutations. In life cycle evolution, inbreeding appears to lower recombination efficiency (Otto & Marks, 1996). In this paper, we investigate the effect of selfing for the evolution of double fertilization, and more generally, for the evolution of asymmetrical gene expression. We investigate sporophytic and gametophytic selfing. Under gametophytic selfing, we assume that a given proportion of fertilizations involve gametes produced by the same haploid individual (generating zygotes that are homozygous at all loci). Gametophytic selfing in fact corresponds to haploid parthenogenesis, which has been recorded in several angiosperms (Solanum nigram, Lilium spp., Orchis maculate, Nicotiana tabacum), but seems to occur sporadically in natural populations of plants. Under sporophytic selfing, a given proportion of fertilizations involves gametes produced by the same diploid individual (as in most selfing plants).

We find that with gametophytic selfing, double fertilization evolves in a wider range of parameter values than without selfing when paternally controled, and a narrower range of parameter values than without selfing when maternally controled. Thus, gametophytic selfing broadens the conditions where the genetic conflict over masking occurs. In the case of sporophytic selfing, maternal and paternal optima converge. In that case, whith full selfing, recombination has no more effect even when it is low, and double fertilization evolves when deleterious mutations are recessives, whatever recombination is. Hence, the male-female conflict disappears.

2 Methods

We consider two locus models exactly similar to Cailleau *et al.* (2010), except that a proportion f of embryos are produced by selfing. The a/A locus is a selected locus, while the m/M locus modifies the probability of double fertilization. Two alleles segregate at each locus (A, a, and m, M, respectively). At the modifier locus, eaither the maternal allele (model with maternal control)

or the paternal allele is expressed (model with paternal control). The probability of double fertilization is ν with allele M, and $\nu + \delta_\nu$ with allele m. At the selected locus, viability is 1 for A or AA endosperms, $1 - s$ for a or aa endosperms and $1 - hs$ for heterozygous Aa endosperms. Equations are first written for genotypes frequencies for each model (the four frequencies of gametes in the gametophytic model, and the 16 frequencies of embryo in the sporophytic model). We use the multi-locus formalism of Barton & Turelli (1991) to write recurrence equations for allele frequencies and for genetic associations, measuring the statistical associations among alleles at different loci and/or on different chromosomes. Definitions of the different parameters and variables are given Table 2. We denote p_a the frequency of allele a at locus a/A in the population, and $q_a = 1 - p_a$. Similarly, we denote p_m the frequency of allele m at locus m/M in the population, and $q_m = 1 - p_m$.

Endosperm make-up	Endosperm genotype (m:p)	Embryo's fitness
1m:0p	A : Ø	1
1m:0p	a : Ø	$1 - s$
1m:1p	A : A	1
1m:1p	A : a	$1 - hs$
1m:1p	a : a	$1 - s$

Table 1: Table of fitnesses

Gametophytic selfing: The life cycle with gametophytic selfing is illustrated Figure 1. Each generation starts after mutation, at the end of the haploid stage. With gametophytic selfing, the population can be described by four haploid genotype frequencies (p_{AM}, p_{aM}, p_{Am}, p_{am}) or by two allelic frequencies (p_a and p_m) and a linkage disequilibrium (noted C_{am}, $C_{am} = p_a p_m - p_{am}$). At syngamy, a proportion f of zygotes are formed by two gametes produced by the same haploid individual (and are thus homozygous at all loci), while a proportion $1 - f$ are formed by gametes sampled at random. Selection occurs during seed development and germination, i.e. at the beginning of the diploid phase. Fitness values are given Table 1. The life cycle continues with mutation. A proportion μ of a alleles mutates in A alleles. The life cycle ends with meiosis. We denote r the recombination rate between the two loci.

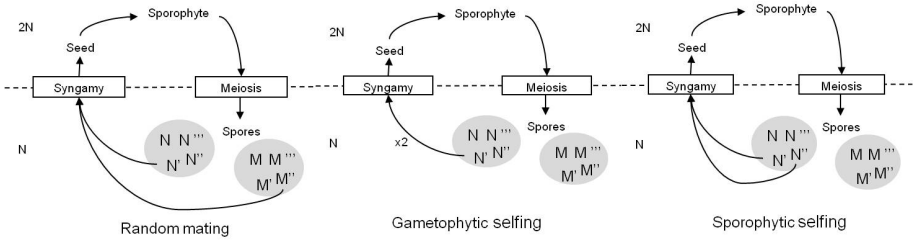

Figure 1: Life cycles with outbreeding, gametophytic selfing, and sporophytic selfing. Grey circles group the spores produced by different individuals.

Sporophytic selfing: The life cycle under sporophytic selfing is illustrated Figure 1. Each generation starts after mutation, at the end of the diploid stage. With sporophytic selfing, we describe the population by sixteen diploid genotype frequencies. Life cycle starts with meiosis (we still denotes recombination with r) and syngamy. A proportion f of fertilizations involves gametes produced by the same diploid individual, while a proportion $1-f$ involves gametes sampled at random from the whole population otherwise. The rest of the life cycle is similar to gametophytic selfing.

Exact simulations: we first assess the effect of selfing by studying the results of exact simulations. In these simulations, we start from a population where all genotype frequencies are equal, and we follow the change in frequency during 1000 generations. At that point, the frequency of the deleterious allele is stable and we record the direction of the change in frequency between the 985th generation and the 1000th generation. All the simulation results we present below are made with $s = 5.10^{-2}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, and $\delta_\nu = 10^{-2}$. For each graph, we simulated 51x11 points, which differ by values of h ($0 < h < 1$ by steps of 0.1) and r ($0 < r < 1/2$ by steps of 0.01).

Approximations: Full recursions are not given here for space reasons. We found useful approximations for the model with gametophytic selfing. These approximations can be obtained when selection is weak relative to recombination: in this case, genetic associations equilibrate fast and the population quickly reaches a state described as *quasi-linkage equilibrium*(QLE), where associations change slowly over time (Otto & Day, 2007). In principle, self-fertilization must not be too high for the QLE approximation to hold, because selfing reduces the effectiveness of recombination. The expression of the linkage disequilibrium at QLE is obtained by solving $C'_{am} = C_{am}$. We make some assumptions concerning the strength of selection and the frequency of self-fertilization; to do this, we introduce a small scaling term ξ and assume that the modifier effect δ_ν is of order ξ^n, while the selection coefficient s and the selfing rate f are of order ξ^n. At QLE, the order of the linkage disequlibrium will depend on the assumptions we make.

Notation	Corresponding variable or parameter
p_a	Frequency of a allele
p_m	Frequency of m allele
p_{am}	Frequency of am genotype
C_{am}	Linkage disequilibrium in the population of gametes
f	Selfing rate
r	Recombination rate
μ	Mutation rate
s	Difference in fitness between homozygous genotypes (AA, A vs aa, a)
h	Dominance of a allele
ν	Rate of double fertilization with M allele
δ_ν	Increase in the rate of double fertilization with m allele

Table 2: Table of variables and parameters

3 Results

3.1 Results of simulations

Gametophytic selfing: With gametophytic selfing, recombination efficiency seems to be diminished, since the effect of gametophytic selfing is the same as the effect of a lower recombination rate : double fertilization evolves in a reduced range of $\{r,h\}$ values as f increases with a maternal expression of the modifier locus, and in a wider range of $\{r,h\}$ values with a paternal expression of the modifier locus (Fig. 2). As a consequence, the extent of the male-female genetic conflict increases as f increases.

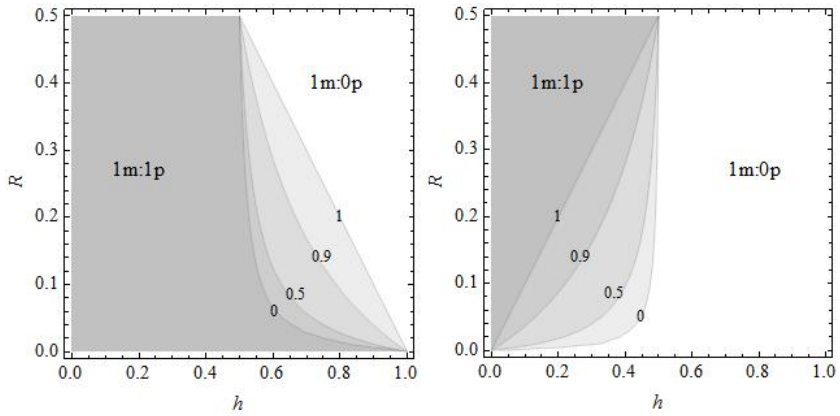

Figure 2: Simulation results for double fertilization evolution, with gametophytic selfing ($s = 5.10^{-2}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$.

Sporophytic selfing: With sporophytic selfing, the contrary happens. Double fertilization evolves in a wider range of $\{r,h\}$ values as f increases with a maternal expression of the modifier locus, and in a narrower range of $\{r,h\}$ with a paternal expression of the modifier locus (Fig. 3). As a consequence, the parameter range with a male-female genetic conflict is reduced as f increases, and when $f = 1$, the value of dominace under which double fertilization evolves does not depend on r anymore, it is $1/2$ (the same situation occurs with a biparental control of double fertilization, see Cailleau *et al.* , 2010). In that case, the conflict has disapeared.

3.2 Gametophytic selfing

In order to understand the main processes underlying the evolution of double fertilization with gametophytic selfing, we derive analytical results following the QLE approximation. However, graphical comparisons between those analytical results and simulation results indicate as expected that this approximation is not accurate when selfing is high (not shown).

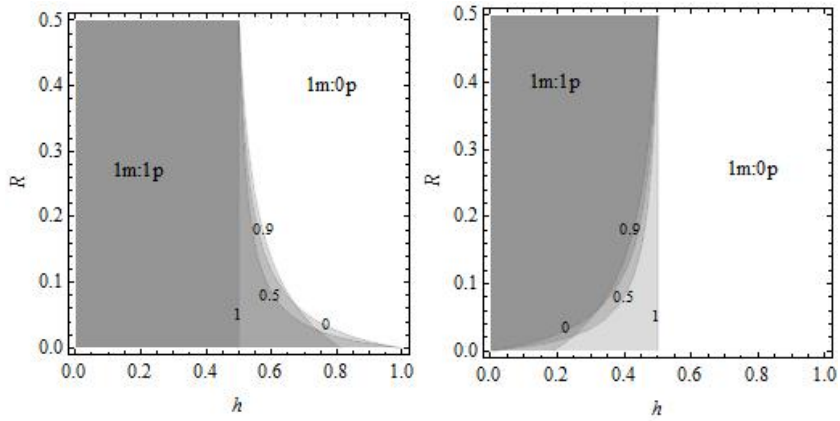

Figure 3: Simulation results for double fertilization evolution, with sporophytic selfing ($s = 10^{-1}$, $\mu = 10^{-3}$, $\nu = 10^{-1}$, $\delta_\nu = 10^{-2}$).

Change in frequency at a/A locus: In the case of gametophytic selfing, we calculate a selection-mutation equilibrium assuming $\mu << s$:

$$P_a^* = \frac{\mu}{\left[\left(\frac{1}{2}(1-\nu) + h\nu\right) + \frac{1}{2}f\left(1 + 2\nu\left(\frac{1}{2} - h\right)\right)\right]s} + o(\mu/s)^1. \tag{1}$$

As expected, less deleterious mutations are accumulated when selfing is more important and when dominance is high.

Change in frequency at m/M locus: Similarly, the change in frequency at modifier locus is given by :

$$\begin{aligned}\Delta p_m = &\left[\delta_\nu(1-f)\left(\frac{1}{2} - h\right)pq_{am}\right]s \\ &- \left[C_{am}\left(\frac{1}{2} - \nu\left(\frac{1}{2} - h\right) + f\left(\frac{1}{2} + \nu\left(\frac{1}{2} - h\right)\right)\right)\right]s.\end{aligned} \tag{2}$$

where $pq_{am} = (1-p_a)p_a(1-p_m)p_m$ and C_{am} is a pseudo-equilibrium calculated by QLE approximation, and can be written as :

$$C_{am}^{QLE} = \delta_\nu \frac{s(K-h)}{2r} pq_{am} \tag{3}$$

where $K = 1 - r$, r and $\frac{1}{2}$ for maternal, paternal, or biparental expression of the modifier locus, respectively.

We can see Eq. 2 that direct selection is proportional to the rate of outbreeding $(1-f)$. Hence, the strength of direct selection diminishes as selfing increases. The indirect effect is more complex than expected. In fact, inbreeding does not diminish recombination efficiency (in that case, $(1-f)$ would be found as a factor of r). The term $\left(\frac{1}{2} + \nu\left(\frac{1}{2} - h\right)\right)$ corresponds to the average exposition of deleterious alleles to selection in the population. This term is found as a factor of $1 + f$. Hence, selfing increases the average exposition by f. This increase in the exposition of alleles reinforces the effect of indirect selection. Overall,

inbreeding strengthens indirect selection and weakens direct selection. The result is that indirect selection counterbalances direct selection more easily with selfing, and thus broadens the range of parameter values where the male/female conflit occurs.

3.3 Sporophytic selfing

We were not able to derive analytical results for sporophytic selfing. Too much variables are involved and show first order effects. However, whe derived a mutation-selection equilibrium using the same approximations as those validated for the model with gametophytic selfing, in order to compare the two equilibria. This mutation selection equilibrium is :

$$P_a^* = \frac{\mu}{\left[\left(\frac{1}{2}(1-\nu)+h\nu\right)+\frac{1}{4}f\left(1+2\nu\left(\frac{1}{2}-h\right)\right)\right]s} + o(\mu/s)^1. \qquad (4)$$

Interestingly, that equilibrium look-like the equilibrium given Equation 1, except that the effect of inbreeding is halved. This appears to be logical, since deleterious alleles transmitted by a given heterozyote are always found at the homozygous state under gametophytic selfing, while they are found at the homozygous state in only half of the seeds produced via sporophytic selfing.

4 Discussion

In this article we model the masking of deleterious mutations in the endosperm, and its influence on the evolution of double fertilization, in the case where selfing occurs. More specifically, we followed the fate of alleles modifying the probability of occurence of double fertilization at a modifier locus with a sex-limited expression. In our models, a second locus is expressed in the endosperm, where deleterious mutations occur.

When selfing do not occur, double fertilization evolves when deleterious mutations are recessives in the case of biparental control of the modifier, when deleterious mutations are recessives *and* recombination between loci is not too low with maternal control of the modifier, and when deleterious mutations are recessives *or* recombination is low in the case of male control of the modifier. The results of the model with gametophytic selfing present important differences with those of the model with sporophytic selfing. With gametophytic selfing, double fertilization evolves in a reduced range of conditions of recombination and dominance as selfing increases with a maternal expression of the modifier locus, and in a wider range of conditions of recombination and dominance with a paternal expression of the modifier locus. With sporophytic selfing, globally, the contrary happens.

Modifier m alleles increase the rate of double fertilization, e.g., masking of deleterious mutations. As a consequence, when recombination is low, m alleles are genetically associated to deleterious a alleles. In the case of gametophytic selfing, a alleles are more frequently exposed and the cost for the m allele to be associated with a alleles increases. As this cost depends of recombination, it increases as recombination decreases.

In the case of sporophytic selfing, things are more complicated. The only difference in the contruction of the two models is that when the mother sporophyte

is heterozygote, there are embryos that are heterozygotes with sporophytic selfing, while there is no such embryos with gametophytic selfing. Hence the effects present in the model of gametophytic selfing are present in the model of sporophytic selfing, and there is an additional effect that is specific to the sporophytic model, due to the occurence of heterozygotes. Since the results of the two models are opposite, the direction of this specific effect is necessarily opposed to that of the effect shared by the two models.

The effect specific to sporophytic selfing is more important and counterbalances in most cases the effect shared by the two models. In some specific conditions (e.g. selfing equals 0.9, and recombination is not low, that is above the point where the curves cross), double fertilization evolves, as with gametophytic selfing, in a reduced range of conditions of recombination and dominance compared with low selfing and maternal expression of the modifier locus, and in a widened range of conditions of recombination and dominance compared with low selfing and paternal expression of the modifier locus. In that specific conditions, we can suspect it is the pressure shared with gametophytic selfing that predominates.

Models with gametophytic selfing are easier to build and study than models with sporophytic selfing. That is why models with gametophytic selfing are often used to assess of selfing on trait evolution. However, gametophytic selfing cannot be used as an approximation for sporophytic selfing in the specific case of double fertilization evolution. This is probably due to the fact that double fertilization modify heterozygosity, and that the occurence of heterozygotes is the difference between the two kinds of models. The similar problem occurs with evolution of recombination under sporophytic or gametophytic selfing (Roze & Lenormand, 2005). Hence, models of gametophytic selfing should be considered with caution.

According to the litterature, parent-offspring conflict is the more likely explanation for the evolution of double fertilization. However, parent-offspring conflict is not predicted to occur in selfing plants, and hardly account for the maintenance of double fertilization in all angiosperm famillies except Podostemaceae. According to our model of sporophytic selfing, double fertilization is advantageous in more and more conditions of dominance and recombination as selfing increases, in the case of maternal control. Double fertilization is advantageous in less and less conditions of dominance and recombination as selfing increases with a paternal control, but those conditions remain wider than the conditions found with maternal control. Hence, in the case where the plant is completely autogamous, the sufficient condition for double fertilization to be advantageous for both sexes is that deleterious mutation are recessives. Deleterious mutations tend to be recessive in diploids, with average dominance ranging between 0.1 and 0.4 depending on the studies (see e.g. Garcia-Dorado & Caballero, 2000; Vassilieva *et al.* , 2000; Szafraniec *et al.* , 2003), altough little data is availaible for plants. Hence, the heterozygosity theory is sufficient to explain the initial evolution and maintenance of double fertilization in both allogamous (Cailleau *et al.* , 2010) and autogamous (this study) species.

References

Barton, N.H., & Turelli, M. 1991. Natural and sexual selection on many

loci. *Genetics*, **127**, 229–255. 3

Cailleau, A., Cheptou, P.-O., & Lenormand, T. 2010. Ploidy and the Evolution of Endosperm of Flowering Plants. *Genetics*, **184**, 439–453. 1, 2, 5, 8

Friedman, W. E. 2001a. Comparative embryology of basal angiosperms. *Current Opinion in Plant Biology*, **4**(1), 14–20. 1

Friedman, W. E. 2001b. Developmental and evolutionary hypotheses for the origin of double fertilization and endosperm. *Comptes Rendus De L Academie Des Sciences Serie Iii-Sciences De La Vie-Life Sciences*, **324**(6), 559–567. 1

Friedman, W. E., & Ryerson, K. C. 2009. Reconstructing the Ancestral Female Gametophyte of Angiosperms: Insights from Amborella and Other Ancient Lineages of Flowering Plants. *American Journal of Botany*, **96**(1), 129–143. 1

Garcia-Dorado, A., & Caballero, A. 2000. On the Average Coefficient of Dominance of Deleterious Spontaneous Mutations. *Genetics*, **155**, 1991–2001. 8

Otto, S.P., & Day, T. 2007. *A Biologist's Guide to Mathematical Modeling in Ecology and Evolution*. Princeton University Press. 4

Otto, S.P., & Goldstein, D.B. 1992. Recombination and the evolution of diploidy. *Genetics*, **131**, 745–751. 1

Otto, S.P., & Marks, J.C. 1996. Mating systems and the evolutionary transition between haploidy and diploidy. *Biological Journal of the Linnean Society*, **57**, 197–218. 2

Roze, D., & Lenormand, T. 2005. Self fertilization and the evolution of recombination. *Genetics*, **170**, 841–857. 8

Szafraniec, K., Wloch, D. M., Sliwa, P., Borts, R. H., & Korona, R. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast Saccharomyces cerevisiae. *Genetical Research*, **82**(1), 19–31. 8

Vassilieva, L. L., Hook, A. M., & Lynch, M. 2000. The fitness effects of spontaneous mutations in Caenorhabditis elegans. *Evolution*, **54**(4), 1234–1246. 8

Annexe D

Manuscrit : Endosperm genetic make-up and inbreeding depression

Endosperm genetic make-up and inbreeding depression

Cailleau A., Lenormand T., Cheptou P.-O.

October 29, 2010

1 Introduction

Plants life-cycle alternates two multicellular life forms, the sporophyte ($2n$) that arise from fertilization and produces spores, and the gametophyte (n) that arises from meiosis and produces gametes. In seed plants, the life cycle is predominantly sporophytic, that is, diploid. In diploids, deleterious mutations may arise at the heterozygous state. In that case, they are generally partly masked, because most deleterious mutations are recessive in diploid tissues (dominances in fitness goes from 0.1 à 0.4, see e.g. Garcia-Dorado & Caballero, 2000; Vassilieva *et al.* , 2000; Szafraniec *et al.* , 2003). Deleterious mutations may arise at the homozygous state. In that case, deffects are more important but deleterious mutations are eliminated more easily from the population. Homozygosity increases with the rate of selfing. This is why autogamous populations are predicted to have a better purging process and a lower genetic load than allogamous ones (Lande & Schemske, 1985). As a consequence, when allogamous species are completely selfed, they are expected to have a greater decrease in fitness than self-fertilizing species (Lande & Schemske, 1985).

All this reasoning raises on the idea that genes have a diploid expression. However, during early developement, and more precisely, in the seed, it is not necessary the case. In gymnosperms, seed nutritive tissue is gametophytic. Only the set of maternal alleles inherited by the embryo are expressed in the endosperm (hereafter, we note m and call maternal contribution a set of maternal alleles inherited by the embryo). We note p (and call paternal contribution) a set of paternal alleles inherited by the embryo. According to this nomenclature, the endosperm of gymnosperms is $1m : 0p$. In angiosperms, the genetic make-up of the endosperm results from double fertilization on the one hand, and the kind of female gametophyte (or embryo sac) on the other hand. Double fertilization is specific to angiosperms. It is the parralel fertilization of the egg cell on the one hand, and of the mother cell of endosperm, the central cell, in the other hand (Guignard, 1899; Nawaschin, 1898). As a consequence of double fertilization, a paternal contribution participates to endosperm genetic make-up. There is seven basic types of embryo sac (Friedman *et al.* , 2008). Each type correspond to a different number of maternal contributions, of which some are inherited by the embryo (m), while some are not (hereafter m').

The seven basic types of embryo sac differ by the number of meiosis products (or megaspores) that participate to their formation, and the number of mitosis that follow meiosis. Let M_1, M_2, M_3 and M_4 be the four chromatids of two

chromosomes at the beginning of meiosis. M_1 and M_2 on the one hand and M_3 and M_4 on the other hand belong to the same chromosome and are identical. When the chromosomes go by pair, M_1 recombinates with M_3, and M_2 recombinates with M_4. As a consequence, if q is the probability of second division segregation (probability that depend on recombination), M_1 shares $1-q$ genes with M_2, no gene with M_3, and q genes with M_4. Let M_1 be the contribution inherited by the embryo. In most angiosperms, the embryo sac is monosporic. In that case, only $M1$ participates to its formation. In some angiosperms, the embryo sac is bisporic. In that case, M_1 and M_2 both participate to its formation. Eventually, in other angiosperms, the embryo sac is tetrasporic. In that case, all meiotic products may participate to its formation.

Two types of embryo sac are monosporic. One is Oenothera: M_1 undergoes two mitosis, and four nuclei are formed. One goes to the egg cell, two go to synergids, and one goes to the central cell. The endosperm of angiosperms with an Oenothera type of embryo sac (Austrobaileyales, Oenothera, Nymphéales) is $1m:1p$. The other is Polygonum : M_1 undergoes three mitosis, and eight nuclei are formed. One goes to the egg cell, two go to synergids, three go to antipods, and two go to the central cell. The endosperm of angiosperms with a Polygonum type of embryo sac (70% of angiosperm species) is $2m:1p$. Only one basic type of embryo sac is bisporic. It is the Allium type: M_1 and M_2 both undergo two mitosis. Among the four nuclei that result from the mitosis of M_1, one goes to the egg cell, two go to synergids, and one goes to the central cell. Among the four nuclei that result from the mitosis of M_2, one goes to the central cell, and the three others go to antipods. The endosperm of angiosperms with an Allium type of embryo sac is $2m:1p$ at some $(1-q)$ loci, and $1m1m':1p$ at other (q) loci. Four basic types of embryo sac are tetrasporic. Two of them are restricted to the Peperomia genus and lead to nonaploid or decapentaploid endosperms. We focus on the two others. One is the Adoxa type: all four products participate to the embryo sac, and each undergoes one mitosis. Among the two nuclei that result from the mitosis of M_1, one goes to the egg cell, and the other goes to the central cell. Among the four nuclei that result from the mitosis of M_3 and M_4, one goes to the central cell, and the three others go to antipods. If the second nuclei that goes to the central cell is derived from M_3, the endosperm is $1m1m':1p$. If the second nuclei that goes to the central cell is derived from M_4, the endosperm is $2m:1p$ at some (q) loci, and $1m1m':1p$ at other $(1-q)$ loci. Eventually, the last (tetrasporic) basic type of embryo sac is the Penaea type: all four product participate to the embryo sac, and each undergoes two mitosis. Among the four nuclei that result from the mitosis of M_1, one goes to the egg cell, two go to synergids, and one goes to the central cell. Among the four nuclei that result from the mitosis of each other spore, one goes to the central cell, and the others go to antipods. In that case, there is four nuclei in the central cell, and each derive from one of the meiotic products. The endosperm is $2m2m':1p$. Other types of embryo sac may arise (Drusa, Plumbago, etc.), but they result in the same endosperm genetic make-up as those seven types (Table 1, Friedman *et al.* , 2008).

The variations in the number of maternal and paternal contributions in the endosperm may affect gene expression, and as a consequence, may affect heterosis. It has been predicted that gymnosperm endosperm should be less heterotic than endosperms that result from monosporic gametophytes in angiosperms, which should be less heterotic than endosperms that results from bisporic ga-

Tissue	Genetic make-up of the embryo sac	Endosperm genetic make-up
Gymnosperm endosperm	M_1	$1m:0p$
Perisperm 2N	M_1M_2*	$1m1m':0p$
Monosporic 2N (Oenothera, Nuphar, Cotylédons)	$M1$	$1m:1p$
Monosporic 3N (Polygonum, Amborella)	$2M_1$	$2m:1p$
Bisporic 3N (Allium, Endymion, Drusa)	M_1M_2	$(1-q)2m:1p$, $q1m1m':1p$
Tetrasporic 3N (Adoxa, Drusa)	$M1M3$ or M_1M_4	$f_{M_3}1m1m':1p$, $f_{M_4}q2m:1p$, $f_{M_4}(1-q)1m1m':1p$
Tétrasporic 5N (Penaea, Plumbago, Fritillaria, Plumbagella)	$M_1M_2M_3M_4$	$2m2m':1p$
Tétrasporic 9N (Peperomia)	$2M_12M_22M_32M_4$	$4m4m':1p$
Tétrasporic 15N (Peperomia)	$2M_14M_24M_34M_4$	$6m8m':1p$

Table 1: The seven basic types of embryo sac. Each one result in a different genetic make-up in the endosperm. The star indicates that we refer to the chromatids before they recombinate. q is the probability of second division segregation. f_{M_3} and f_{M_4} are the frequency of M_3 and M_4 participating to central cell nuclei in Drusa type of embryo sac, respectively. $f_{M_3} = 1 - f_{M_4}$. Modified from Friedman *et al.* (2008).

metophytes, which should themselves be less heterotic than endosperms derived from tetrasporic gametophytes (Friedman *et al.* , 2008; Willson & Burley, 1983). This is hypothesized to be one of the possible causes for endosperm evolution (Brink & Cooper, 1940; Willson & Burley, 1983). However, the variations in the number of maternal and paternal contributions in the endosperm should also affect the frequency of deleterious mutations in the population (Cailleau *et al.* , 2010), and as a consequence, may affect inbreeding depression.

As seed characteristics are expected to impact seed and individual development, and as a consequence, fitness, we expect inbreeding to impact differentially the fitness depending on the genetic make-up of the endosperm.

In this paper, we derive population genetics predictions for inbreeding depression for allogamous populations, and for different numbers of maternal and paternal contributions. When then ask whether the variations in the number of maternal and paternal contributions can account for variations in inbreeding depression, and the extent to which heterozygosity theory is supported by this analysis.

2 Predictions of inbreeding depression

We consider a locus A/a where deleterious mutations a arise at a rate μ. The fitness of a given individual is determined by the genotype of its nutritive tissue. Selection occur because the genotype of that nutrtive tissue is correlated to the genotype of the individual. For example, in the case of a monosporic $2m:1p$ endosperm, the genotype is AAa if A is maternally inherited and a is paternally

inherited by the embryo. In the case of a polysporic $2m2m' : 1p$ endosperm, the genotype is $AAaaa$ if the mother sporophyte is Aa and the a allele is paternally inherited.

Individuals that are homozygous for the functional allele (the endosperm is A, AA, AAA, or $AAAAA$ depending on the number of contributions) have a fitness of 1. Individuals that are homozygous for the deleterious allele (a, aa, aaa or $AAAAA$) have a fitness of $1 - s$. In the case of a diploid $1m : 1p$ endosperm, heterozygous individuals (aA) have a fitness $1 - hs$. In the case of a triploid $2m : 1p$ or $1m1m'1p$ endosperm, heterozygous individuals have a fitness $1 - h_1 s$ if the genotype is aaA, and $1 - h_2 s$ if it is aAA. In the case of a pentaploid $2m2m' : 1p$ endosperm, heterozygous individuals have a fitness $1 - h_A s$ if the genotype is $AAAAa$, $1 - h_B s$ if the genotype is $AAaaA$, $1 - h_C s$ if the genotype is $AAaaa$, and $1 - h_D s$ if the genotype is $aaaaA$ (Table 2).

Genotype	Fitness	Ratio a/n
A, AA, AAA, $AAAAA$	1	1
$AAAAa$	$1 - h_A s$	$4/5 = 0.8$
AAa	$1 - h_2 s$	$2/3 = 0.66$
$AAaaA$	$1 - h_B s$	$3/5 = 0.6$
Aa	$1 - hs$	0.5
$AAaaa$	$1 - h_C s$	$2/5 = 0.4$
Aaa	$1 - h_1 s$	$1/3 = 0.33$
$aaaaA$	$1 - h_D s$	$1/5 = 0.2$
a, aa, aaa, $aaaaa$	$1 - s$	0

Table 2: Fitnesses as a function of genotypes. Genotypes are classified following an increasing a/n ratio (number of a allele/ ploidy). We hypothesize dominances values are classified the same way.

Our goal is to predict the values of inbreeding depression that would result from the different kind of endosperm. Do do this, we first calculate mutation-selection equilibria for populations that exhibit different genetic make-up for their endosperm. We then write the average fitness of these populations as a function of these equilibria, which allows us to write the values of inbreeding depression corresponding to these genetic make-up.

2.1 Mutation-selection equilibria

In order to find mutation-selection equilibria, we model the evolution of the frequencies of embryos (AA, Aa, aa) in panmictic populations exhibiting the various genetic make-up we described before. We follow the different events in the life cycle over one generation. We start at the sporophytic stage. Meiosis occurs and produces gametes. Those gametes fuse and form embryos. The embryos undergo selection on the basis of the genotypes of their nutritives tissues. Then mutation occurs (A mutates to a with a probability μ), and the cycle starts again.

These models allow us to write exact recursions describing the change in frequency for the a allele in each population. These recursions are complicated expressions. To simplify them, we assume that μ/s is small. We then solve these recursions for the change in frequency to be null, which provides the seven

mutation-selection equilibria:

$$p_a^{1m0p} = \frac{2\mu}{s} + O(\mu/s)^1 \tag{1}$$

$$p_a^{1m1p} = \frac{\mu}{hs} + O(\mu/s)^1 \tag{2}$$

$$p_a^{2m1p} = \frac{2\mu}{(h_1+h_2)s} + O(\mu/s)^1 \tag{3}$$

$$p_a^{1m1m'1p} = \frac{\mu}{h_2 s} + O(\mu/s)^1 \tag{4}$$

$$p_a^{2m2m'1p} = \frac{2\mu}{(h_A+h_B)s} + O(\mu/s)^1 \tag{5}$$

With the $1m:1p$ nutritive tissue, the tissue is diploid and we retrieve the classic equilibrium frequency of a deleterious mutation under diploid selection. With the $1m:0p$ nutritive tissue, the tissue is haploid but the equilibrium is twice the classical equilibrium of a deleterious mutation under haploid selection. Indeed, the efficiency of selection is halved compared to the classical case of haploid selection, because paternal alleles are not exposed to selection (they are in the embryo, but not in the endosperm, see Cailleau *et al.* , 2010 for details). With the $2m:1p$ nutritive tissue, the mutation selection equilibrium ressemble a lot that with a $1m:1p$ endosperm, with the difference that h is replaced by the average values of triploid dominances, $h_m = (h_1+h_2)/2$.

With the $1m1m':1p$ nutritive tissue, the mutation selection equilibrium does not depend on all the values of dominance. It depends only on the dominance of a in the $AA:a$ endosperm, h_2. This is due to the fact that aaA endosperms are rare compared to AAa endosperm (in the best case, i.e. if selection is very low, the frequency of aaA endosperms is $(1-p_a)p_a^2$, while the frequency of AAa endosperms is $(1-p_a)^2 p_a$, with $p_a < 1-p_a$)

With the $2m2m':1p$ nutritive tissue, the mutation selection equilibrium again does not depend on all the values of dominance. It depends only on the dominance of a in the $AAAAa$ and $AAaaA$ endosperms, h_A and h_B. This is due to the same reason, that aa mother sporophytes are rare because of selection at the preceding generation.

2.2 Mean fitnesses at mutation-selection equilibirum

Let W_O be the average fitness in a given population after one generation of outbreeding, and W_S be the average fitness in a given population after one generation of selfing. We have :

$$W_O^{1m0p} = (1-p_a)^2 + p_a(1-p_a) + \left[(p_a^2 + p_a(1-p_a)\right](1-s) \tag{6}$$

$$W_O^{1m1p} = (1-p_a)^2 + 2p_a(1-p_a)(1-hs) + p_a^2(1-s) \tag{7}$$

$$\begin{aligned} W_O^{2m1p} &= (1-p_a)^2 + p_a(1-p_a)(1-h_1 s) \\ &\quad + p_a(1-p_a)(1-h_2 s) + p_a^2(1-s) \end{aligned} \tag{8}$$

$$W_O^{1m1m'1p} = f_{11}(1-p_a) + (f_{22}(1-p_a) + f_{12}p_a + f_{21}p_a)(1-h_1 s) \tag{9}$$
$$+ (f_{12}(1-p_a) + f_{21}(1-p_a) + f_{11}p_a)(1-h_2 s) + f_{22}p_a(1-s)$$

$$W_O^{2m2m'1p} = f_{11}(1-p_a) + f_{11}p_a(1-h_A s) \tag{10}$$
$$+ (f_{12}(1-p_a) + f_{21}(1-p_a))(1-h_B s)$$
$$+ (f_{12}p_a + f_{21}p_a)(1-h_C s) + f_{22}(1-p_a)(1-h_D s) + f_{22}p_a(1-s)$$

$$W_S^{1m0p} = W_O^{1m0p} \tag{11}$$

$$W_S^{1m1p} = (1-p_a)^2 + \frac{1}{4}p_a(1-p_a) \tag{12}$$
$$+ p_a(1-p_a)(1-hs) + \left[p_a^2 + \frac{1}{4}p_a(1-p_a)\right](1-s)$$

$$W_S^{2m1p} = (1-p_a)^2 + \frac{1}{4}p_a(1-p_a) + \frac{1}{2}p_a(1-p_a)(1-h_1 s) \tag{13}$$
$$+ \frac{1}{2}p_a(1-p_a)(1-h_2 s) + \left[p_a^2 + \frac{1}{4}p_a(1-p_a)\right](1-s)$$

$$W_S^{1m1m'1p} = f_{11} + \left(\frac{f_{12}}{2} + \frac{f_{21}}{2}\right)(1-h_1 s) \tag{14}$$
$$+ \left(\frac{f_{12}}{2} + \frac{f_{21}}{2}\right)(1-h_2 s) + f_{22}(1-s)$$

$$W_S^{2m2m'1p} = f_{11} + \left(\frac{f_{12}}{2} + \frac{f_{21}}{2}\right)(1-h_B s) \tag{15}$$
$$+ (f_{12}p_a + f_{21}p_a)s(1-h_C s) + f_{22}(1-s)$$

where f_{11} is the frequency of AA maternal sporophytes, f_{12} and f_{21} are the frequency of Aa (A maternally inherited) and aA (A paternally inherited) maternal sporophytes, and f_{22} is the frequency of aa maternal sporophytes. Maternal sporophytes have undergone selection, but if we assume that s is low, we can suppose $f_{11} = (1-p_a)^2$, $f_{12} = (1-p_a)p_a$, $f_{21} = (1-p_a)p_a$ and $f_{22} = p_a^2$. In that case, we have :

$$W_O^{1m1m'1p} = (1-p_a)^3 + 3(1-p_a)p_a^2(1-h_1 s) \tag{16}$$
$$+ 3(1-p_a)^2 p_a(1-h_2 s) + p_a^3(1-s)$$

$$W_O^{2m2m'1p} = (1-p_a)^3 + (1-p_a)^2 p_a(1-h_A s) + 2(1-p_a)^2 p_a(1-h_B s) \tag{17}$$
$$+ 2(1-p_a)p_a^2(1-h_C s) + (1-p_a)p_a^2(1-h_D s) + p_a^3(1-s)$$

$$W_S^{1m1m'1p} = (1-p_a)^2 + (1-p_a)p_a(1-h_1 s) \tag{18}$$
$$+ (1-p_a)p_a(1-h_2 s) + p_a^2(1-s)$$

$$W_S^{2m2m'1p} = (1-p_a)^2 + (1-p_a)p_a(1-h_B s) \tag{19}$$
$$+ 2(1-p_a)p_a^2 s(1-h_C s) + p_a^2(1-s)$$

2.3 Inbreeding depression at specific loci

2.3.1 Values of inbreeding depression

Inbreeding depression is defined as :

$$\delta = 1 - \frac{W_S}{W_O} \tag{20}$$

Therefore, in all the allogamous populations we described, inbreeding depression δ is, at mutation-selection equilibrium :

$$\delta^{1m0p} = 0 \tag{21}$$

$$\delta^{1m1p} = \frac{\mu(1-2h)}{2h} \tag{22}$$

$$\delta^{2m1p} = \frac{\mu(1-h_1-h_2)}{h_1+h_2} \tag{23}$$

$$\delta^{1m1m'1p} = \frac{\mu(h_1-2h_2)}{h_2} \tag{24}$$

$$\delta^{2m2m'1p} = \frac{2\mu(h_C-h_A-h_B)}{h_A+h_B} \tag{25}$$

2.3.2 Locus-specific predictions

As classically envisionned, in diploid $1m:1p$ cases, there is no inbreeding depression when $h = 1/2$. Deleterious mutations are generally recessive. In these conditions (i.e. $h < 1/2$), inbreeding depression is more important as h is low. In the haploid $1m:0p$ case, there is no inbreeding depression.

In the triploid $2m:1p$ case, inbreeding depression follows the same function as with $1m:1p$, with the difference that $h_m = (h_1+h_2)/2$ replaces h. On should expect h to be more or less intermediary between h_1 and h_2 (Table 2). Hence, h should not be very different from h_m, and inbreeding depression in species with a $1m:1p$ nutritive tissue should not be very different from inbreeding depression in species with a $2m:1p$ nutritive tissue.

In a $1m1m':1p$ nutritive tissue, inbreeding depression is null if $h_1-2h_2=0$, that is, if $h_2 = h_1/2$. Inbreeding depression is particularly high if h_2 is low.

$2m:1p$ and $1m1m':1p$ are both triploid cases, so that the predictions for their values of inbreeding depression involve the same parameters. Hence, we can try to assess the conditions where inbreeding depression in one of these cases outweights inbreeding depression in the other cases. Inbreeding depression with $1m1m':1p$ is superior to inbreeding depression with $2m:1p$ if :

$$\frac{\mu(h_1-2h_2)}{h_2} > \frac{\mu(1-h_1-h_2)}{h_1+h_2} \tag{26}$$

$$h_1^2 - h_2 - h_2^2 > 0 \tag{27}$$

Whether or not that condition is attained depends accurately on the unknown values of h_1 and h_2. Hence, we cannot make a general prediction.

In the case where the nutritive tissue is $2m2m' : 1p$, the variations of inbreeding depression are not comparable with other expressions.

Globally, if we do not make assumptions for the values of dominances, we can only predict that the lowest (null) values of inbreeding depression are expected with $1m : 0p$. All other cases correspond to non-null levels of inbreeding depression. However, the relative levels of inbreeding depression for these tissues vary with parameters for which estimations are currently unavailaible. In the following section, we make some assumptions for the values of dominance, which allows us to make qualitative predictions for inbreeding depression.

2.3.3 Biological predictions

The metabolic theory of dominance (Kacser & Burns, 1981) stipulates that mutations change the rate of metabolic activity, which in turn determines dominance. According to this theory, metabolic activity follows a saturating fonction of catalytic products. Hence, even if the homozygous AA genotype produces 100% of catalytic products, the homozygous aa genotype 0% of catalytic products, and the heterozygous Aa genotype 50% of catalytic products, metabolic activity is not halved in Aa genotypes, but only reduced of a quantity σ, which results in a phenotype altered at a level σ. σ is the dominance of the deleterious allele in the heterozygous Aa genotype.

Using the assumptions of the metabolic theory of dominance, we can model the fitness of a genotype as a function of the ratio of the deleterious allele a in the genotype (R) as :

$$1 - \sigma s = 1 - s + s\left(1 - R^{\frac{\text{Log}(1/h)}{\text{Log}[2]}}\right) \tag{28}$$

That function is represented Figure 1 assuming h (the dominance of a in a diploid Aa genotype) is 0.25. We choose that value because it is the average value of all known measures of dominance available so far (T. Lenormand, pers. comm.).

If we solve the function of Equation 28 for σ with $R = a/n$ (a/n given Table 2), which gives $h_A = 0.04$, $h_2 = 0.1089$, $h_B = 0.16$, $h_c = 0.36$, $h_1 = 0.4356$, $h_D = 0.64$.

When pluging these values of dominance in the predictions of inbreeding depression Equation 22-25, whe can estimate that $\delta^{1m1p} \approx \mu$, $\delta^{2m1p} \approx 0.8\mu$, $\delta^{1m1m'0p} \approx 2\mu$, $\delta^{2m2m'0p} \approx 1.6\mu$. Hence, we can predict that $\delta^{1m0p} < \delta^{1m1p} < \delta^{2m1p} < \delta^{1m1m'0p} < \delta^{2m2m'0p}$.

2.4 Inbreeding depression in the endosperm

Our locus-specific predictions can be extended to predict inbreeding depression at the level of the genome. Let $U = \mu$ be the rate of genomic mutation. If the mutation rate at loci do not varies a lot across the genome, we can predict that :

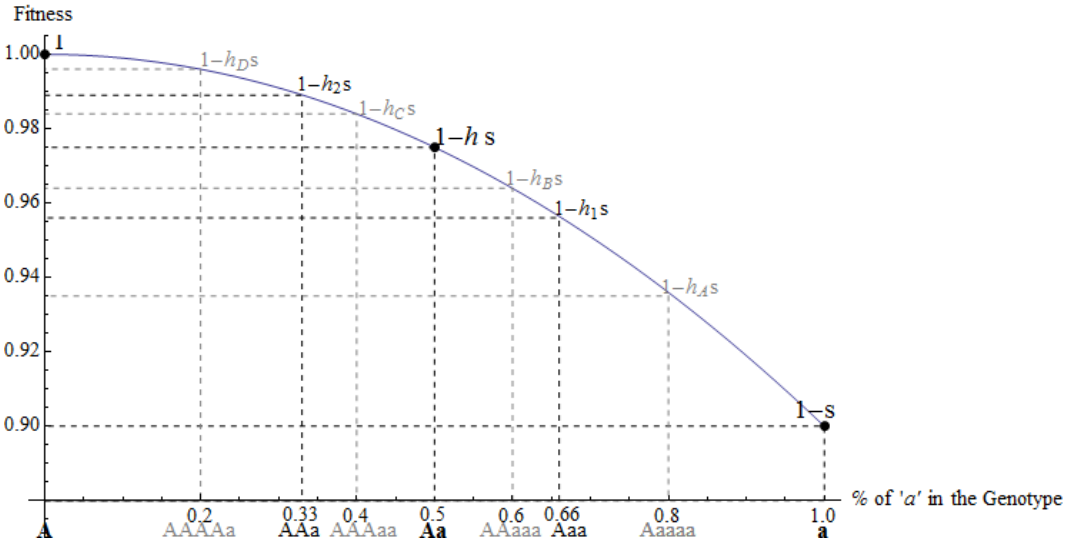

Figure 1: Illustration of the metabolic theory of dominance, with $s = 10^{-1}$. Homozygous genotypes for A have the best fitness, fixed to 1. Homozygous genotypes for a have the worst fitness, fixed to $1-s$. Heterozygous Aa genotypes have a fitness $1 - 0.25s$. These informations allow to draw the function between dominance and fitness, and to estimate dominance in non-diploid tissues.

$$\delta^{Gymnosperms} = 0 \tag{29}$$

$$\delta^{Oenothera} = \frac{U(1-2h)}{2h} = U \tag{30}$$

$$\delta^{Polygonum} = \frac{U(1-h_1-h_2)}{h_1+h_2} = 0.8U \tag{31}$$

$$\delta^{Penea} = \frac{2U(h_C-h_A-h_B)}{h_A+h_B} = 1.6U \tag{32}$$

With Allium and Adoxa types of embryo sac, some loci are $1m1m' : 1p$, while others are $2m : 1p$. With the Allium type of embryo sac, $1m1m' : 1p$ should occur with a frequency $1-q$ and $2m : 1p$ with a frequency q. As a consequence, inbreeding depression with the Allium type of embryo sac should be :

$$\delta^{Allium} = q\frac{U(1-h_1-h_2)}{h_1+h_2} + (1-q)\frac{U(h_1-2h_2)}{h_2}; \tag{33}$$

$$\delta^{Allium} = 0.8qU + 2(1-q)U; \tag{34}$$

In the case where the recombination rate is very important, we expect $q = 1/2$, so that $\delta^{Allium} = 1.4U$. In the case were the recombination rate is very low, $q = 0$, and $\delta^{Allium} = 2U$.

With the Adoxa type of embryo sac, $1m1m' : 1p$ should occur with a frequency of 1 if this is M_3 that participates to central cell nuclei, and $1-q$ if this is M_4 that participates to central cell nuclei (Table 1). If we assume that M_3 participates to the genetic make-up of the central cell as often as M_4, $1m1m' : 1p$ should occur with a frequency $(1+q)/2$. As a consequence, inbreeding depression with the Adoxa type of embryo sac should be :

$$\delta^{Adoxa} = \left(1-\frac{1+q}{2}\right)\frac{U(1-h_1-h_2)}{h_1+h_2} + \frac{1+q}{2}\frac{U(h_1-2h_2)}{h_2}; \tag{35}$$

$$\delta^{Adoxa} = 0.8\left(1-\frac{1+q}{2}\right)U + 2\frac{1+q}{2}U; \tag{36}$$

In the case where the recombination rate is very important, $\delta^{Adoxa} = 1.7U$. In the case were the recombination rate is very low, $q = 0$, and $\delta^{Adoxa} = 1.4U$.

To sum up, we can predict :

$$\delta^{Gymnosperms} = 0 \tag{37}$$

$$\delta^{Oenothera} = U \tag{38}$$

$$\delta^{Polygonum} = 0.8U \tag{39}$$

$$\delta^{Adoxa} = 1.4U - 1.7U \tag{40}$$

$$\delta^{Allium} = 1.4U - 2U \tag{41}$$

$$\delta^{Penea} = 1.6U \tag{42}$$

As a consequence, we can predict that inbreeding depression with a polysporic embryo sac (Adoxa, Allium, or Penea) is expected to outweigh inbreeding depression with a monosporic embryo sac (Oenothera or Polygonum).

3 Test

There is various measures of inbreeding depression in the litterature. The protocol is well defined : in a natural population, a group of individuals is selfed, while another group is crossed. Proxi of population fitnesses are taken from seeds and embryos that result from these fertilizations (seed weight, survival, etc.), and inbreeding depression is calculated as defined Equation 20.

The survival and development of the embryo depend not only on the genotype of the embryo, but also on the genotype of the endosperm (Brink & Cooper, 1940, 1947). As a consequence, we expect these measures of inbreeding depression to be affected by the type of endosperm. More precisely, we expect measures of inbreeding depression to be more affected by the type of endosperm as the stage of development is precocious, because early stages of development should be under depeer endosperm control.

We compiled a dataset (Table 4, Appendix) for 21 angiosperm species (Fig. 3 in the Appendix) and twelve gymnosperm species (*Abies procera*, *Picea glauca*, *Picea mariana*, *Pinus banksiana*, *Pinus elliottii*, *Pinus ponderosa*, *Pinus resinosa*, *Pinus strobus*, *Pinus sylvestris*, *Pinus taeda*, *Pinus thunbergii*, *Thuja plicata*). This dataset is informed for the type nutritive tissue (1m0p, Monosporic, Bisporic, or Tetrasporic, angiosperm data from Carman, 1997), the rate of outbreeding (C.G. Eckert database), and inbreeding depression at different stages (Seed, Germination, Survival, Fecundity, Nescent group unpub. dataset), for several populations of each species. Values of inbreeding depression at seed stage correspond to a compilation of values of inbreeding depression for the number of seed per plant, or, when not available, the proportion of fruits set from each plant, or, when not available, seed weight. Values of inbreeding depression at germination stage correspond to the proportion of seeds germinating. Values of inbreeding depression at survival stage correspond to the proportion of plants that attain maturity. Values of inbreeding depression at fecundity stage correspond to a compilation of values of inbreeding depression for the seed set, and when not available, vegetative size.

Since the values of inbreeding depression can be supposed to evolve rapidly, and the data is widely distributed among the phylogeny, we analyse the dataset by performing ANOVAs. We describe inbreeding depression (δ) as a function of the models given Table 3. We base the comparison of these models on their AIC.

Stage	Seed	Germination	Survival	Fecundity
Number of Species	25	18	14	18
Number of populations	52	38	48	49
Modele	AIC			
δ = OutRate + NutritiveTissue + NutritiveTissue:OutRate	5.804428	-34.07291	-25.78609	NA
δ = OutRate + NutritiveTissue	6.662957	**-37.71250**	-25.54413	3.149749
δ = OutRate	30.87999	-32.57680	-28.50059	5.346682
δ = NutritiveTissue	**5.126904**	-31.59878	-27.53949	**1.696773**
δ = 1	30.48778	-30.97805	**-30.31895**	3.54736

Table 3: Models. All models include a random effect that correspond to populations of each species.

The best model for seed stage and fecundity is the model where the nutritive

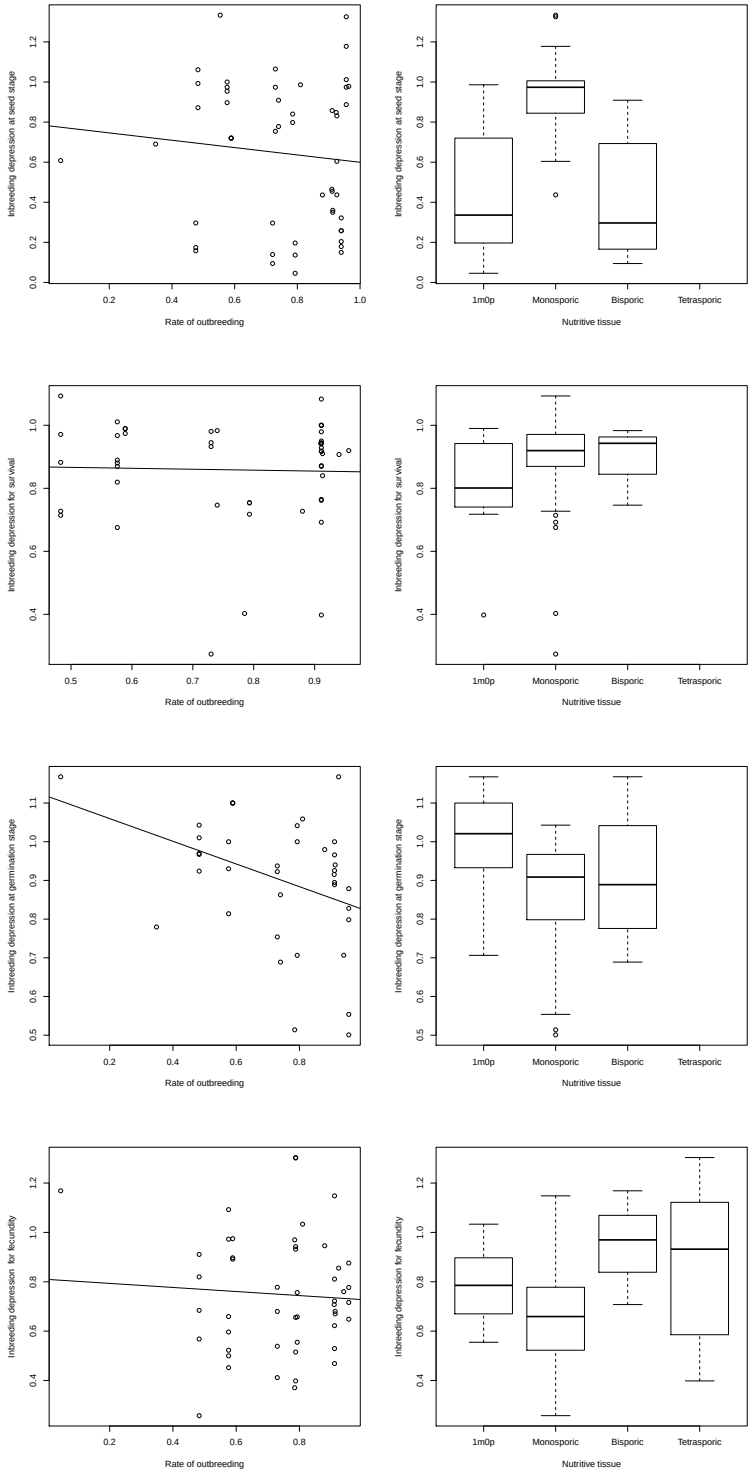

Figure 2: Inbreeding depression as a function of breeding system and the type of nutritive tissue.

tissue explains more variability than the breeding system. The best model for germination is the model where both the nutritive tissue and the breeding system explain variability. We can see how inbreeding depression varies as a function of breeding system and the type of nutritive tissue in Figure 2.

4 Discussion

We predict that the genetic make-up of the endosperm should affect inbreeding depression, in particular for early stages of development. More precisely, if we follow the assumptions of the metabolic theory of dominance, we can predict that inbreeding depression should be lower in gymnosperms than in angiosperms species with a monosporic embryo sac, and that inbreeding depression in angiosperm species with bi and tetrasporic embryo sacs should outweight inbreeding depression of gymnosperms species and angiosperms species with a monosporic embryo sac.

We analyse a dataset and show that as expected, the genetic make-up of the endosperm affects inbreeding depression at seed and germination stage, but not adulthood survival. The genetic make-up of the endosperm also affect inbreeding depression at a late stage, fecundity. This may be considered surprising at first sight, but fecundity involve measures of seed set. Now, seed set may be more affected by the fact that the sporophyte is inbred in species whith a polysporic nutritive tissue, because in this case the nutritive tissue is made of the two related maternal contributions, than in species whith a monosporic nutritive tissue, because in this case the genetic contributions in the nutritive tissue are no more related. Globally, these results are thus consistent with our prediction that inbreeding depression should vary with the type of nutritive tissue.

They indicate that changes in heterozygosity have to be taken in account to understand the evolution of seed nutritive tissues.

References

Brink, R. A., & Cooper, D. C. 1940. Double fertilization and development of the seed in angiosperms. *Bot. Gaz.*, **102**, 1–25. 3, 11

Brink, RA, & Cooper, DC. 1947. The endosperm in seed development. *Botanical review*, **132**, 423–541. 11

Cailleau, A., Cheptou, P.-O., & Lenormand, T. 2010. Ploidy and the Evolution of Endosperm of Flowering Plants. *Genetics*, **184**, 439–453. 3, 5

Carman, J. G. 1997. Asynchronous expression of duplicate genes in angiosperms may cause apomixis, bispory, tetraspory, and polyembryony. *Biological Journal of the Linnean Society*, **61**(1), 51–94. 11

Friedman, W. E., Madrid, E. N., & Williams, J. H. 2008. Origin of the fittest and survival of the fittest: Relating female gametophyte development to endosperm genetics. *International Journal of Plant Sciences*, **169**(1), 79–92. 1, 2, 3

Garcia-Dorado, A., & Caballero, A. 2000. On the Average Coefficient of Dominance of Deleterious Spontaneous Mutations. *Genetics*, **155**, 1991–2001. 1

Guignard, L. 1899. Sur les anthérozoides et la double copulation sexuelle chez les végétaux angiospermes. *C. R. Acad. Sci. Paris*, **128**, 864–871. 1

Kacser, H., & Burns, J. A. 1981. The Molecular-Basis of Dominance. *Genetics*, **97**(3-4), 639–666. 8

Lande, R., & Schemske, D. W. 1985. The Evolution of Self-Fertilization and Inbreeding Depression in Plants .1. Genetic Models. *Evolution*, **39**(1), 24–40. 1

Nawaschin, S.G. 1898. Resultate einer Revision der Befruchtungsvorg´ange bei Lilium Martagon und Fritillaria tenella. *Bull. Acad. Sci. St Petersburg*, **9**, 377–382. 1

Szafraniec, K., Wloch, D. M., Sliwa, P., Borts, R. H., & Korona, R. 2003. Small fitness effects and weak genetic interactions between deleterious mutations in heterozygous loci of the yeast Saccharomyces cerevisiae. *Genetical Research*, **82**(1), 19–31. 1

Vassilieva, L. L., Hook, A. M., & Lynch, M. 2000. The fitness effects of spontaneous mutations in Caenorhabditis elegans. *Evolution*, **54**(4), 1234–1246. 1

Webb, C. O., & Donoghue, M. J. 2005. Phylomatic: tree assembly for applied phylogenetics. *Molecular Ecology Notes*, **5**(1), 181–183. 15

Willson, MF, & Burley, N. 1983. *Mate Choice in Plants: Tactics, Mechanisms, and Consequences*. Princeton, New Jersey: Princeton university press. 3

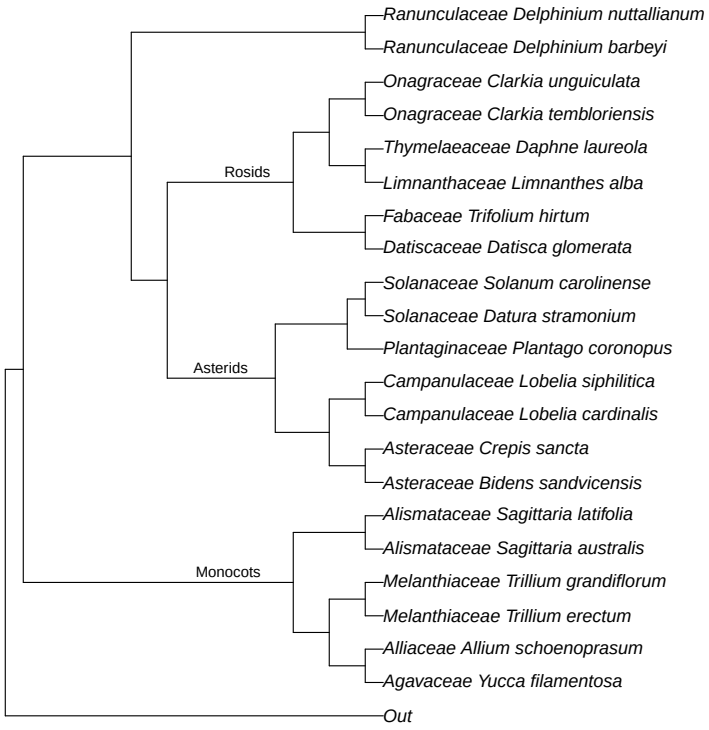

Figure 3: Phylogeny of the angiosperm species used in the analysis. The relationships between families were obtained from Webb & Donoghue (2005), at Phylocom

Species	OutRate	NutritiveTissueQualit	ID.Seed	ID.Germination	ID.Survival	ID.Fecundity
Delphinium nuttallianum	0.926	Monosporic	0.624	-	-	-
Delphinium barbeyi	0.554	Monosporic	1.333	-	-	-
Clarkia unguiculata	0.964	Monosporic	0.978	-	-	-
Clarkia tembloriensis	0.483	Monosporic	0.996	0.983	0.878	0.648
Daphne laureola	0.348	Monosporic	0.690	0.780	-	-
Limnanthes alba	0.788	Tetrasporic	-	-	-	0.864
Datisca glomerata	0.785	Bisporic	0.840	-	-	0.970
Datura stramonium	0.045	Bisporic	0.608	1.168	-	1.169
Plantago coronopus	0.576	Monosporic	0.975	0.963	0.873	0.685
Lobelia cardinalis	0.730	Monosporic	0.966	0.858	0.783	0.602
Crepis sancta	0.911	Monosporic	0.858	0.937	0.929	0.717
Bidens sandvicensis	0.785	Monosporic	0.798	0.514	0.403	0.371
Sagittaria latifolia	0.740	Bisporic	0.843	0.776	0.865	-
Trillium grandiflorum	0.721	Bisporic	0.177	-	-	-
Trillium erectum	0.476	Bisporic	0.210	-	-	-
Allium schoenoprasum	0.910	Bisporic	0.465	0.915	0.943	0.708
Yucca filamentosa	0.956	Monosporic	1.075	0.712	0.920	0.754
Picea glauca	0.793	1m0p	0.126	0.916	0.742	0.656
Picea mariana	0.911	1m0p	0.455	0.926	0.580	0.717
Pinus banksiana	0.880	1m0p	0.437	0.980	0.727	0.946
Pinus elliottii	0.940	1m0p	0.150	0.707	0.907	0.760
Pinus ponderosa	0.913	1m0p	0.355	0.940	0.875	0.675
Pinus resinosa	0.810	1m0p	0.986	1.059	-	1.033
Pinus strobus	0.924	1m0p	0.848	1.168	-	0.855
Pinus sylvestris	0.940	1m0p	0.244	-	-	-
Thuja plicata	0.589	1m0p	0.720	1.100	0.984	0.921

Table 4: Table of variables values. The values of inbreeding depression are averaged across populations.

Annexe E

Tables des valeurs sélectives de la partie modèles

Les différents types d'embryons possibles sont :

AM|AM	aM|AM	Am|AM	am|AM
AM|aM	aM|aM	Am|aM	am|aM
AM|Am	aM|Am	Am|Am	am|Am
AM|am	aM|am	Am|am	am|am

Avec à gauche, le gamete reçu maternellement et à droite, reçu paternellement par le sporophyte n

Leurs fréquences sur chaque plante dépend du génotype du sporophyte n-1, et des fréquences des gamètes de chaque type (f_t) dans la population de pollen.

		Gamete reçu maternellement par le sporophyte n-1															
		AM				aM				Am				am			
Gamete reçu paternellement par le sporophyte n-1	AM	f_{AM}	0	0	0	$f_{AM}/2$	$f_{AM}/2$	0	0	$f_{AM}/2$	0	$f_{AM}/2$	0	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$
		f_{aM}	0	0	0	$f_{aM}/2$	$f_{aM}/2$	0	0	$f_{aM}/2$	0	$f_{aM}/2$	0	$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$
		f_{Am}	0	0	0	$f_{Am}/2$	$f_{Am}/2$	0	0	$f_{Am}/2$	0	$f_{Am}/2$	0	$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$
		f_{am}	0	0	0	$f_{am}/2$	$f_{am}/2$	0	0	$f_{am}/2$	0	$f_{am}/2$	0	$(1-r)\ f_{am}/2$	$r\ f_{am}/2$	$(1-r)\ f_{am}/2$	$r\ f_{am}/2$
	aM	$f_{AM}/2$	$f_{AM}/2$	0	0	0	f_{AM}	0	0	$r\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$	0	$f_{AM}/2$	0	$f_{AM}/2$
		$f_{aM}/2$	$f_{aM}/2$	0	0	0	f_{aM}	0	0	$r\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$	0	$f_{aM}/2$	0	$f_{aM}/2$
		$f_{Am}/2$	$f_{Am}/2$	0	0	0	f_{Am}	0	0	$r\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$	0	$f_{Am}/2$	0	$f_{Am}/2$
		$f_{am}/2$	$f_{am}/2$	0	0	0	f_{am}	0	0	$r\ f_{am}/2$	$(1-r)\ f_{am}/2$	$(1-r)\ f_{am}/2$	$r\ f_{am}/2$	0	$f_{am}/2$	0	$f_{am}/2$
	Am	$f_{AM}/2$	0	$f_{AM}/2$	0	$r\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$	0	0	f_{AM}	0	0	0	$f_{AM}/2$	$f_{AM}/2$
		$f_{aM}/2$	0	$f_{aM}/2$	0	$r\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$	0	0	f_{aM}	0	0	0	$f_{aM}/2$	$f_{aM}/2$
		$f_{Am}/2$	0	$f_{Am}/2$	0	$r\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$	0	0	f_{Am}	0	0	0	$f_{Am}/2$	$f_{Am}/2$
		$f_{am}/2$	0	$f_{am}/2$	0	$r\ f_{am}/2$	$(1-r)\ f_{am}/2$	$(1-r)\ f_{am}/2$	$r\ f_{am}/2$	0	0	f_{am}	0	0	0	$f_{am}/2$	$f_{am}/2$
	am	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$	$(1-r)\ f_{AM}/2$	$r\ f_{AM}/2$	0	$f_{AM}/2$	0	$f_{AM}/2$	0	0	$f_{AM}/2$	$f_{AM}/2$	0	0	0	f_{AM}
		$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$	$(1-r)\ f_{aM}/2$	$r\ f_{aM}/2$	0	$f_{aM}/2$	0	$f_{aM}/2$	0	0	$f_{aM}/2$	$f_{aM}/2$	0	0	0	f_{aM}
		$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$	$(1-r)\ f_{Am}/2$	$r\ f_{Am}/2$	0	$f_{Am}/2$	0	$f_{Am}/2$	0	0	$f_{Am}/2$	$f_{Am}/2$	0	0	0	f_{Am}
		$(1-r)\ f_{am}/2$	$r\ f_{am}/2$	$(1-r)\ f_{am}/2$	$r\ f_{am}/2$	0	$f_{am}/2$	0	$f_{am}/2$	0	0	$f_{am}/2$	$f_{am}/2$	0	0	0	f_{am}

Perisperme

La valeur sélective des embryons dépend du génotype du sporophyte maternel (la double bordure délimite l'unité qui est répétée à l'échelle du tableau)

Gamete reçu paternellement par le sporophyte n-1	Gamete reçu maternellement par le sporophyte n-1															
	AM				aM				Am				am			
AM	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
aM	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
Am	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1	1	1	1	1-*hs*	1-*hs*	1-*hs*	1-*hs*
am	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*
	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*	1-*hs*	1-*hs*	1-*hs*	1-*hs*	1-*s*	1-*s*	1-*s*	1-*s*

Tissu polysporique

En ce qui concerne le tissu polysporique, c'est un peu plus complexe, car la valeur sélective dépend d'un génotype triploïde auquel contribue le génotype paternel

Lorsque le sporophyte maternel est *AA* , l'embryon peut être *AAA* ou *Aaa* selon le génotype paternel. Lorsque le sporophyte maternel est *Aa,* il peut être *AAa* ou *Aaa,* etc.

		Gamete reçu maternellement par le sporophyte n-1															
		AM				aM				Am				am			
Gamete reçu paternellement par le sporophyte n-1	AM	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
	aM	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$
	Am	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	1	1	1	1	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
	am	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$
		$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_2 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$
		$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-h_1 s$	$1-s$	$1-s$	$1-s$	$1-s$

Liste des figures

1.1 Cycles de vie des plantes à graine et phylogénie des végétaux vasculaires . . . 3
1.2 Gamétophytes chez les conifères . . . 3
1.3 La fleur et les gamétophytes angiospermes . . . 4
1.4 Les types de gamétophytes angiospermes . . . 5
1.5 La méiose . . . 13
1.6 Développement du gamétophyte femelle chez les plantes à graine . . . 15
1.7 Hypothèses concernant l'évolution de la génétique de l'albumen . . . 16
1.8 Phylogénie de l'évolution des contributions parentales . . . 17
1.9 Phylogénie de la polysporie . . . 18
1.10 Phylogénie des E(z)-like genes . . . 19
1.11 Evolution du ratio Embryon : Graine chez les angiospermes . . . 21
1.12 Evolution du ratio Embryon : Graine détaillé pour les eudicots . . . 22
1.13 Evolution du ratio Embryon : Graine chez les gymnospermes . . . 23
1.14 Scénarios possibles de l'évolution des graines angiospermes . . . 24
1.15 Modèle de Smith et Fretwell . . . 27
1.16 La règle de Hamilton. . . . 29
2.1 Évolution de la double fécondation et des contributions maternelles . . . 40
2.2 Évolution couplée de DF et MCD . . . 41
2.3 Simulations de l'évolution de la double fécondation : modèle gamétophytique 44
2.4 Simulations de l'évolution de la double fécondation : modèle sporophytique . 45
2.5 Simulations du modèle Perisp1 . . . 49
2.6 Simulations du modèle Perisp2 . . . 50
2.7 Simulations du modèle Perisp3 . . . 50
2.8 Simulations du modèle Polysp1 . . . 51
2.9 Simulations du modèle Polysp2 . . . 51
2.10 Simulations du modèle Polysp2 . . . 52
2.11 Simulations du modèle PP . . . 52
3.1 Systèmes de reproduction . . . 61
3.2 Dépression de consanguinité . . . 69
3.3 Effets mâles sur la taille des graines . . . 72
4.1 Cartographie d'un épi de maïs . . . 76
4.2 Effets de l'épi et de la position sur l'épi sur la masse des grains . . . 77
4.3 Variogramems de la masse des graines sur un épi . . . 78
4.4 Effets des génotypes mâles sur la masse des grains . . . 79
5.1 Croisements altérant le ratio des gènes dans l'albumen. . . . 90

6.1 Effets non additifs dans l'albumen triploïde 110

Liste des tables

1.1 Les gènes à empreinte parentale dans l'albumen 10
1.2 Génétique des différents tissus nourriciers 12
2.1 Modèles de l'évolution des contributions parentales 36
2.2 Table des valeurs sélectives 37
3.1 Base de donnée 64
3.2 Base de donnée 65
3.3 Valeur sélective pour chaque génotype, en fonction de l'effet des allèles délétères (s), et de la dominance de l'allèle délétère a dans ces génotypes. 67
3.4 Effets mâles sur la taille des graines 71
4.1 Matériel expérimental 74
4.2 Plan expérimental 75
4.3 Estimations du modèle d'analyse des cartographies 80
5.1 Contrôle de la mise en place des empreintes 88
5.2 Éxpressions génétiques et hypothèse de différentiation cellulaire 99

Liste des encadrés

1.1 La double fécondation chez les gymnospermes 7
1.2 Les empreintes génétiques dans le monde vivant 9
3.1 Le développement de l'albumen 63
5.1 Régulateurs de l'empreinte parentale chez *A. thaliana* 87
5.2 Altérations du ratio des contributions parentales et empreintes 90
5.3 L'autre fonction des FIS 91
5.4 Mise en place de l'apomixie chez les angiospermes 95

Évolution de la composition génétique du tissu nourricier de la graine

Chez les plantes à graine, l'albumen est un tissu nourricier surprenant, puisqu'il résulte de la double fécondation, qui est la fécondation concomitante de l'oosphère d'une part, et de la cellule mère de l'albumen, la cellule centrale, d'autre part. Dans cette thèse, nous étudions les pressions de sélection qui déterminent l'évolution de l'albumen et pourraient expliquer l'évolution (1) de la double fécondation, (2) d'un doublement des contributions maternelles dans la cellule centrale, (3) de la polysporie, qui consiste en la participation de plusieurs produits de méiose à la formation du gamétophyte, et (4) de l'empreinte parentale, l'expression différentielle des allèles maternels et paternels. Ces innovations modifient l'hétérozygotie dans le tissu nourricier et par conséquent, ont le potentiel de changer l'hétérosis de la graine. Dans cette thèse, nous commençons par étudier comment les changements génétiques qui découlent de la double fécondation, du doublement des contributions maternelles, de la polysporie et de l'empreinte parentale modifient l'hétérosis, ce qui peut jouer en faveur ou en défaveur de leurs évolutions. Puis, nous faisons une revue des données disponibles dans la littérature pour tester si ces traits sont le résultat d'un conflit mâle-femelle sur l'allocation des ressources. Enfin, nous étudions de manière expérimentale les patrons de l'allocation des ressources chez le maïs, pour tester si les embryons sont en compétition pour les ressources, ce qui est une des conditions nécessaires pour qu'un conflit sur l'allocation des ressources ait lieu. Nos modèles théoriques nous permettent de décrire un conflit mâle-femelle sur l'exposition des allèles délétères dans les tissus pour lesquels l'expression des gènes est asymétrique. Ce conflit n'avait jamais été décrit auparavant, et ouvre de nouvelles perspectives pour la compréhension de l'évolution de l'expression génétique. L'analyse des données indique que les théories alternatives à la théorie du conflit sur l'allocation des ressources ont parfois un bon pouvoir explicatif, et méritent par conséquent d'être d'avantage explorées. Enfin, notre étude expérimentale sur le maïs montre que la compétition entre embryons est prédominante lors de l'allocation des ressources chez cette espèce, ce qui est concordant avec les prédictions de la théorie du conflit sur l'allocation.

Mots-clés : Double fécondation | Empreinte parentale |Ressources | Mutations délétères | Conflits | Allocation

---

Evolution of the genetic make-up of seed nutritives tissues

In seed plants, the endosperm is a surprising nutritive tissue, because it results from double fertilization, an eccentricity which results from the parallel fertilization of the egg cell on the one hand, and of the mother cell of the endosperm, the central cell, on the other hand. In this thesis, we study the selective pressures which drive the evolution of the endosperm and may explain the evolution of (1) double fertilization, (2) a doubling of maternal contributions in the central cell, (3) polyspory, the participation of several meiotic products to the gametophyte and (4) imprinting, the differential expression of maternal and paternal alleles. These innovations modify heterozygosity in the endosperm and as a consequence, have the potential to change heterosis in the seed. In this thesis, we first investigate how genetic changes that result from double fertilization, doubling of maternal contribution, polyspory and imprinting modify heterosis, which may play in favour or against the evolution of these traits. Second, we review the available data to test whether these traits are the result of a male-female conflict over resource allocation. Finally, we study experimentally patterns of resource allocation in maize to assess whether embryos compete for resources, which is a necessary condition for the conflict over resource allocation to occur. Our theoretical models allow us to describe a male-female conflict over the exposition of deleterious alleles in tissues with asymmetrical gene expression. This conflict had never been described before and opens perspectives for understanding the evolution of gene expression. We conclude from our analysis of data that theories which are alternative to the conflict theory over resource allocation may have a better explanatory power and therefore deserve to be further explored. Finally, our experimental study in maize shows that competition between embryos drives resource allocation in this species, which is consistent with predictions of the conflict over resource allocation theory.

Keywords: Double fertilization | Imprinting | Resources | Deleterious mutations | Conflicts | Allocation

---

Discipline : Biologie des populations et écologie

Centre d'Écologie Fonctionnelle et Évolutive (UMR 5175), Montpellier