Variabilité naturelle des mucilages des graines de Brassicaceae : étude comparative inter et intra espèces

Sébastien Viudes

Université Fédérale
Toulouse Midi-Pyrénées

THÈSE

En vue de l'obtention du
DOCTORAT DE L'UNIVERSITÉ DE TOULOUSE

Délivré par l'Université Toulouse 3 - Paul Sabatier

Présentée et soutenue par

Sébastien VIUDES

Le 3 novembre 2021

Variabilité naturelle des mucilages des graines de Brassicaceae : étude comparative inter et intra espèces

---

Ecole doctorale : **SEVAB - Sciences Ecologiques, Vétérinaires, Agronomiques et Bioingenieries**

Spécialité : **Développement des plantes**

Unité de recherche :
LRSV - Laboratoire de Recherche en Sciences Végétales

Thèse dirigée par
Christophe DUNAND et Vincent BURLAT

Jury

Mme Gwyneth Ingram, Rapporteure
Mme Annie Marion Poll, Rapporteure
Mme Maïté Vicré-Gibouin, Examinatrice
Mme Fabienne Vailleau, Examinatrice
M. Christophe DUNAND, Directeur de thèse
M. Vincent Burlat, Co-directeur de thèse

Résumé

La myxospermie et la myxocarpie (regroupées sous le terme myxodiasporie) désignent respectivement la capacité d'une graine ou d'un fruit à extruder du mucilage polysaccharidique lors de l'imbibition. Les espèces myxodiasporiques sont largement répandues parmi les plantes à fleurs, mais sont souvent étroitement apparentées aux espèces non myxodiasporiques. Cette répartition rend mystérieuse l'histoire évolutive de la myxodiasporie : Est-ce un trait ancestral qui s'est perdu plusieurs fois chez les espèces non myxodiasporiques? Ou est-ce une fonctionnalité qui est apparue plusieurs fois indépendamment chez les espèces myxodiasporiques? De plus, les rôles écologiques de la myxodiasporie sont très divers et semblent dépendre de l'espèce. La génétique contrôlant cette caractéristique est encore inégalement connue, en dehors de l'espèce modèle *Arabidopsis thaliana*. En effet, cette plante modèle est une Brassicaceae myxospermique qui a été bien caractérisée pour la morphologie, le développement, la composition et la génétique de son mucilage de graine et de ses cellules sécrétrices de mucilage (CSM). Cette thèse met en lumière l'évolution, la génétique et l'écologie de la myxospermie dans la famille des Brassicaceae au niveau inter- et intra-espèce.

Au niveau inter-espèces, les graines et leurs cellules épidermiques d'une trentaine d'espèces de Brassicaceae ont été comparées à celles d'*A. thaliana* pour étudier l'évolution de la myxospermie dans cette famille. Chaque espèce de Brassicaceae a été soigneusement phénotypée à l'échelle macroscopique (graine entière) et microscopique (CSM) avant et après imbibition. De plus, les orthologues de chaque gène d'*A. thaliana* connus pour être impliqués dans la myxospermie ont été recherchés dans les génomes disponibles des espèces étudiées. La disponibilité de données transcriptomiques pour quelques espèces ont permis d'approfondir l'étude de ces gènes orthologues. Dans la famille des Brassicaceae, une grande diversité de structures de mucilages a été observée à des niveaux macro et même microscopiques. D'après les données morphologiques, il semble y avoir une origine ancestrale du mucilage chez les Brassicaceae, mais elle est très différente de celle d'*A. thaliana*. Malgré les morphologies contrastées observées, les acteurs génétiques candidats n'étaient pas corrélés avec le phénotype. Une grande majorité des gènes orthologues de ceux impliqués dans la myxospermie d'*A. thaliana* étaient présents et partiellement exprimés même pour les espèces non myxospermiques.

Au niveau intra-espèce, 166 populations naturelles d'*A. thaliana* ont été analysées pour leur taille de mucilage adhérent et non adhérent. Ces populations sont originaires d'une région du sud de la France et possèdent des données génétiques (fréquences des polymorphismes mono-nucléotidiques (SNP)) et environnementales associées. L'association de leur phénotype de mucilage aux fréquences des SNP a permis de mettre en évidence de nouveaux gènes candidats prometteurs impliqués dans les deux couches du mucilage d'*A. thaliana*. De plus, des corrélations intéressantes avec les données environnementales orientent vers de nouveaux rôles écologiques putatifs à découvrir pour le mucilage d'*A. thaliana*.

Mots clés : *Arabidopsis thaliana*, Brassicaceae, mucilage des graines, myxospermie, cellules sécrétrices de mucilage (CSM), évolution, gènes orthologues, étude d'association a l'échelle du génome (GWAS), rôle écologique.

Abstract

The myxospermy and myxocarpy (grouped under the term myxodiaspory) designates the ability of a seed or a fruit to extrude mucilage upon imbibition, respectively. Myxodiasporous species are widely spread among flowering plants but are often closely related to non-myxodiasporous species. This repartition makes mysterious the evolutionary scenario of myxosdiaspory: Is it an ancestral trait which was lost several times in non-myxodiasporous species? Or is it a feature that appeared several times independently in myxodiasporous species? Additionally, the ecological roles of myxodiaspory are very diverse and seem to be species-dependent. Also, the genetics controlling this feature is still inequally known except for the model plant *Arabidopsis thaliana.* Indeed, this myxospermous Brassicaceae species has been well characterized for the morphology, development, composition, and genetics of its seed mucilage and mucilage secretory cells (MSCs). This thesis sheds light on the evolution, genetics, and ecology of myxospermy in the Brassicaceae family at both the inter- and intra- species level.

At the inter-species level, the seeds and their epidermal cells from about thirty Brassicaceae species were compared to those of *A. thaliana* to investigate the myxospermy evolution in this family. Every Brassicaceae species was thoroughly phenotyped at the macroscopic (whole seed) and microscopic (MSCs) scale before and after imbibition. Additionally, ortholog from every *A. thaliana* gene known to be implicated in myxospermy was tracked in the available genomes of the studied species. The availability of transcriptomic data for few species allowed further investigation of these orthologous genes. Among the Brassicaceae family, a wide diversity of mucilage structures was observed at the macro- and microscopic level. Based on morphological data, the origin of myxospermy in Brassicaceae appears ancestral, but the trait is very different from that of *A. thaliana*. Despite the observed contrasted morphologies, the putative genetic actors were not correlated with the phenotype. A large majority of orthologous genes to those related to myxospermy in *A. thaliana* were present and partially expressed even in non-myxospermous species.

At the intra-species level, 166 natural populations of *A. thaliana* were analysed for their adherent and non-adherent mucilage size. These populations come from a South France region and have available associated genetic (single nucleotide polymorphisms (SNPs) frequencies) and environmental data. The association of their mucilage phenotype with the SNPs frequencies highlights promising new candidate genes putatively implicated in both mucilage layers of *A. thaliana*. In addition, interesting correlations with environmental data point to new putative ecological roles to be discovered for the mucilage of *A. thaliana*.

Key words: *Arabidopsis thaliana*, Brassicaceae, seed mucilage, myxospermy, mucilage secretory cells (MSCs), evolution, orthologous genes genome wide association study (GWAS), ecological role.

Remerciements

Le problème des remerciements réside dans l'impossibilité d'être exhaustif. Je vais néanmoins essayer de remercier nommément un maximum des personnes qui ont participé à rendre cette thèse possible.

Tout d'abord merci à mes deux encadrants. Merci à Christophe pour m'avoir fait confiance dès la première année de Master en me prenant en stage sans aucune preuve préalable de mes compétences, puis d'avoir conservé cette confiance tout au long de la thèse. La liberté que tu m'as laissée, a été motrice de mon évolution. Merci à Vincent pour ta pédagogie, diplomatie, et pertinence scientifique mais aussi pour ton investissement sur tous les fronts. Tu as su canaliser ma curiosité débordante pour me permettre de concrétiser mes recherches.

Merci aux membres de mon comité de thèse pour leurs conseils avisés : Helen North, Pierre-Marc Delaux, Jean Keller (merci aussi à toi pour les nombreuses fois ou tu m'as patiemment conseillé et pour m'avoir prêté tes scripts).

Merci aux membres de mon jury de thèse, Gwyneth Ingram, Annie Marion-Poll, Maité Vicré et Fabienne Vailleau pour avoir accepté d'examiner mon manuscrit.

Merci aux membres de la plateforme d'imagerie TRI-genotoul pour leur sympathie et leur disponibilité.

Merci à mes deux stagiaires Thomas et Werner pour leur investissement et l'aide apportée.

Merci à toute l'équipe pédagogique du LRSV avec qui j'ai pu interagir lors de mes enseignements, si l'expérience a été aussi bonne c'est avant tout grâce à vous.

Merci ensuite à tous les thésards et post-doctorants avec qui j'ai pu, même un tant soit peu, passer du temps, les échanges scientifiques ou nos expériences respectives ont été très enrichissantes. Un merci tout particulier à Maxime et Julie pour tout, Bastien pour ton éternelle bonne humeur (depuis 2015 quand même !), Ali pour m'avoir apporté la preuve vivante qu'on peut être très similaire en ayant grandi dans deux pays plus que contrastés, et puis Harold, Ayla, Quentin, Margot, mes anciens collègues représentants SEVAB, et à tous les autres qui se reconnaitront.

Merci à ma famille de ne pas (trop ?) m'en vouloir d'être allé les voir bien trop peu souvent.

Et merci à celle qui a supporté au quotidien mes excès de fatigue et mes surplus d'encombrement cérébral.

Table des matières

1. Introduction générale

L'évolution est un processus itératif qui ne conserve que les individus les plus adaptés à leur environnement. La science est un processus itératif qui conserve les hypothèses les plus étayées pour s'approcher progressivement de la vérité. J'aime à penser que la méthode scientifique est inspirée du mécanisme biologique qui a créé l'humanité, ainsi que toutes les autres formes de vie qui, entre autres choses, lui permettent de subsister.

1.1. L'étude de l'évolution

1.1.1. L'évolution, un mécanisme complexe à la découverte précoce

L'évolution est le nom attribué au mécanisme responsable de la diversité et de la complexité du monde vivant que l'on peut apprécier chaque jour sur la Terre. La vie est tellement vaste et merveilleuse qu'elle a longtemps été dépeinte comme étant d'origine divine. Pourtant, depuis le XIX^{ème} siècle nous savons que l'ensemble de cette diversité est le résultat d'une diversification progressive des êtres vivants à partir d'un ancêtre commun unique. Cette diversification repose sur le mécanisme de sélection naturelle qui favorise les individus les plus adaptés à leur environnement génération après génération. C'est avec l'application de plus en plus rigoureuse de la méthode scientifique que des similarités morphologiques ont pu être observées, reportées, et mises en relation afin de commencer à réfléchir aux mécanismes sous-jacents à l'évolution. En effet, la morphologie étant le paramètre le plus facilement accessible, c'est grâce aux observations macroscopiques que ce mécanisme a pu être compris et publié pour la première fois en 1859 par Charles Darwin dans *L'origine des espèces* (Darwin, 1859). Résumé simplement, si deux organismes partagent le même trait morphologique, ils l'ont donc probablement hérité de leur ancêtre commun. En revanche, s'ils ont une spécificité, elle doit répondre à une contrainte présente dans leur environnement. Il est toujours étonnant de constater a posteriori à quel point les Hommes sont doués pour comprendre et décrire un mécanisme dans son ensemble, quand bien même il manque une information clef pour confirmer l'hypothèse. L'information manquante dans le cas de Darwin était l'existence et le rôle du code génétique, puisque à son époque les recherches sur l'hérédité étaient encore balbutiantes (bien que Mendel y travaillait déjà dans l'ombre) et que l'existence des chromosomes et leur lien avec l'hérédité n'ont eu de réel succès qu'au début du XX^{ème} siècle. Il est d'ailleurs fascinant de mettre en perspective *L'origine des espèces* de Darwin avec *Le gène égoïste* de Dawkins (Dawkins 1976) pour prendre conscience à quel point la connaissance de la biologie moléculaire peut changer notre perception d'un même mécanisme. C'est bien le même mécanisme qui est décrit dans les deux ouvrages, mais l'unité de base sur lequel ce mécanisme s'exerce change. Dawkins propose de penser le gène comme unité de l'évolution, ce qui a pour conséquence de considérer les individus (unité de l'évolution de Darwin) comme de « simples » machines à survie au service des gènes.

1.1.2. La progression de la compréhension du matériau de base de l'évolution

Pour étudier l'évolution, la morphologie se révèle souvent peu digne de confiance, car elle ne reflète pas toujours clairement les relations de parenté entre les espèces. C'est particulièrement vrai en cas d'apparitions indépendantes de morphologies qui se ressemblent car elles ont convergé vers la même fonction en réponse à des contraintes similaires. Ces phénomènes de convergence évolutive sont induits par l'exposition longue à une même contrainte environnementale forte qui va provoquer une adaptation similaire chez des espèces pourtant bien différentes à l'origine. Un exemple frappant dans le monde végétal est l'adaptation aux milieux chauds et arides de certaines espèces de Cactaceae sur le continent Américain et d'Euphorbiaceae et Aizoaceae (entre autre) sur le continent Africain que l'on confond bien souvent du fait de leur ressemblance morphologique (Alvarado-Cárdenas et al., 2013) (**Figure 1**). Or, si les adaptations convergentes ont des origines différentes elles auront aussi des bases génétiques différentes. L'avènement de la génétique a donc permis de déjouer ces pièges pour un établissement plus exact de la phylogénie du vivant. Les arbres de parenté ainsi basés sur la similarité des séquences génomiques (arbres phylogénétiques) ont donc permis de se rapprocher de plus en plus des réelles histoires évolutives. Deux modalités sont alors possibles pour la construction d'arbre phylogénétique. La première utilisera un mélange de séquences génétiques présentes chez toutes les espèces à étudier. Les *internal transcribed spacer*s (ITS) ont longtemps été utilisés pour leur conservation (en présence et en séquence) dans le vivant permettant de se faire une assez bonne idée de l'apparentement entre les espèces sans avoir besoin de séquencer une grande portion du génome (Baldwin et al., 1995). La deuxième possibilité consiste à reconstruire la phylogénie d'une famille de gènes provenant d'une ou plusieurs espèces. Cette approche permet d'appréhender l'évolution de la famille de ce gène au cours du temps, et éventuellement à travers les espèces mais ne suivra pas forcement la phylogénie des espèces à cause de divergence propre à ce gène. Dans le premier cas ce sera donc des noms d'espèces qui seront en feuilles de l'arbre alors que dans le deuxième ce seront des noms de gènes. Cette dualité fait écho avec les deux natures d'unité de l'évolution évoquées précédemment.

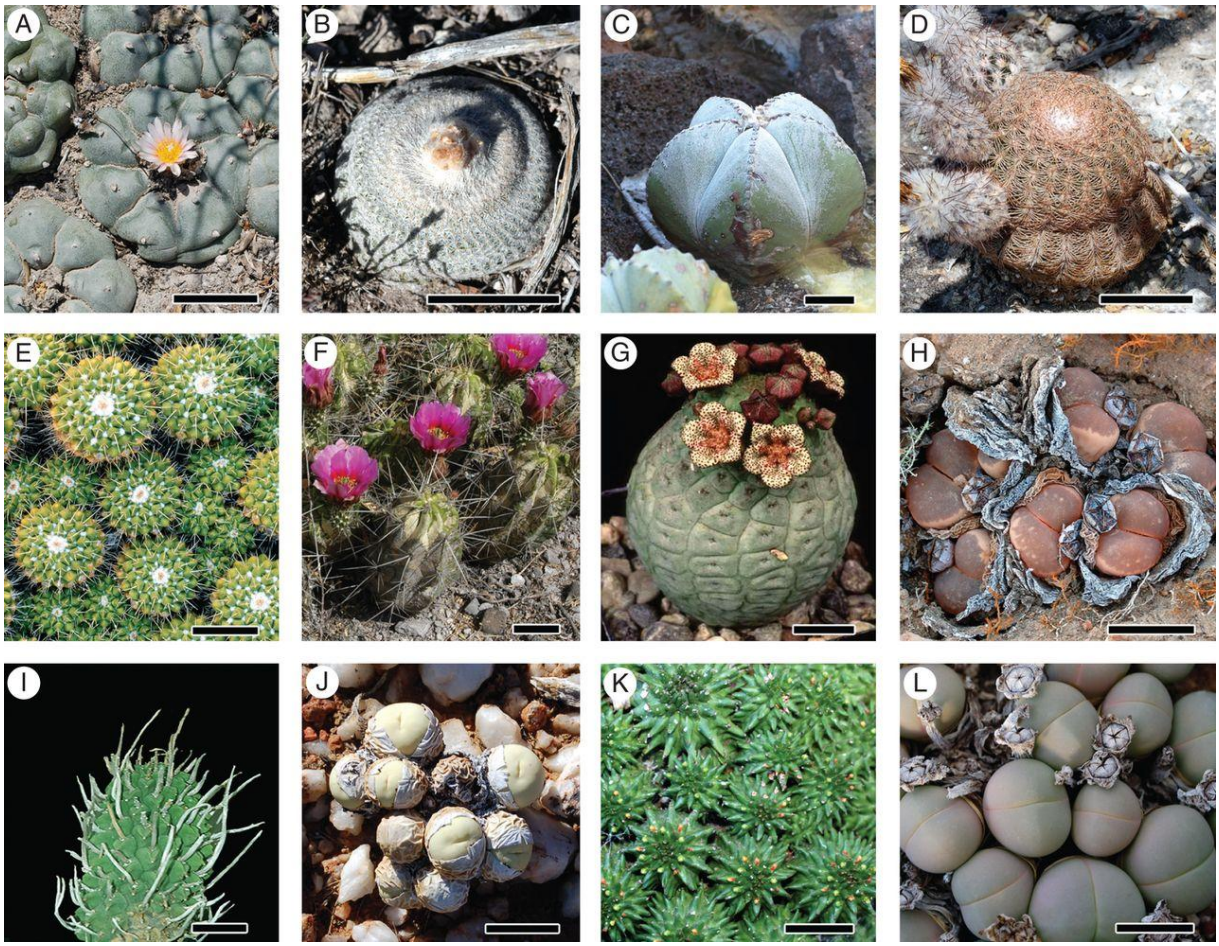


Figure 1 : Exemples de convergence évolutive vers la forme globulaire et la succulence dans des espèces appartenant à des familles éloignées géographiquement et phylogénétiquement (extrait de Alvarado-Cárdenas et al., 2013). (A–F) Plantes mexicaines (Cactaceae) : (A) *Lophophora williamsii*; (B) *Epithelantha micromeris*; (C) *Astrophytum myriostigma*; (D) *Echinocereus pectinathus*; (E) *Mammillaria compressa*; et (F) *Echinocactus horizonthalonius*. (G–L) Plantes d'Afrique du sud : (G) *Larryleachia cactiformis* (Apocynaceae); (H) *Lithops herrei* (Aizoaceae); (I) *Euphorbia fasciculata* (Euphorbiaceae); (J) *Conophytum calculus* (Aizoaceae); (K) *Euphorbia susannae* (Euphorbiaceae); et (L) *Gibbaeum heathii* (Aizoaceae). Barre d'echelle = 2 cm. Photographes: (B, F) C. Gómez, (G) J. Trager.

1.1.3. Une disponibilité de données grandissante, en types comme en nombre

L'innovation technologique constante du XXI^{ème} siècle et la diminution des coûts de séquençage ont permis de faciliter progressivement l'obtention des données génomiques. En conséquence, de plus en plus d'espèces ont des données génomiques disponibles et elles deviennent de plus en plus exploitables grâce au progrès des séquenceurs mais aussi du traitement informatique de cette colossale masse d'informations. Concernant l'étude de l'évolution, le traitement de ces données par des logiciels appropriés, peut permettre de construire des phylogénies d'espèces plus précises en utilisant un grand nombre de gènes simultanément pour compenser leurs divergences évolutives intrinsèques (Nikolov et al., 2019). Ces banques de données permettent aussi de mettre en regard les données génétiques disponibles avec des données phénotypiques déjà publiées pour retracer l'évolution d'un réseau de

gènes et du trait qu'il contrôle à très grande échelle phylogénétique (Radhakrishnan et al., 2020). Ce genre d'approche tire parti des données d'expression des gènes (transcriptomique) et/ou de leur traduction (protéomique) car elles permettent de s'approcher de la réalité spatiotemporelle des traits que l'on souhaite étudier. Dans un même organisme l'expression sera dépendante de l'organe, des tissus ou même du type cellulaire étudié, et sera de plus variable en fonction des conditions environnementales (Zhang et al., 2017). De plus, une simple modification post traductionnelle d'une protéine peut complètement modifier sa fonction biologique (Li et al., 2018). Le croisement de tous ces niveaux d'informations rend donc les analyses plus complexes mais plus puissantes.

Le principal défaut de cette ressource génomique vient de sa nature même. Comme l'information est colossale, elle est traitée informatiquement afin d'automatiser le plus de tâches possibles. Des erreurs se glissent lors de ces processus informatiques et, malgré les multiples innovations dans la correction automatiques d'erreurs, certaines peuvent persister (Laehnemann et al., 2016). De plus, l'automatisation est dépendante de nos connaissances actuelles sur la structure et le fonctionnement génomique. Pour prendre un exemple récent, des ARN considérés jusqu'à présent comme non codant se sont révélés être codants pour des peptides aux effets biologiques importants (Ren et al., 2021). L'annotation et les prédictions de la traduction de ces ARN devraient donc être mises à jour pour tous les organismes déjà séquencés. En plus d'être colossale, l'information est asymétriquement accessible à cause de limitations techniques. Par exemple, certaines zones des chromosomes sont beaucoup plus difficiles à séquencer que d'autres comme l'attestent les efforts déployés pour obtenir les quelques pourcents encore inconnus ou erronés du génome humain, 20 ans après la publication de sa première version (Nurk et al., 2021).

1.1.4. Phylogénie simplifiée du vivant et phylogénie ciblée des plantes

La vie sur Terre peut être divisée en trois groupes contenant chacun une diversité inégale (**Figure 2**). Le groupe le plus diversifié, et ce de très loin, est constitué des bactéries. Les bactéries sont des organismes unicellulaires sans noyau à l'instar des Archeae, même si ils sont extrêmement différents génétiquement. Les Eucaryotes, organismes composé de cellules avec un système endomembranaire dont le noyau, est le groupe qui va nous intéresser ici car il contient, entres autres, les plantes. Il est fascinant de constater que la majorité de la vie que nous avons identifié jusqu'à présent est restée invisible jusqu'au XVII^{ème} siècle (Porter, 1976), et incompréhensible jusqu'au XIX^{ème} siècle où les travaux de Robert Koch et Louis Pasteur ont provoqué l'avènement de la bactériologie.

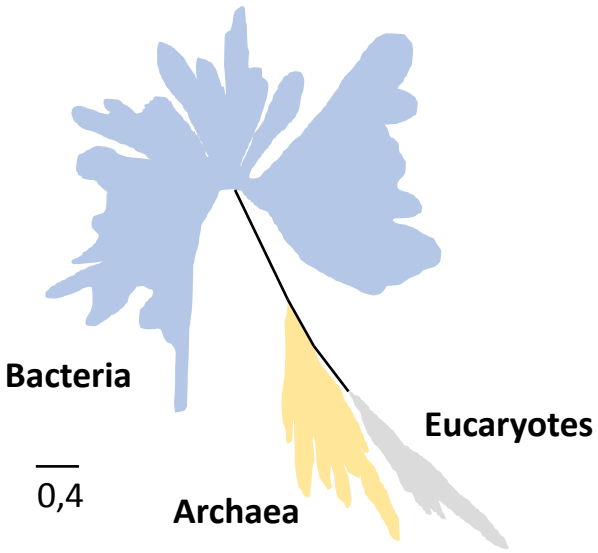


Figure 2 : Phylogénie de la vie simplifiée à l'extrême (adapté de Hug et al., 2016). Les trois grands règnes du vivant y sont représentés groupés, donnant un aperçu de l'abondance relative d'espèces et de la diversité de *phyla* qu'ils contiennent. La taille des formes colorées et de la branche qui les relie montrent la divergence génétique qui les sépare. Une distance de la taille de la barre d'échelle représente 0,4 substitution par sites. L'homogénéité génétique et la faible diversification des Eucaryotes est ici frappante comparé aux bactéries.

Les Eucaryotes, malgré une grande diversité de groupes, sont bien plus homogène génétiquement que les bactéries (**Figure 3**). Bien que les Eucaryotes soient le groupe le plus anciennement étudié, leur phylogénie évolue encore et elle comporte toujours des incertitudes. La découverte de nouveaux sous-groupes impose la réorganisation des relations anciennement établies. Aussi étonnant que ça puisse paraitre, pour des groupes formés depuis longtemps comme les Archeaplastidia, il est encore aujourd'hui difficile de savoir qui des Rhodophyta (algues rouges), des Glaucophyta (algues « glauques »), ou des Chloroplastidia (plantes vertes) ont divergé avant les autres. Un constat identique peut être fait pour les Opisthokonta (contient les animaux et les champignons) et ses deux groupes frères. Il est bon de garder à l'esprit que la phylogénie est à l'image de la science, vraie en fonction des données actuellement disponibles et des méthodes présentement utilisées. Les nouvelles données génomiques continuant d'affluer et la méthodologie évoluant sans cesse (voir partie 1.1.3), la phylogénie va donc encore beaucoup changer.

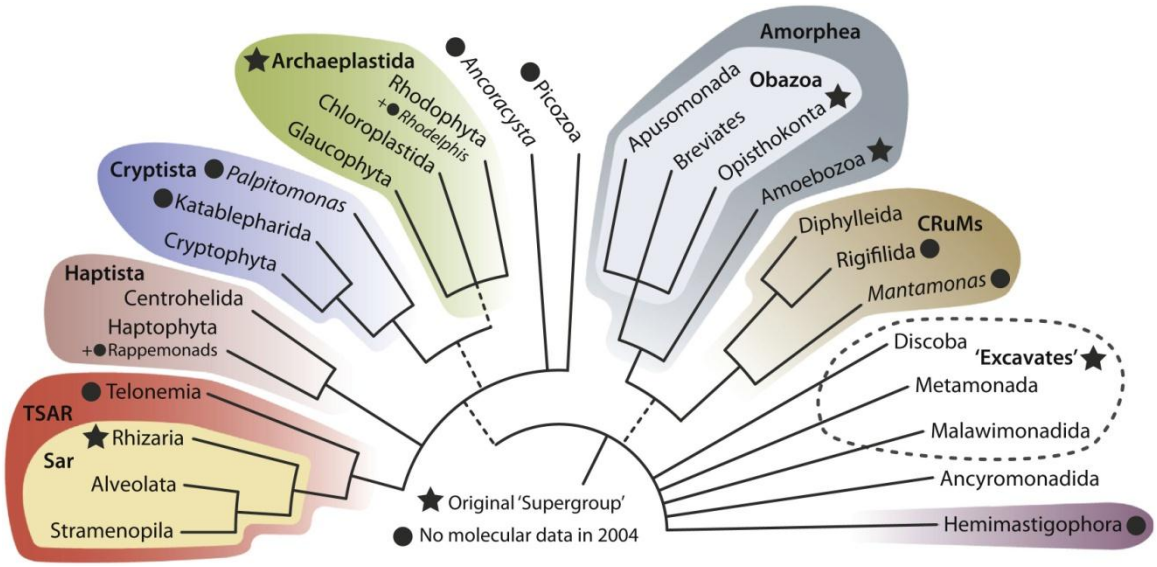


Figure 3 : Cladogramme théorique des relations phylogénétiques des Eucaryotes (extrait de Burki et al., 2020). C'est une compilation du consensus émanant des études phylogénétiques récentes. Les embranchements multiples montrent les nœuds encore non résolus. Les lignes pointillées symbolisent l'incertitude concernant la nature monophylétique de certains groupes. Les ronds noirs identifient les groupes qui n'avaient pas encore de données moléculaires en 2004. Les étoiles noires mettent en avant les « supergroupes » qui constituées les anciens modèles de phylogénies. Tous sauf les Archaeplastidia sont ici remis en question soit par leur dissolution, soit par leur inclusion dans de nouveaux groupes. Les animaux et champignons sont contenu dans les Opistochonta et les plantes dans les Chloroplastidia.

Les Archeplastidieae contiennent donc les algues rouges et « glauques » et les plantes vertes citées précédemment. Les Chlorophytes et les Streptophytes sont deux groupes monophylétiques d'algue verte qui ont divergé avant l'ancêtre commun des plantes terrestres (**Figure 4**). Les Streptophytes ayant divergé plus tardivement que les Chlorophytes, ils sont donc les plus proches des plantes terrestres. Viennent ensuite par ordre de divergences les Embryophytes (plantes terrestres) qui regroupent les bryophytes (mousses, hépatiques, et anthocérothes) et les plantes vasculaires. Les plantes vasculaires contiennent les Lycophytes et Fougères au sens large (fougères et prèles) et les Spermatophytes (plantes à graines). Les Spermatophytes contiennent les Gymnospermes (conifères, Cycas, Ginkos) et les Angiospermes (plantes à fleurs). Ce sont bien évidement ces plantes à graines qui vont être l'objet de cette thèse car elle traite du mucilage des graines. Néanmoins il est toujours intéressant de prendre du recul sur un sujet en observant les parallèles qui peuvent exister avec d'autres objets d'études.

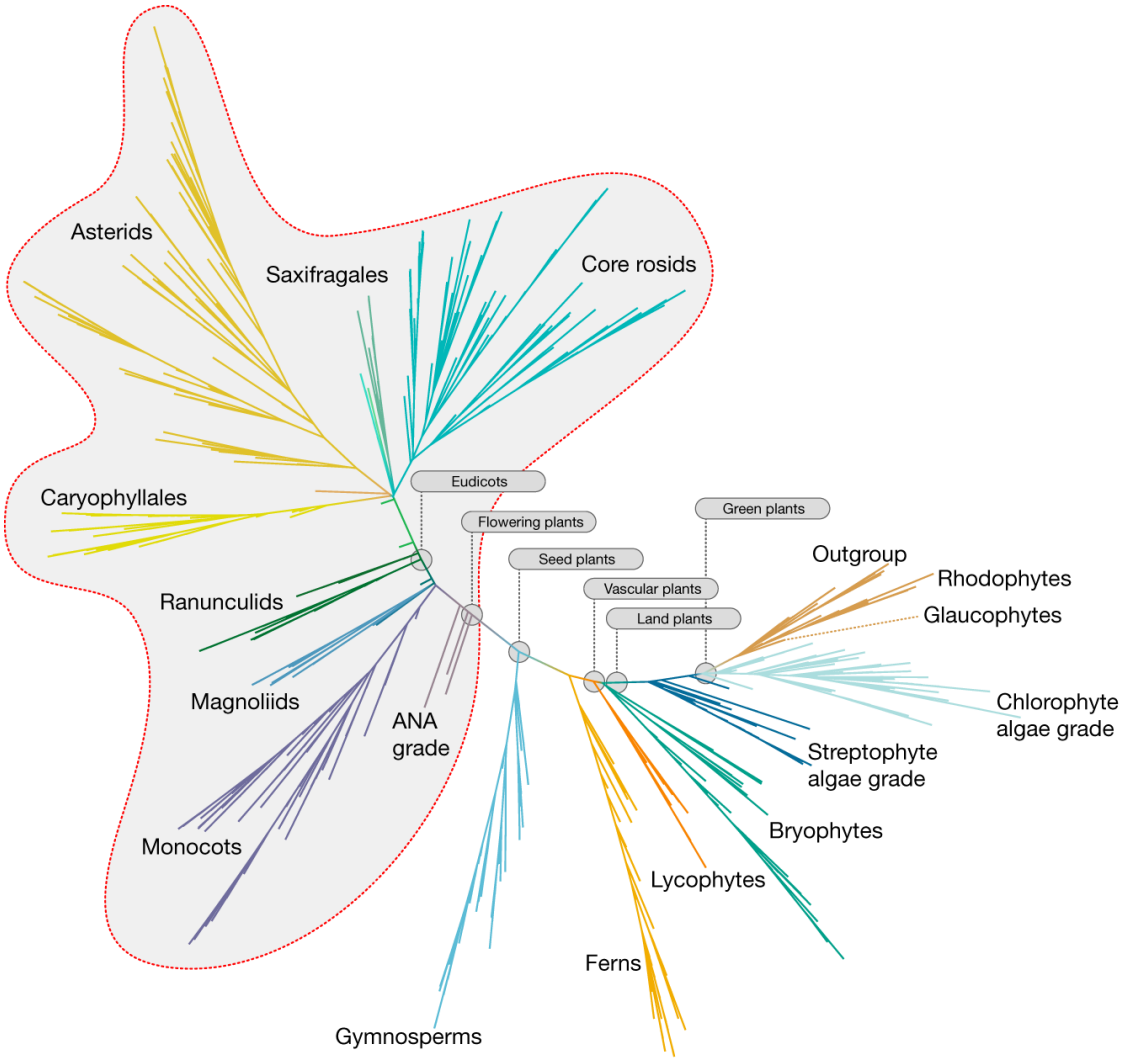


Figure 4 : Phylogénie des Embryophytes extraite de Leebens-Mack et al., 2019. Elle provient de l'analyse de 410 gènes à travers 1153 espèces. A proximité immédiate des outgroups se trouvent les Rhodophytes (algues rouges), Glaucophytes (algues unicellulaires glauques), et Chlorophytes (algues vertes) qui constitue les trois grands groupes monophylétiques qui ont divergé avant l'ancêtre commun des plantes vertes. Les Streptophytes sont aussi regroupées avec les Chlorophytes sous l'appellation commune d'algues verte bien qu'ils soient tout deux des groupes monophylétiques indépendants. Viennent ensuite les plantes terrestres non vasculaires, puis vasculaires, à graine, et à fleurs.

1.2. Les mucilages, une paroi macroscopique multifonction ?

1.2.1. Historique et définition du mucilage

Le mot mucilage proviendrait du français « musillage » ou « muscilage » emprunté au latin *mucilago* (mucosité) et désigne dès le XIV^{ème} siècle les parties visqueuses des plantes ou par extension les préparations pharmaceutiques élaborées avec (CNRTL 2012). C'est donc une définition très large mais qui montre que le terme mucilage est historiquement inféodé au règne végétal. Il a néanmoins aussi été utilisé au XIX^{ème} siècle pour désigner les matières visqueuses d'origine animale plus couramment

appelées mucus (CNRTL 2012). Actuellement, le terme mucilage désigne le plus souvent une matrice apoplastique gélatineuse viscoélastique composée de polysaccharides de haut poids moléculaire qui est secrétée dans l'environnement (Sasse et al., 2018). Néanmoins, il existe des matrices polysaccharidiques gélatineuses qui ne sont pas secrétées mais qui ont été extraites de tissus internes de plante (comme pour l'intérieur des feuilles d'*Aloe vera*) et qui sont quand même appelées mucilages (Minjares-Fuentes et al., 2017). Chez les plantes, le mucilage est une paroi modifiée (Haughn and Western, 2012) et il convient de développer la notion de paroi avant de se focaliser sur les mucilages des plantes.

1.2.2. Les parois du vivant

Les bactéries, champignons, et plantes possèdent des matrices extracellulaires polysacharidiques mais leurs natures et leurs agencements varient d'un règne à l'autre (**Figures 5, 6, et 7**). Chez les bactéries ce sont systématiquement des peptidoglycanes qui surmontent directement la membrane plasmique, ils peuvent être seuls ou associés avec des lipopolysacharides, des arabinogalactanes, ou des membranes additionnelles (**Figure 5)**. Chez les champignons, de la chitine puis des glucanes se situent généralement à proximité immédiate de la membrane plasmique, la partie la plus externe de la paroi varie ensuite fortement d'une espèce à l'autre (**Figure 6)**. Chez les plantes la composition polysaccharidique des parois peut se résumer à cellulose, hémicelluloses, pectines et protéines mais elle sera détaillée dans la partie suivante. Cette différence de composition indique que les parois de ces trois règnes ont des origines évolutives distinctes mais ont cependant convergées pour répondre à des contraintes en partie similaires imposées par le mode de vie sessile commun à ces règnes (Niklas, 2004).

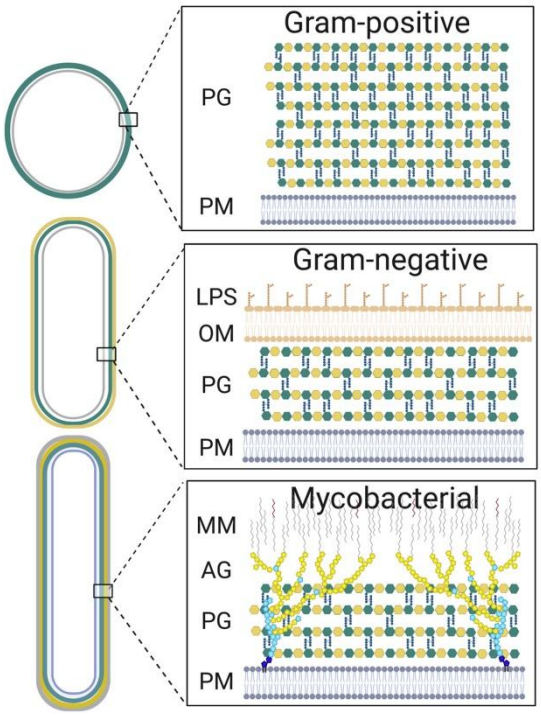


Figure 5 : Structure et composition des parois de bactéries Gram-positives, Gram-négatives, et mycobactériennes (extrait de Brown et al., 2020). La membrane plasmique (PM) est surmontée de peptidoglycanes (PG) et, en fonction du type bactérien, de lipopolysaccharides (LPS), d'une mycomembrane (MM), d'arabinogalactanes (AG), ou d'une membrane externe (OM).

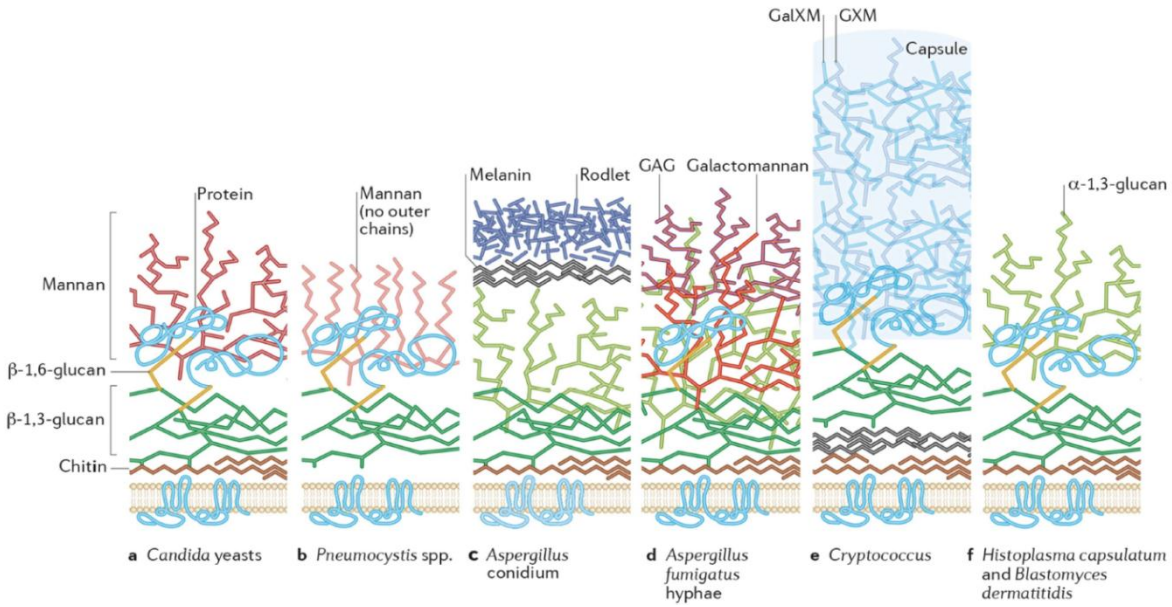


Figure 6 : Exemples de structures et compositions de parois fongiques chez six espèces (extrait de Gow et al., 2017). Les composants les plus fréquents sont la chitine et les β-(1,3) et β-(1,6) glucanes. Les autres composants qui se lient ensuite à eux peuvent énormément varier entre les espèces.

1.2.3. Les parois végétales

Les cellules végétales s'entourent de cellulose, hémicelluloses, et de pectines dans un agencement bien particulier (**Figure 7**). La cellulose constitue le squelette de la matrice extracellulaire grâce à sa structure en microfibrilles. Les celluloses synthases transitent depuis le Golgi pour venir se positionner au travers de la membrane plasmique. Elles forment des complexes en forme de rosettes qui leur permettent de synthétiser plusieurs chaine de β-(1-4)-glucane simultanément qui vont s'associer supramoléculairement entre elles via un réseau de ponts hydrogène pour former les microfibrilles de cellulose cristalline (Lampugnani et al., 2018). L'interaction entre les celluloses synthases et les microtubules vont guider le positionnement orienté des microfibrilles pour former un maillage organisé (Lampugnani et al., 2018). Les hémicelluloses (xyloglucanes, arabinoxylanes, mannanes…) et pectines (homogalacturonanes rhamnogalacturonanes I (RGI) et RGII) vont être synthétisés directement dans le Golgi puis être délivrés dans la paroi par fusion de vésicules Golgiennes à la membrane plasmique (Anderson and Kieber, 2020). Les xyloglucanes vont se lier entre les microfibrilles de cellulose par des ponts hydrogène. Les homogalacturonanes vont par endroits accueillir des chaines latérales complexes formant alors des rhamnogalacturonanes de type II (**Figure 8**). Les pectines comprennent aussi des rhamnogalacturonanes de type I qui sont formées d'un squelette de rhamnose et d'acide galacturonique en alternance et de chaine latérales de galactanes,

d'arabinanes et d'arabinogalactanes (**Figure 8).** Il existe des protéines spécialisées dans la synthèse des polysaccharides comme les glycosyltransferases incluant les celluloses synthases mais aussi des protéines impliquées dans la sécrétion des polysaccharides non-cellulosiques depuis le Golgi jusqu'à l'extérieur de la membrane plasmique (**Figure 7** ; Driouich et al., 2012). La paroi végétale contient aussi diverses protéines qui lui confèrent de la plasticité pour répondre aux contraintes développementales de sa cellule ainsi que des capacités de signalisation et de perception pour permettre l'interaction avec l'environnement (Albenne et al., 2014). Plusieurs familles de protéines sont donc à l'œuvre pour synthétiser et remodeler les parois par la modification et l'assemblage des polysaccharides qu'elles contiennent comme les peroxydases de classe III par exemple (Francoz et al., 2015).

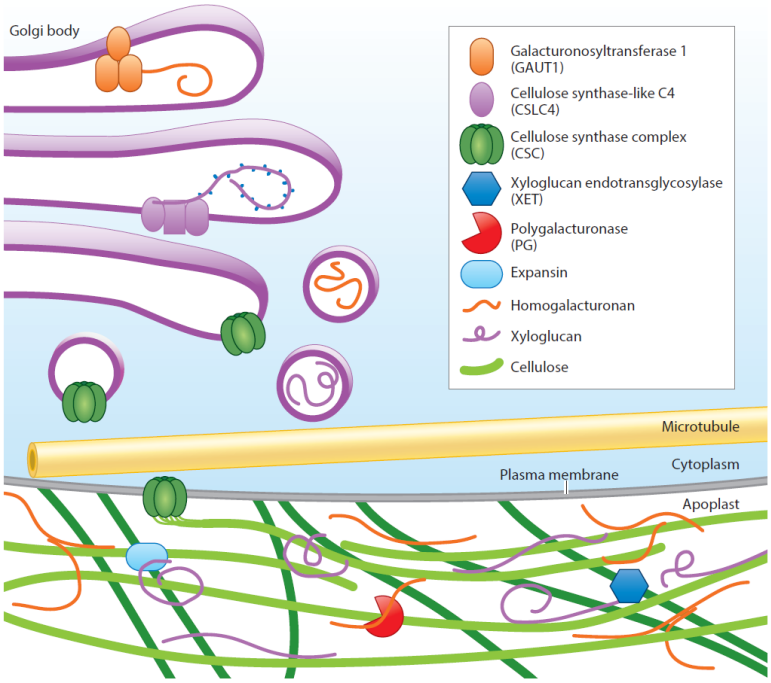


Figure 7 : Schéma simplifié de la périphérie d'une cellule végétale montrant la synthèse et l'agencement des polysaccharides qui composent la paroi (extrait de Anderson and Kieber., 2020). GAUT1 et CSLC4 sont des exemples de glycosyltransferases Golgiennes impliquées respectivement dans la synthèse d'homogalacturonane (pectine) et de xyloglucane (hemicellulose). CSC est le complexe de cellulose synthase organisé en rosette pour synthétiser les microfibrilles de cellulose au niveau de la membrane plasmique. PG, Expansine et XET sont des exemples de protéines pariétales impliquées dans le remodelage de la paroi.

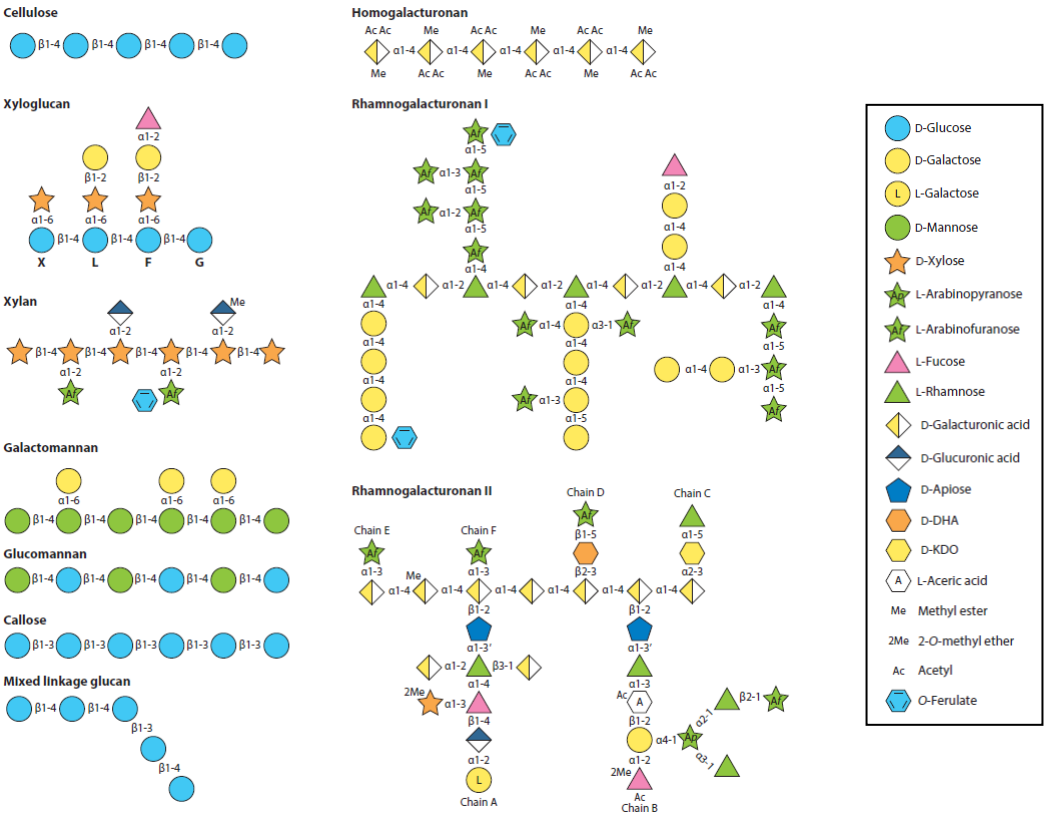


Figure 8 : Exemples schématiques de structures de polysaccharides pariétaux (extrait de Anderson and Kieber., 2020). Leur composition en monosaccharides, leur agencement, et leur type de liaisons sont détaillées. Notez que la structure des polysaccharides, et particulièrement de leurs chaines secondaires sont incomplètes.

Au cours de l'évolution des plantes, les parois ont changées créant une grande diversité de compositions. La cellulose se dégage comme un composant ubiquitaire des organismes photosynthétiques comprenant les Archeaeplastidieae, les cyanobactéries, les Phaeophyceae (algues brunes), mais aussi certain groupes non-photosynthétiques comme les Oomycètes (Niklas, 2004). Certain de ces groupes sont très éloignés en terme d'apparentement, le cas extrême étant les cyanobactéries qui ne sont même pas des Eucaryotes. L'explication viendrait des évènements d'endosymbiose. En effet, l'ancêtre commun des Archeaeplastidieae à acquis sa capacité à faire de la photosynthèse suite à l'endosymbiose primaire d'une cyanobactérie (Bacteria) (Archibald, 2015). Cet évènement a aussi provoqué un transfert horizontal de gènes qui aurait contenu l'ancêtre des celluloses synthases que l'on retrouve aujourd'hui chez les Archeaeplastidieae, expliquant la similarité génétique des celluloses synthases et la présence de cellulose chez ces deux groupes d'organismes très éloignés (Niklas, 2004). Pour les autres groupes comme les Oomycetes, cela proviendrait d'un évènement d'endosymbiose secondaire sur lequel de nombreuses interrogations persistent (Poppers et al., 2011; Archibald, 2015). Les xyloglucanes et les pectines sont présentes chez toute les plantes vertes sauf pour les Chlorophyceae où les pectines sont remplacées par les ulvanes (Poppers et al., 2011).

1.2.4. Les mucilages des plantes sans graines

Comme les mucilages des plantes sont composés des mêmes trois grands types de polysaccharides que les parois, toute la machinerie impliquée dans sa formation (synthèse, transport, modification) pourrait donc être le socle à partir duquel les mucilages végétaux ont évolué. Il est intéressant de noter le parallèle qui existe entre les mucilages végétaux et les matrices polymériques extracellulaires des bactéries (biofilm) qui participent à la formation et au maintien de communauté bactérienne complexe (Satpathy et al., 2016). De façon intrigante considérant l'évènement d'endosymbiose primaire précédemment évoqué, des matrices polysaccharidiques extracellulaires existent chez certaines cyanobactéries actuelles. *Anabaena azollae* se recouvre de mucilage pour rester au contact de la surface de sa plante hôte symbiotique : les azollas (Robins et al., 1986). *Nostoc commune* est capable de former des colonies macroscopiques enrobées d'une matrice extracellulaire composée de glycanes qui leur confère une résistance hors du commun à la sècheresse (Potts, 1994). Cette espèce est aussi capable de secréter des polysaccharides et des pigments en réponse à des rayonnements UV (Ehling-Schulz et al., 1997).

Dans les différents groupes de plantes verte il est possible de trouver du mucilage dans une large diversité d'organe et de rôles. Les algues vertes appartenant à l'ordre des Desmidiales (Streptophytes) comme *Closterium acerosum*, *Micrasterias denticula*, ou encore *Netrium digitus* sont capables de secréter du mucilage de façon localisée et/ou de façon abondante (Domozych, 1999; Eder and Lütz-Meindl, 2008; Eder and Lütz-Meindl, 2010). Des algues appartenant aussi aux Streptophytes mais ayant colonisé le milieu terrestre sont capables de secréter du mucilage pigmenté en guise d'écran solaire (Busch and Hess, 2021).

Les Anthocérothes (Bryophytes) secrètent des mucilages au niveau de plusieurs organes (Frangedakis et al., 2021). La cellule apicale de leur thalle qui assure leur croissance est recouverte de mucilage pour éviter son dessèchement. Les thalles dans leur ensemble sont aussi mucilagineux à l'extérieur mais aussi à l'intérieur de cavités et de canaux débouchant sur l'extérieur permettant à l'algue symbiotique Nostoc de pénétrer et de s'installer (Frangedakis et al., 2021).

Les hépatiques (Bryophytes) possèdent des papilles à mucilage sur la face inferieure de leurs thalles (Galatis and Apostolakos, 1977). Elles ont aussi du mucilage à l'intérieur des tissus au niveau de leur parenchyme. Dans le genre Treubia, il peut même être secrété en large quantité à l'extérieur et présente une composition différente de celle du mucilage secrété par les papilles (Duckett et al., 2006). Le genre Haplomitrium peut aussi secréter du mucilage par ses axes souterrains, un organe spécifique du taxon, ce qui pourrait aider des champignons symbiotiques à les coloniser (Carafa et al., 2003).

Les mousses (Bryophytes) ont des poils axillaires sécréteurs de mucilage, similaires aux papilles des hépatiques, qui participent à la protection des jeunes feuilles et que l'on retrouve jusqu'aux espèces de mousses qui ont divergées le plus tôt (Zhang et al., 2016).

Amauropelta cheilanthoides, Amauropelta pachyrhachis, et *Steiropteris decussata* sont trois especes de fougères qui sont enveloppées d'une couche de mucilage très hydrophile au niveau de leurs méristèmes aériens, des *primordia* foliaires, et les jeunes feuilles (Oliveira et al., 2017). Ce mucilage est produit dans des trichomes glandulaires et protège la plante de l'excès de soleil et d'une trop grande évapotranspiration.

Ce tour d'horizon (non exhaustif) des mucilages existant chez les différents groupes de plantes montre que la sécrétion de mucilage n'est pas rare mais bien diverse en terme de fonction et d'organes concernés.

1.2.5. Les mucilages chez les plantes à graines

A la surface des feuilles de la plupart des espèces arbustives des Spermatophytes (et particulièrement chez les Gymnospermes) se trouve du mucilage qui remplit les petites cavités situé au-dessus des stomates pour participer à la régulation du flux d'eau (Westhoff et al., 2009). Chez les plantes à fleurs, le mucilage peut être détecté au niveau des tiges, des feuilles, des racines et de la graine (**Figure 9**, Galloway et al., 2020). C'est le plus souvent une matrice gélatineuse viscoélastique de polysaccharides de haut poids moléculaire qui est secrétée dans l'environnement. Sa nature et sa fonction dépendent de l'espèce et de l'organe. Par exemple, les petites goutes collantes *a priori* purement pectiques au bout des trichomes des feuilles des Droséras (Drosera) ou les larges quantités contenues dans les feuilles spécialisées des Népenthes (Nepenthaceae) sont capables d'immobiliser des insectes (Adlassnig et al., 2010). Toujours dans un rôle d'adhésion mais au niveau des tiges, des espèces grimpantes comme le lierre (*Hedera helix*) secrètent un mucilage composé de pectines et de protéines à arabinogalactanes (AGP) contre leur support pour s'y fixer solidement (Melzer et al., 2010). La cuscute (Cuscuta), une plante parasite de plante, secrète non seulement du mucilage par ses tiges pour aider son adhésion à la tige de la plante parasitée mais aussi pour protéger la tige modifiée qui lui sert à ponctionner la sève de la plante dont elle se nourrit (Schaffner, 1979; Vaughn, 2002). Ce mucilage peut être composé de pectines, callose, cellulose, AGP et extensine (Vaughn, 2002). Chez de nombreuses espèces, les racines secrètent du mucilage dans la zone du sol à proximité immédiate appelée la rhizosphère. Le mucilage racinaire peut impacter à la fois la composition physicochimique du sol mais aussi les communautés microbiennes qui s'y développent (Sasse et al., 2018). Sa composition polysaccharidique peut varier en fonction des espèces mais ils sont souvent composés de galactose, fucose, mannose, glucose, arabinose, xylose et d'acides uroniques (comme l'acide galacturonique et l'acide glucuronique) (Nazari et al., 2021). Chez *Arabidopsis thaliana*, le mucilage est accumulé dans quelques cellules de la coiffe racinaire pour ensuite être relâché grâce à une dégradation partielle de leur paroi (Maeda et al., 2019). Chez le soja (*Glycine max*), de nombreuses cellules de la coiffe racinaire sont capables de se détacher et de produire un mucilage composé de pectines, cellulose, hémicellulose (xyloglucanes et heteromannanes) mais aussi d'ADN extracellulaire et d'histones qui

empêchent l'Oomycète pathogène *Phytophtora parasitica* de coloniser la racine (Ropitaux et al., 2020).

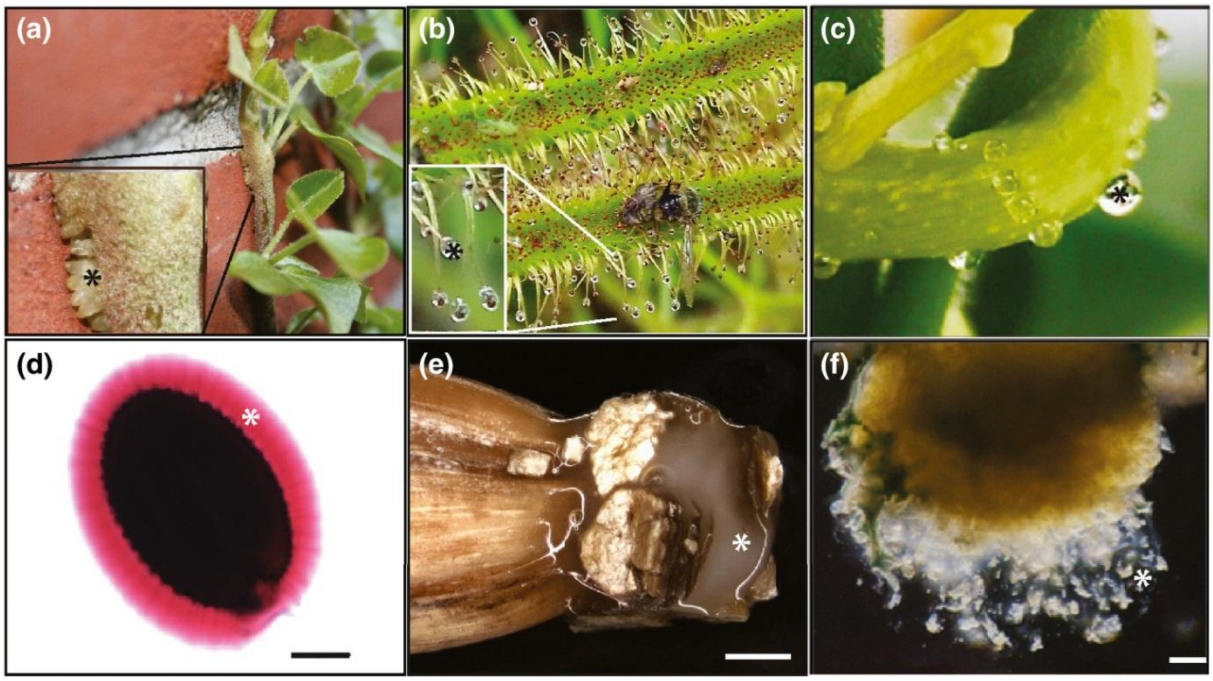


Figure 9 : Exemples de mucilages collants secrétés par des plantes à fleurs (extrait de Galloway et al., 2020). (a) mucilage de lierre lui permettant de s'accrocher à la brique; (b) goutes collantes d'une drosera ayant piégé un insecte; (c) goute de mucilage sur la tige d'une cuscute; (d) mucilage d'une graine d'*Arabidopsis thaliana* coloré au rouge de ruthénium; (e) racine primaire d'orge tout juste sortie de la graine nappé de mucilage qui maintien collé de la vermiculite a son extrémité; (f) Haustorium de cuscute en coupe qui secrète du mucilage contenant des enzymes pour dissoudre des parois. *, mucilage ; Barre d'échelle = 0,1 mm.

1.3. Les mucilages des graines (Myxospermie)

1.3.1. Les origines évolutives de la graine et de son manteau

La graine est apparue une seule fois au cours de l'évolution et a, entre autres adaptations, permis l'évolution radiative impressionnante des plantes à graines comparée aux autres divisions des plantes terrestres (Doyle, 2013) (**Figure 4**). Son apparition découle de l'hétérosporie (petit gamétophyte male et gros gamétophyte femelle), du développement de structures pour recevoir le pollen, et de l'apparition du tégument (Linkies et al., 2010). Issue du tégument de l'ovule, le tégument (aussi appelé testa ou manteau de la graine) est un tissu diploïde génétiquement maternel (Evenari, 1984). De par sa position externe dans la graine, le testa va se retrouver en contact direct avec l'environnement. C'est donc lui qui va assurer la protection de la graine face aux contraintes biotiques (pathogènes et prédateurs) et abiotiques (radiations, humidité, température, etc…) (Rajjou et Debeaujon, 2008). Chez les Angiospermes le testa varie fortement en épaisseur entre les espèces, impliquant des structures tissulaires ou des tailles cellulaires variables (Coen and Magnani, 2018). Les différentes couches

cellulaires du testa peuvent même varier en nombre et en taille entre des espèces proches comme chez les Brassicaceae (Vaughan et Whitehouse, 1971). Chez la plante modèle *Arabidopsis thaliana* le manteau de la graine comporte cinq couches de cellules correspondant aux deux couches principales du tégument qui restent facilement visibles après le dessèchement de la graine qui entraine la compaction des cellules (**Figure 10**). La couche cellulaire la plus interne (ii1) est riche en proanthocyanidines (pigments) qui, en plus de donner leur couleur aux graines, participent à la protection face au rayonnement UV et augmentent l'imperméabilité des graines pour maintenir leur viabilité dans le temps (Rajjou et Debeaujon, 2008). La couche épidermique la plus externe (oi2) est composée des cellules à l'origine de la sécrétion du mucilage de la graine qui font l'objet des parties suivantes.

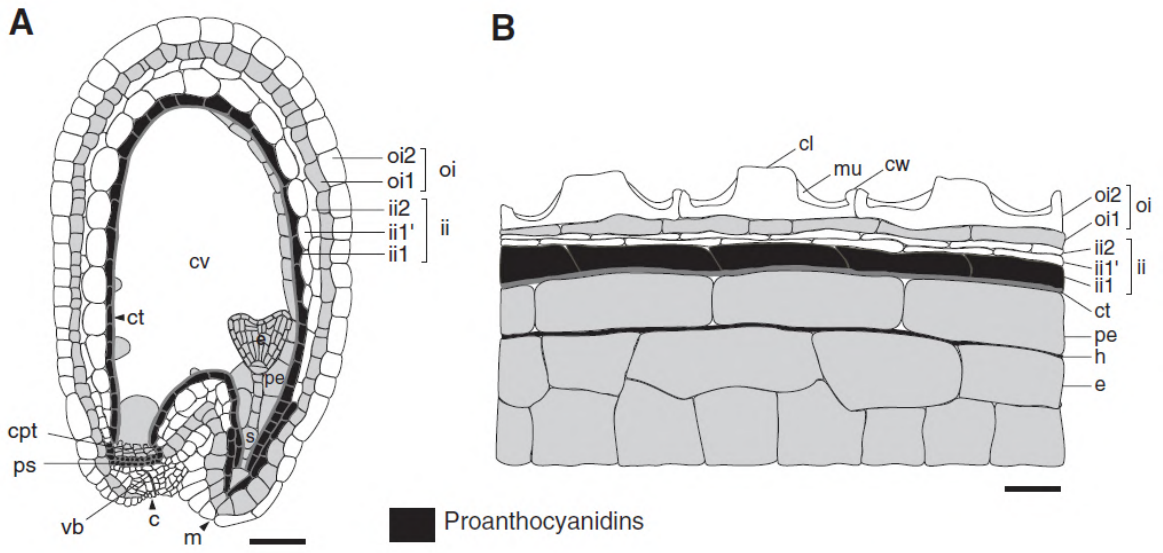


Figure 10 : Structure du testa d'*Arabidopsis thaliana* (extrait de Debeaujon et al., 2007). (A) Coupe longitudinale d'une graine en développement au stade embryonnaire « cœur ». Barre d'échelle = 40µm **(B)** Coupe d'un testa mature et sec. Barre d'échelle = 7 µm. La couche cellulaire contenant des proanthocyanidines est coloré en noir. Abréviations : (oi) tégument externe, (ii) tégument interne, (cv) vacuole centrale, (ct) cuticule, (pe) endosperme périphérique, (cpt) tissus du chalazal en prolifération, (ps) brin de pigment, (vb) faisceau vasculaire, (c) chalaza, (m) micropyle, (cl) columelle, (mu) mucilage, (cw) paroi cellulaire, (h) couche hydraline, (e) embryon.

1.3.2. L'évolution et les rôles écologiques de la myxospermie

Préface à l'article de revue

Cette partie a été publiée sous forme de revue en 2020 dans *Plant, Cell and Environment* lors d'une issue spéciale portant sur l'évolution des réponses aux stress environnementaux chez les plantes. Elle est éditée au format du journal et contiendra donc sa propre bibliographie pour les références qui y sont citées. Cet article est le fruit de mes recherches bibliographiques qui avaient pour but de trouver les potentiels points communs évolutifs de la myxodiasporie (capacité à secréter du mucilage autour

du fruit ou de la graine) au niveau de son contrôle génétique, du développement des cellules sécrétrices du mucilage, et des rôles écologiques associés. En effet dans la littérature scientifique il existe de nombreuses descriptions de mucilage concernant des espèces appartenant à la plupart des Ordres d'Angiospermes (**Figure 11**). Ils sont très divers morphologiquement et biochimiquement (Western, 2012; Phan and Burton, 2018) mais toutes ces formes de myxospermie chez les Angiospermes rendent l'hypothèse qu'elles seraient le fruit d'autant d'apparitions indépendantes très peu parcimonieuses. De plus, l'avancé des connaissances génétiques pour plusieurs espèces d'Angiosperme relativement bien reparties phylogénétiquement commence à permettre leur recoupement pour tenter de comprendre l'évolution des mucilages au travers des gènes qui les contrôlent. C'est donc aux questions suivantes que cet article esquisse une réponse par une revue de la littérature : Quand le mucilage des graines est-il apparu ? Comment se forme-t-il (contrôle génétique et développement cellulaire) ? Pour remplir quels rôles écologiques ?

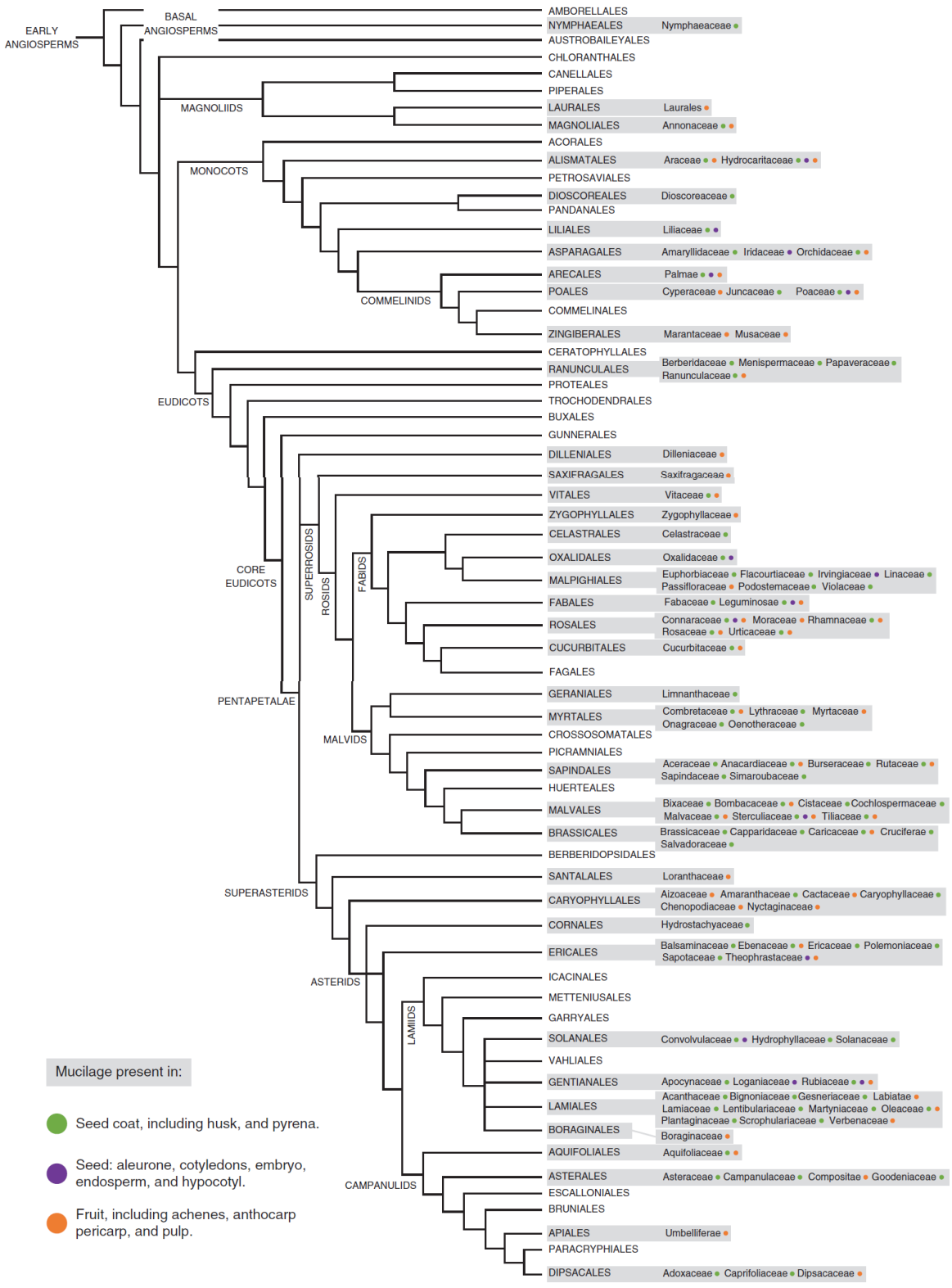


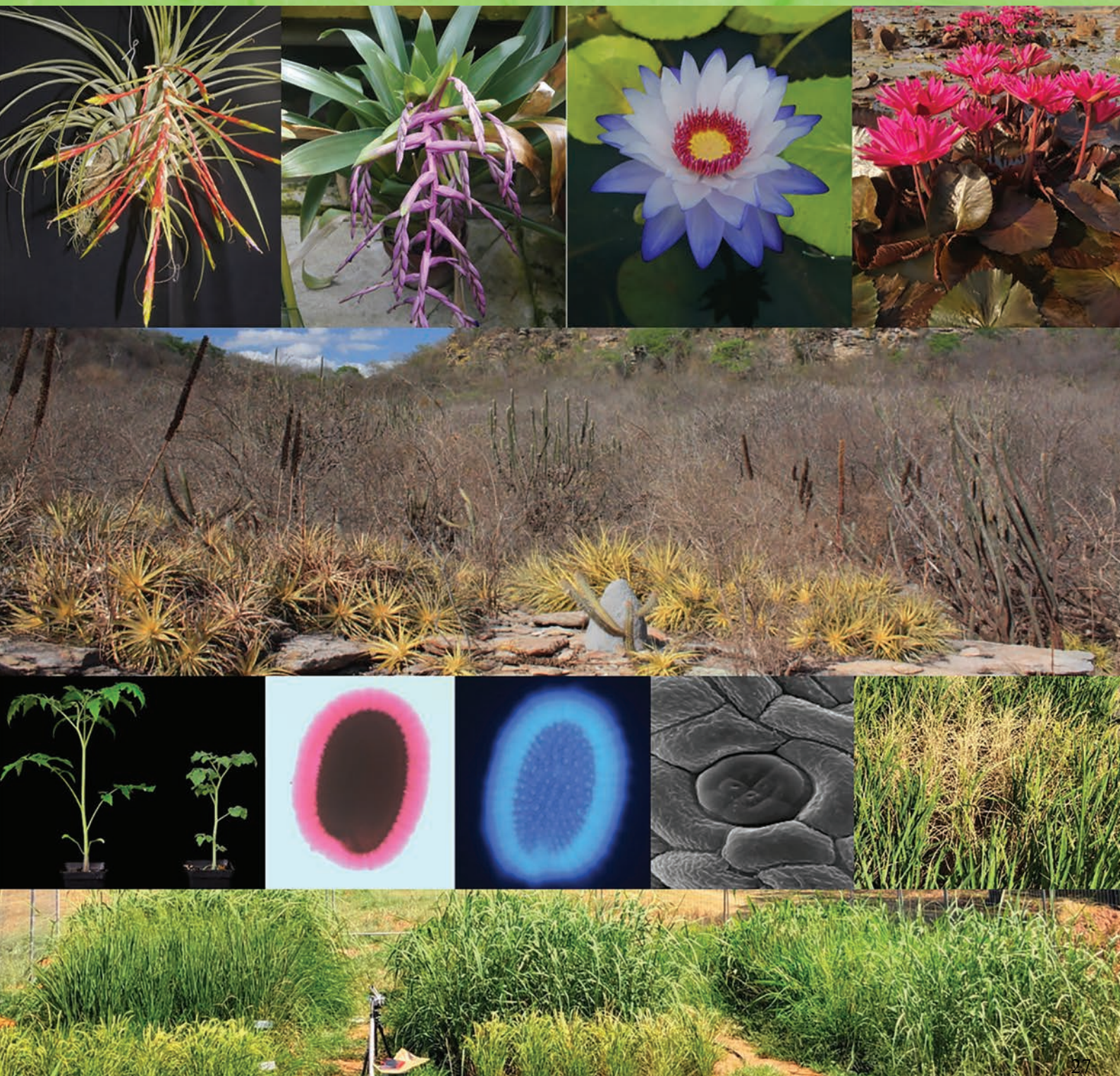
Figure 11 : Les ordres d'Angiospermes ainsi que le détail des familles décrites comme étant myxodiasporiques dans la littérature (extrait de Phan et Burton, 2018). A droite du nom des familles les points vert signalent du mucilage présent dans le testa, les points violet dans un autre compartiment de la graine, les points orange au niveau du fruit.

ISSN 0140 7791

Plant, Cell & Environment

Volume 43 Number 12
December 2020

WILEY



Received: 21 March 2020 | Revised: 12 June 2020 | Accepted: 12 June 2020

DOI: 10.1111/pce.13827

REVIEW

Plant, Cell & Environment WILEY

Seed mucilage evolution: Diverse molecular mechanisms generate versatile ecological functions for particular environments

Sébastien Viudes | Vincent Burlat | Christophe Dunand

Laboratoire de Recherche en Sciences Végétales, CNRS, UPS, Université de Toulouse, Castanet-Tolosan, France

Correspondence
Vincent Burlat and Christophe Dunand, Laboratoire de Recherche en Sciences Végétales, CNRS, UPS, Université de Toulouse, 31326 Castanet-Tolosan, France.
Email: burlat@lrsv.ups-tlse.fr (V. B.) and dunand@lrsv.ups-tlse.fr (C. D.)

Funding information
Agence Nationale de la Recherche, Grant/Award Number: ANR-18-CE20-0007; French Laboratory of Excellence, Grant/Award Numbers: ANR-10-LABX-41, ANR-11-IDEX-0002-02

Abstract

Plant myxodiasporous species have the ability to release a polysaccharidic mucilage upon imbibition of the seed (myxospermy) or the fruit (myxocarpy). This is a widespread capacity in angiosperms providing multiple ecological functions including higher germination efficiency under environmental stresses. It is unclear whether myxodiaspory has one or multiple evolutionary origins and why it was supposedly lost in several species. Here, we summarize recent advances on three main aspects of myxodiaspory. (a) It represents a combination of highly diverse traits at different levels of observation, ranging from the dual tissular origin of mucilage secretory cells to diverse mucilage polysaccharidic composition and ultrastructural organization. (b) An asymmetrical selection pressure is exerted on myxospermy-related genes that were first identified in *Arabidopsis thaliana*. The *A. thaliana* and the flax intra-species mucilage variants show that myxospermy is a fast-evolving trait due to high polymorphism in a few genes directly acting on mucilage establishment. In *A. thaliana*, these actors are downstream of a master regulatory complex and an original phylogenetic overview provided here illustrates that this complex has sequentially evolved after the common ancestor of seed plants and was fully established in the common ancestor of the rosid clade. (c) Newly identified myxodiaspory ecological functions indicate new perspectives such as soil microorganism control and plant establishment support.

KEYWORDS

ecological roles, inter-species natural variability, intra-species natural variability, MBW master regulator, MSC toolbox genes, myxocarpy, myxodiaspory, myxospermy

1 | INTRODUCTION

Mucopolysaccharides, also called mucilage, are found to be produced in early diverging non-vascular plant groups such as hornworts (Anthocerotophyta) which extrude it around organs for various functions such as dehydration protection during growth and reproduction (Renzaglia, Duff, Nickrent, & Garbary, 2000). In flowering plants, several kinds of mucilage with cell wall-like compositions can be secreted by a wide range of organs such as seeds, fruits, roots, leaves and stems, conferring an impressive diversity of physical properties (Galloway, Knox, & Krause, 2020). The term myxodiaspory designates the ability to extrude mucilage upon imbibition from the seed coat or the fruit pericarp (Ryding, 2001; Figure 1). The ability of species to release seed mucilage from the seed coat epidermis is called myxospermy, while the same ability coming from the fruit epicarp outermost cell layer is called myxocarpy (Figure 1).

Seed mucilage presence was reported in a Charles Darwin's letter (Weitbrecht, Müller, & Leubner-Metzger, 2011). The description of

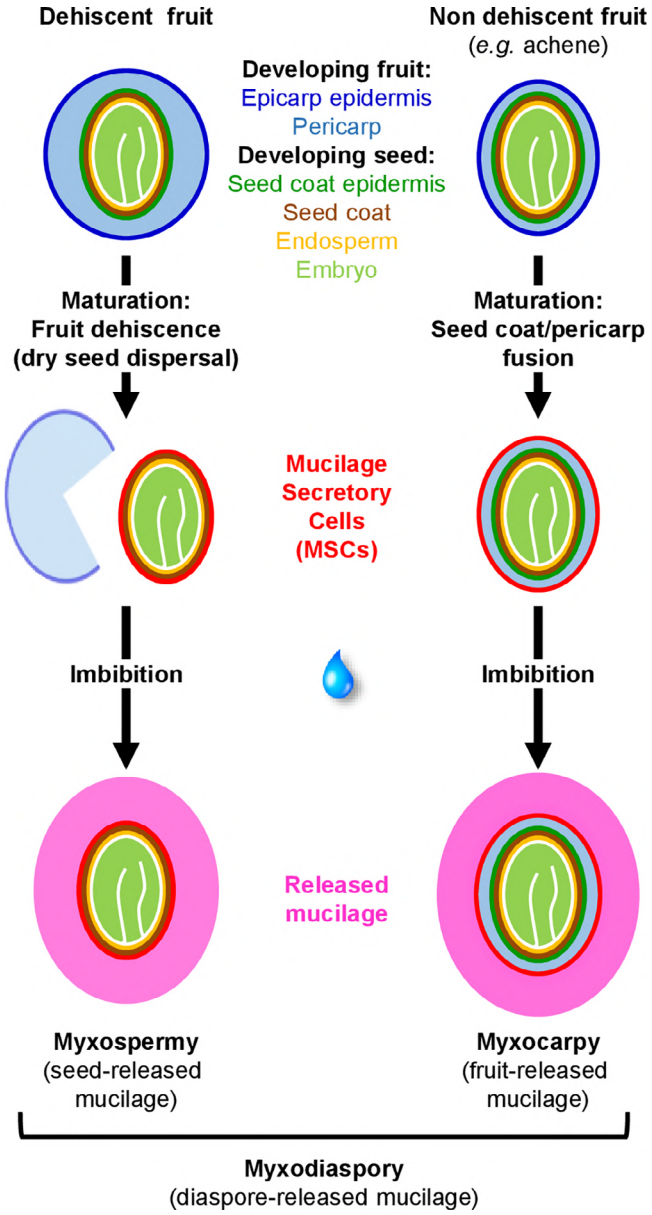


FIGURE 1 Developmental scheme comparing the two epidermal origins of mucilage secretory cells enabling mucilage release in myxospermous versus myxocarpous species. Note that the different layers are not drawn to scale and that more than one seed per fruit is commonly found in myxospermous species

seed mucilage morphology in several species and its putative ecological functions has drawn the interest of the scientific community for over a century. In the last 50 years, myxodiaspory was described in a majority of angiosperm orders associated with the diversity of mucilage secretory cells (MSCs) and the polysaccharidic mixtures they release (reviewed in Western, 2012; Phan & Burton, 2018). Recent reviews have explored the biochemical composition and cell wall dynamics of MSCs during seed development, and the molecular, biochemical and structural characterization of mucilage, in particular in *Arabidopsis thaliana* (Golz et al., 2018; Griffiths & North, 2017; Phan & Burton, 2018; Šola, Dean, & Haughn, 2019). Seed mucilage can represent a significant metabolic cost as, for instance, it accounts for 2–3% of the *A. thaliana* seed mass (Macquet, Ralet, Kronenberger, Marion-Poll, & North, 2007). In the large array of studied species, this metabolic investment was documented to fulfil numerous ecological functions such as seed protection and seed dispersal by direct physical modification of the local environments (reviewed in Western, 2012; Yang, Baskin, Baskin, & Huang, 2012; North, Berger, Saez-Aguayo, & Ralet, 2014). Mucilage specific physical and chemical properties and the easy extraction of these substances led to several applications in the pharmaceutical and food industry, as dietary supplement and biopolymer, respectively (Mirhosseini & Amid, 2012; Soukoulis, Gaiani, & Hoffmann, 2018).

Myxospermy has been reported for at least one species in 81 out of the 416 angiosperm families. This trait is widespread in core Eudicot species with at least one occurrence in 38 rosid families and in 28 asterid families, but also in 9 monocot families. It has been even detected in Annonaceae and Nymphaeaceae, two early diverging angiosperm families (Phan & Burton, 2018; Western, 2012). Interestingly, myxocarpy is less documented but is similarly widespread in angiosperm species (at least one species in 19 rosid, 11 asterid and 8 monocot families). Co-occurrence of myxospermy and myxocarpy appears in up to 25% of these families (Phan & Burton, 2018). In the deeply studied Brassicaceae family, 68 out of the 213 studied species release mucilage around seeds, with a relatively even distribution within the different genera (Vaughan & Whitehouse, 1971). Therefore, it is unclear whether myxodiaspory has one or multiple evolutionary origins and for which reason it was, supposedly, lost in several species. Full answer to these evolutionary questions would benefit from a thorough multidisciplinary mapping of these traits across available phylogenetic trees. Such rare studies concern a few species and are restricted to a single research field (e.g., ..genetic characterization, physiology, development or ecology). The overall aim of this review is to focus on the under-studied field of myxodiaspory evolution. So far, a deep characterization of the molecular actors implicated in seed mucilage establishment was performed primarily in *A. thaliana* constituting the so-called MSCs toolbox (Francoz, Ranocha, Burlat, & Dunand, 2015). Computational analyses could help identify orthologous relationships between *A. thaliana* MSCs toolbox genes and genes in other species, in turn clarifying the origin of this toolbox, and when it could have started to be associated with myxospermy. However, numerous MSC toolbox genes form a complex regulatory network; these genes have unequal contributions to myxospermy and individual genes can display pleiotropic functions (Golz et al., 2018). Therefore, it is very difficult to clearly identify genes associated with myxospermy, merely from computational analyses. Indeed, in *A. thaliana* seeds, a regulatory complex dictates carbon partitioning between storage lipids, flavonoid pigments and the seed mucilage (Golz et al., 2018; Li, Zhang, Chen, Ji, & Yu, 2018). Lipids accumulating in the embryo and the endosperm will provide the required nutriments for proper embryo development, while pigments will confer impermeability and radiation protection (Baroux & Grossniklaus, 2019), and the mucilage will bring additional ecological advantages depending on the species.

The objective of this review is to complement the fields covered by the previous myxospermy-dedicated reviews that we acknowledged above. We aim to gather morphological, evolutionary and ecological information to give an overview of the current state of the literature to better understand seed mucilage evolution. First, we briefly describe the diversity among the various myxodiaspory morphological traits (MSC origins and development patterns, cell wall dynamics involved in mucilage release and mucilage chemical diversity). In the second part, we used the *A. thaliana* MSC toolbox genes to investigate natural variability at both the intra- and inter-species level. This provides some clue on the evolutionary history of the underlying genes. We paid particular attention to a well-studied master regulatory complex that not only regulates mucilage, but also other seed traits including storage lipid and protective pigments. In the last part, we discuss recent insights obtained, since the last reviews (Western, 2012; Yang, Baskin, Baskin, & Huang, 2012), on myxodiaspory numerous ecological functions facing abiotic and biotic constraints, as well as its impact on plant development.

2 | DIVERSITY OF MYXODIASPORY

The capacity for seeds or fruits to release mucilage (myxodiaspory) is influenced by, and composed of, more discrete traits, all displaying a large phenotypic variability. These traits include the MSC origin, the MSC morphology, the cell wall dynamics during MSC development, the mucilage extrusion mode, the mucilage polysaccharidic composition and ultrastructural organization. However, these traits are not evenly documented among the studied species (Phan & Burton, 2018; Western, 2012). Most studies report macroscopic phenotypes attesting to the presence or the absence of released mucilage (Western, 2012). Few species, such as the pioneer model species *A. thaliana* and the emerging model species *Linum usitatissimum* (flax) provide multi-level information, illustrating the diversity within the myxodiaspermy-associated traits.

MSCs correspond to the outermost seed coat or fruit pericarp epidermal cell layer enabling mucilage extrusion to the environment (Figures 1 and 2). As exemplified by *A. thaliana* and *L. usitatissimum* (flax), this specific cell layer is differentiated during seed development to become a dead layer at the end of seed maturation (Figure 2; Western, 2001; Miart et al., 2019). During the MSC development, the *A. thaliana* seed mucilage is trapped, with no apparent sub-layering, between the outer periclinal primary wall and a volcano-shaped polarized secondary wall called columella (Figure 2). In flax, a complex multilayered mucilage is sequentially deposited in the MSCs and becomes visible after extrusion (Figure 2). Beyond these two models, the mature MSC morphological diversity is also illustrated in the various species studied over the years (Phan & Burton, 2018), as exemplified by the pioneer morphological survey of mature MSCs conducted on 200 Brassicaceae species covering 90 genera (Vaughan & Whitehouse, 1971).

Upon seed imbibition, sequential events occur within seconds in all species. First, the hydrophilic nature of the mucilage polysaccharide mixture allows water absorption. This induces a mucilage swelling pressure breaking the MSC primary cell wall (Figure 2). Finally, this leads to mucilage extrusion outwards from the seed. As a result, the seed mucilage volume and mass increase up to 75-fold in *Capsella bursa-pastoris* (Deng, Jeng, Toorop, Squire, & Iannetta, 2012) and 5.5-fold in *Henophyton deserti* (Gorai, el Aloui, Yang, & Neffati, 2014), respectively. However, subtle differences occur in seed mucilage extrusion modes contributing to the diversity of this trait (Figure 2). Indeed, the rupture of peculiar primary wall domains occurs either simultaneously in all *A. thaliana* MSCs or sequentially in adjacent flax MSCs (Figure 2). This organized explosion is carefully prepared earlier during MSC seed development by differential cell wall polysaccharide deposition and localized modifications. The polysaccharidic-proteinaceous molecular scaffold enabling the *A. thaliana* primary wall domain loosening starts to be uncovered (Francoz et al., 2019; Kunieda et al., 2013; Saez-Aguayo et al., 2013). In flax, the MSC polysaccharide composition and its internal organization are proposed to play a role in proper MSC opening (Miart et al., 2019).

The structure and polysaccharide composition of seed mucilage directly contribute to the observable diversity of myxospermy. In *A. thaliana*, the released mucilage is separated between an adherent layer bound to the seed and a non-adherent layer, both enriched in poorly branched type I rhamnogalacturonan pectin (Figure 2; Macquet, Ralet, Kronenberger, et al., 2007; Poulain, Botran, North, & Ralet, 2019). In flax, the mucilage is composed of four contrasted layers enriched in type I rhamnogalacturonan, arabinoxylans and xyloglucans/cellulose, respectively (Figure 2; Kreitschitz & Gorb, 2017; Miart et al., 2019). This type of seed mucilage variable layering and composition also exists in other species such as for example, *Lepidium perfoliatum* (Huang, Wang, Yuan, Cao, & Lan, 2015), *Neopallasia pectinata* (Kreitschitz & Gorb, 2017) or *Plantago ovata* (Tucker et al., 2017; Yu et al., 2017). The comparison among multiple species has demonstrated the diversity of seed mucilage microstructure observed by scanning electron microscopy (Kreitschitz & Gorb, 2018). Studying seed mucilage allowed characterization of polysaccharide–polysaccharide specific interactions (Yu et al., 2018) making seed mucilage an excellent model for cell wall dynamics understanding (Arsovski, Haughn, & Western, 2010).

Interestingly, for myxocarpous species such as *Salvia* and *Artemisia* species, the mucilage is extruded by the outermost fruit cell layer, named the achene (non-dehiscent fruit) pericarp, and not by the seed integument (Ryding, 2001). These could be examples of evolutionary convergences leading to the similar differentiation into MSCs of different types of outer cells facing the environment. Indeed, the *Salvia hispanica* (chia) achene mucilage and MSCs show interesting parallels with myxospermous species such as *A. thaliana* and *L. usitatissimum*. Mucilage accumulates in the outer pericarp epidermal cells during chia seed development. After extrusion, the mucilage remains indirectly attached to the seed via the inner pericarp-seed tegument contact (Geneve, Hildebrand, Phillips, Al-Amery, & Kester, 2017).

Finally, an additional peculiarity may exist in *Medicago truncatula* and *M. orbicularis*, the cell wall of the endosperm forms a mucilage gel between the seed coat and the embryo (Song et al., 2017).

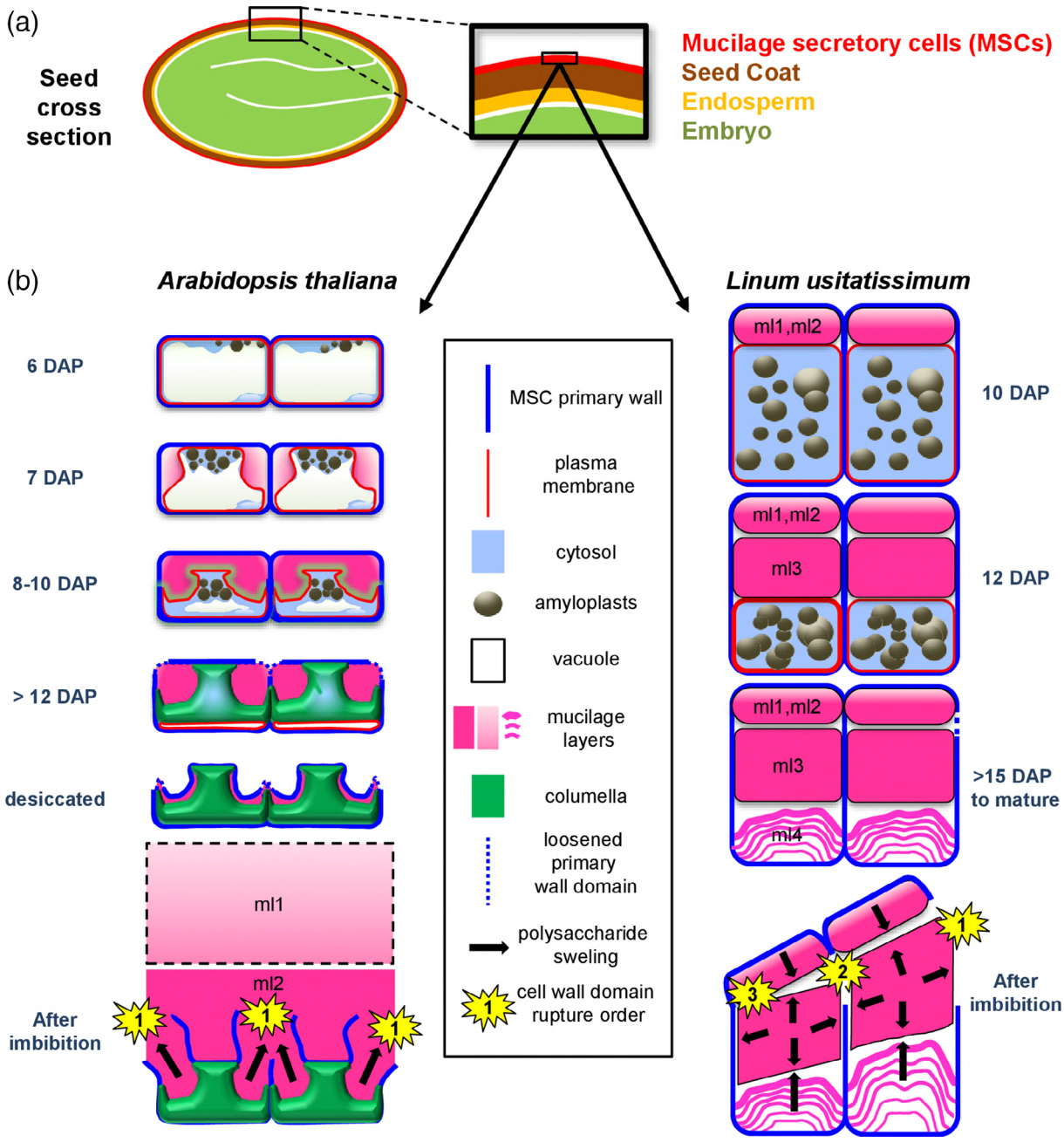


FIGURE 2 *A. thaliana* and *L. usitatissimum* (flax) as model species for seed mucilage secretory cell (MSC) development illustrating the diversity of cell wall dynamics, intracellular polysaccharidic mucilage organization, and extrusion mechanisms upon water imbibition. (a) conceptual seed cross section wide view and magnification illustrating the localization of seed MSCs. (b) Kinetics of MSC development in the historical model *A. thaliana* on the left (adapted from (Francoz et al., 2019; Western, 2001), and in the emerging model *L. usitatissimum* on the right (Adapted from reference Miart et al. (2019).). The 30–50 μm wide cells from both species are drawn to similar scale, the released mucilage layers are not drawn to scale. Note the numerous differences between both models: Major features in *A. thaliana*: highly complex cell wall dynamics with the presence of a volcano-shaped columella, the simultaneous rupture of primary cell wall domain in each MSC and the distinction between adherent mucilage (am) and non-adherent mucilage layers (nam). Major features in *L. usitatissimum*: sequential synthesis of four highly different mucilage layers (m1–m4) and sequential rupture of primary wall domains expending from cell to cell. DAP, day after pollination

Therefore, myxodiaspory is a trait that encompasses multiple levels of diversity, including MSC structure, mucilage extrusion mode and mucilage polysaccharidic composition and structural organization. However, there is no simple clear-cut distribution of the myxodiasporous/non-myxodiasporous traits along the angiosperm families and even within a given family as exemplified for Brassicaceae (Vaughan & Whitehouse, 1971). For this reason, in the following part, we will shed light on molecular mechanisms underlying this morphological diversity through intra-species and inter-species comparative studies for seed mucilage evolution understanding.

3 | EVOLUTION OF THE MOLECULAR ACTORS UNDERLYING THE INTRA- AND INTER SPECIES MYXOSPERMY NATURAL VARIABILITY

3.1 | The mucilage secretory cell (MSC) toolbox genes may be separated into two groups—Upstream regulators and downstream actors

Twenty years of forward and reverse genetics together with more global approaches have allowed functional characterization of numerous genes involved in seed mucilage and mucilage secretory cell (MSC) physiology in *A. thaliana*. Recent reviews sequentially reported 54 genes (Francoz et al., 2015) and 82 genes (Phan & Burton, 2018). Since then, 12 additional genes have been characterized (Fabrissin et al., 2019; Kunieda, Hara-Nishimura, Demura, & Haughn, 2019; Li et al., 2018; Shimada et al., 2018; Šola, Gilchrist, et al., 2019; Takenaka et al., 2018; van Wijk et al., 2018; Voiniciuc et al., 2018; Wang et al., 2019; Yang et al., 2019) bringing the list to 94 genes so far. They constitute the continuously growing MSC toolbox that participate to a proper seed mucilage production and release in *A. thaliana* (Francoz et al., 2015; Voiniciuc, Yang, Schmidt, Günl, & Usadel, 2015). A majority of these genes are transcription factors including well characterized upstream master regulators that will be further discussed hereafter, and less characterized regulatory genes whose position in the gene regulatory network is still puzzling (Golz et al., 2018). The downstream targets mainly encode direct actors responsible for seed mucilage synthesis, assembly and secretion (Francoz et al., 2015) (Figure 3).

3.2 | The myxospermy intra-species natural variability reveals a strong selection pressure targeted on a few MSC toolbox downstream actors

Detecting and understanding the effects of selection pressure on the MSC (MSC) toolbox genes should be easier when considering intra-species rather than inter-species natural variability, because changes are still relatively recent on the evolutionary timescale and are scarcer. Indeed, the natural diversity occurring in *L. usitatissimum* cultivars (Liu et al., 2016; Miart et al., 2019), or in *A. thaliana* natural ecotypes (Saez-Aguayo et al., 2014; Voiniciuc et al., 2016) shows a gradient of mucilage abundance and myxospermy efficiency. This can culminate

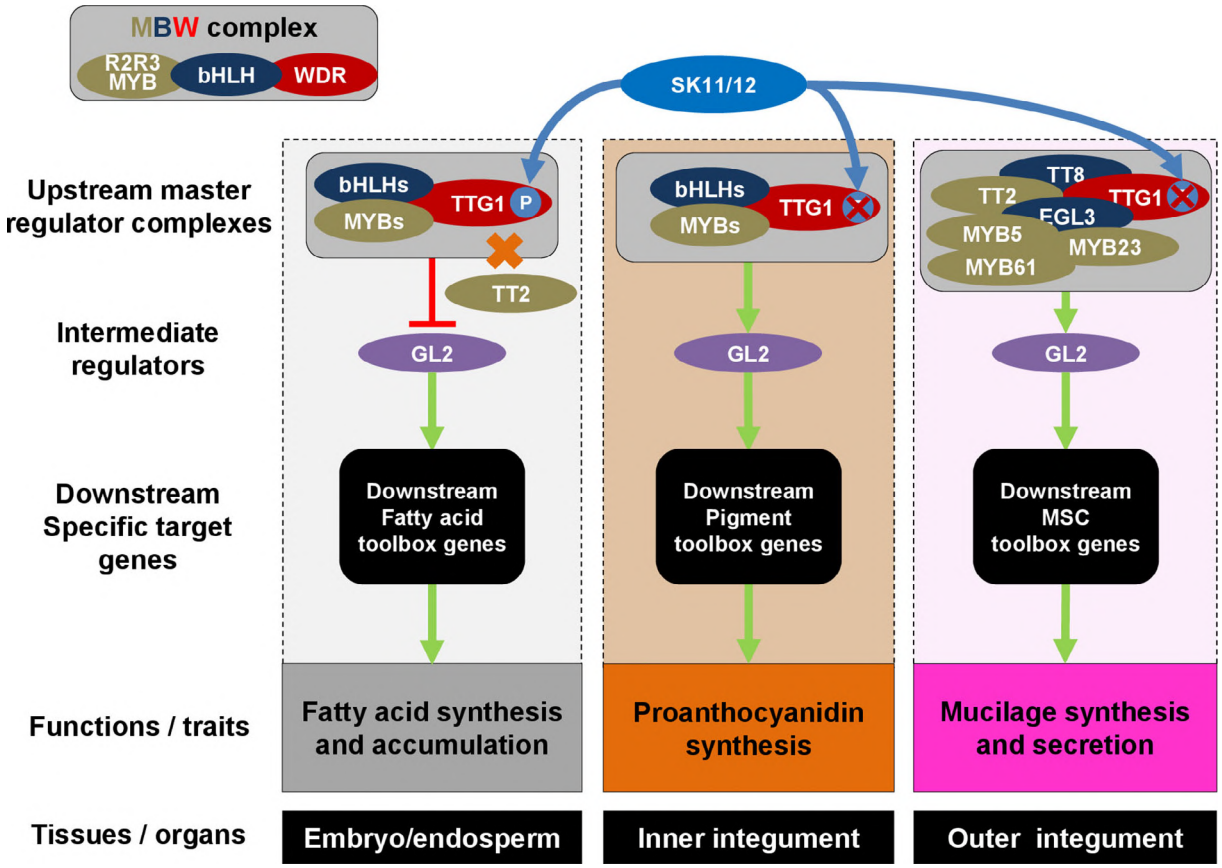


FIGURE 3 The MYB-bHLH-WDR (MBW) complexes regulate the spatiotemporal carbon partitioning in *A. thaliana* seeds through a complex modularity. The conserved ancestral master regulator WDR member (TTG1) and the phosphorylation status of a conserved serine (P), together with a bHLH and MYB modularity of competitive interaction enable the regulation of the spatiotemporal specificity of various seed traits including mucilage production and release (Chen & Wang, 2019; Golz et al., 2018; Li et al., 2018)

in a complete loss of mucilage extrusion for *A. thaliana* in some natural populations, such as Shahdara (Sha) originating from 3,400-m high mountains in Tadjikistan (Macquet, Ralet, Loudet, et al., 2007; https://www.arabidopsis.org/) or the loss of adherent mucilage extrusion for Rakit-1 (Rak-1) originating from 220-m high pine forests in southern central Russia (Saez-Aguayo et al., 2014; https://www.arabidopsis.org/). Interestingly, the absence of adherent mucilage extrusion does not necessarily mean a lack of mucilage synthesis since Rak-1 releases even more non-adherent mucilage than Col-0 reference ecotype (Saez-Aguayo et al., 2014). In *A. thaliana* natural populations, all characterized myxospermy-related mutations leading to pseudogenization converge on three downstream genes of the MSCs toolbox, namely *PECTINMETHYLESTERASE INHIBITOR6* (*PMEI6*) for Djarly (Dja, 2,500-m rocky slopes in Kyrgyzstan, Saez-Aguayo et al., 2013); *MUCILAGE-MODIFIED2/BETA-GALACTOSIDASE6* (*MUM2/BGAL6*) for Sha (Macquet, Ralet, Loudet, et al., 2007) and *PEROXIDASE36* (*PRX36*) for Sk-1 (soil from cliff in old cultivated landscape in Norway, Saez-Aguayo et al., 2014) (Figure 4). This represents a low number of genes

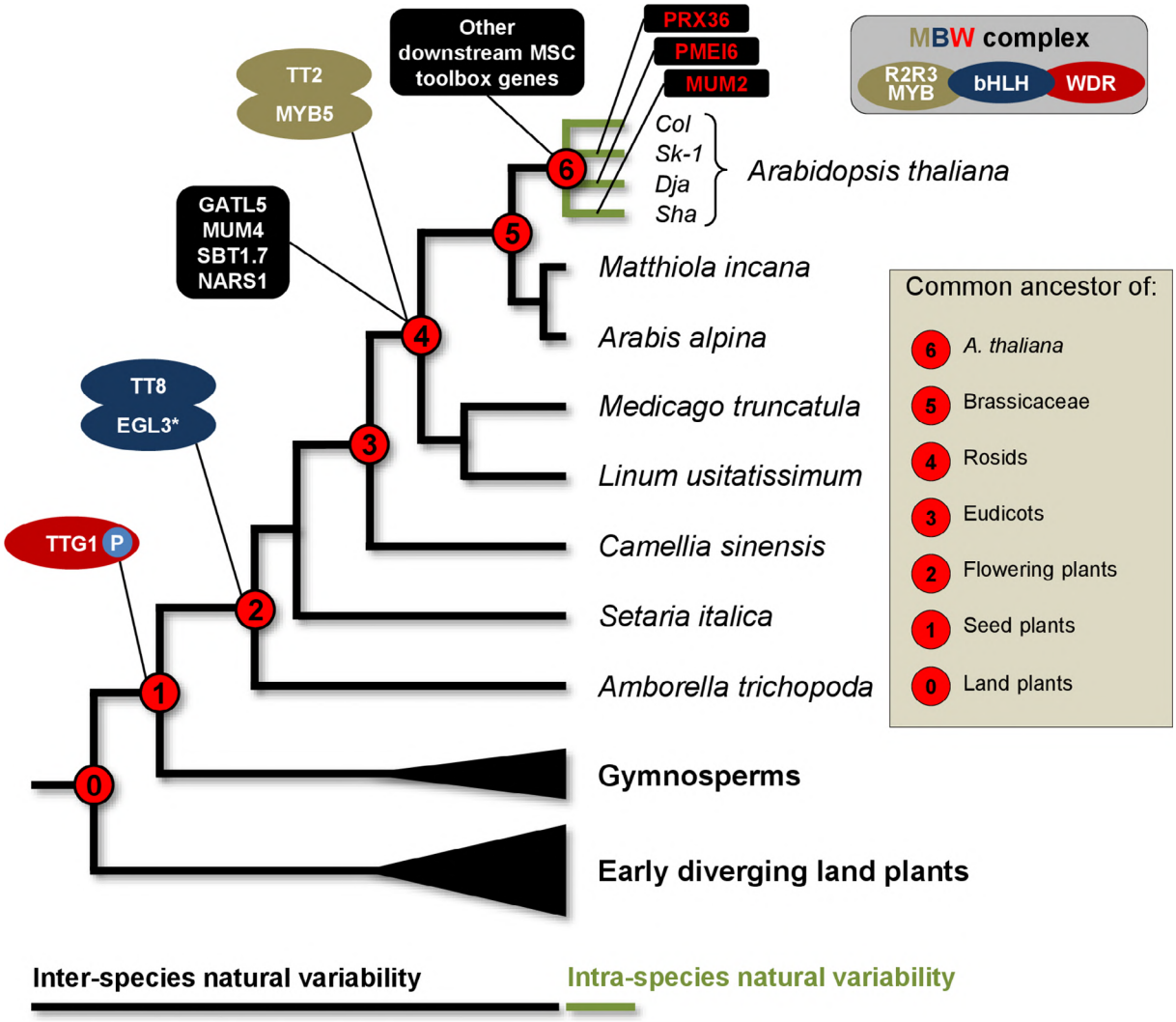


FIGURE 4 Simplified phylogenetic tree of land plants giving an overview of the sequential evolution of the *A. thaliana* mucilage secretory cell (MSC) toolbox based on the available knowledge at both the inter- and intra-species level. Note that the members of the MBW complex controlling the myxospermy trait in *A. thaliana* sequentially evolved and were complete at the rosid node. Note also that except *GALACTURONOSYLTRANSFERASE-LIKE 5* (GATL5), *MUCILAGE MODIFIED4* (MUM4), *SUBTILISIN-LIKE SERINE PROTEASE 1.7* (SBT1.7) and *NAC-REGULATED SEED MORPHOLOGY 1* (NARS1) that are functional in both *A. thaliana* and *L. usitatissimum*, most of the other downstream enzymatic actors of the MSC toolbox, until counter-example is provided, could be specific to *A. thaliana*. Species names on the tree leaves are the ones use in the reviewed studies. Gene names in white represent their apparition (last duplication/speciation) whereas genes names in red show their pseudogenization in *A. thaliana* ecotypes. EGL3* is the ancestor gene of AtEGL3, AtGL3 and AtMYC1 that were later duplicated and diverged within the Brassicaceae family. TTG1 phosphorylation site (P) seems to be conserved together with the gene itself. The genome wide association, phylogenomic, transcomplementation and functional characterization studies used to build this figure come from Airoldi, Hearn, Brockington, Webb, and Glover (2019), Chopra et al. (2014), Dressel and Hemleben (2009), Doroshkov, Konstantinov, Afonnikov, and Gunbin (2019), Pang et al. (2009), Li et al. (2018); Liu, Jun, and Dixon (2014), Liu et al. (2017, 2018), Saez-Aguayo et al. (2014), Soto-Cerda et al. (2018), Zhang, Chopra, Schrader, and Hülskamp (2019) and Zhang and Hülskamp (2019)

as compared to the MSC toolbox size and considering the nearly 70 studied natural accessions, all mutated for myxospermy (Saez-Aguayo et al., 2014). This strongly suggests that downstream MSC toolbox genes undergo strong selection pressure, but also that myxospermy is a fast-evolving trait as exemplified by the seven independent *MUM2/BGAL6* natural mutants found in only two different geographic areas (Saez-Aguayo et al., 2014). *PMEI6*, *MUM2/BGAL6* and *PRX36* encode enzymes necessary for proper seed mucilage hydration and extrusion (Dean et al., 2007; Kunieda et al., 2013; Saez-Aguayo et al., 2013) and PMEI6 and PRX36 functions are tightly and sequentially related (Francoz et al., 2019). As non-myxospermous seeds have a much better buoyancy efficiency and considering that the habitat of Sha is close to a river, seed dispersal by water run-off is one of the seed mucilage functions proposed to explain the loss of myxospermy (Macquet, Ralet, Loudet, et al., 2007; Saez-Aguayo et al., 2014). Unfortunately, no clear association can be established between the natural population habitats and their mucilage phenotypes (Voiniciuc et al., 2016) but the observed relative selection pressure towards a mucilage disappearance highlights its ambivalent function for a myxospermous species such as *A. thaliana*.

More recently, genome wide association studies (GWAS) conducted on *A. thaliana* (Fabrissin et al., 2019) and *L. usitatissimum* (Soto-Cerda et al., 2018) allowed the identification of the statistically most relevant single nucleotide polymorphisms explaining the observed seed mucilage phenotype. In *L. usitatissimum*, out of six loci implicated in mucilage content, five loci highlight putative orthologs of previously characterized *A. thaliana* MSC toolbox genes (Soto-Cerda et al., 2018). In *A. thaliana*, these five genes are two regulators and three downstream direct actors in seed mucilage synthesis or modification (for more details see the following part). In *A. thaliana*, through a very precise molecular phenotyping, GWAS revealed only eight peaks significantly above the high background of less implicated positions, pointing to seed mucilage being a polygenic trait (Fabrissin et al., 2019). Upon the eight candidates, two genes were identified, one already known to belong to the MSC toolbox and another one characterized as implicated in the production of seed mucilage pectin content (Fabrissin et al., 2019). These results highlight the fact that the MSC toolbox starts to be well characterized in *A. thaliana* and that it can be used to investigate whether functional orthologs are present in other species. It also suggests that the selective pressure on myxospermy mostly applies on downstream target genes of the MSC toolbox. Hopefully, this conservation of upstream genes of the toolbox regulation network will allow to study the evolutionary origins of part of the MSC toolbox genes across angiosperms.

3.3 | Angiosperm inter-species natural variability reveals the genetic conservation of a master regulatory complex involved in myxospermy

In *A. thaliana*, some of the master regulators belonging to the MSC (MSC) toolbox also regulate the formation of trichomes and root hairs (Jones & Dolan, 2012), the flavonoid biosynthesis (anthocyanidins and proanthocyanidins) (Lloyd et al., 2017), and the seed carbon partitioning (Chen & Wang, 2019; Golz et al., 2018; Li et al., 2018). The combinations of specific MYB and bHLH transcription factors together with TRANSPARENT TESTA GLABRA 1 (TTG1), a WD40 domain repeats (WDR) transcription factor allows the regulation of each of these traits. They constitute the MYB-bHLH-WD40 repeat (MBW) regulatory complexes (Figure 3). It is important to note that, to our knowledge: (a) TTG1 is common to all aforementioned traits, (b) each bHLH protein is involved in the regulation of two or more traits and (c) each MYB mostly controls only one trait (Chen & Wang, 2019; Zhang et al., 2019; Figure 3). Accordingly, the *A. thaliana ttg1* mutant lacks root hairs and trichomes, has a reduced level of anthocyanins and proanthocyanins and does not accumulate mucilage in the MSCs (Galway et al., 1994; Walker et al., 1999; Western, 2001). In wild type *A. thaliana*, TTG1 phosphorylation by SHAGGY-like kinases 11/12 (SK11:12) prevents TTG1 interaction with TRANSPARENT TESTA2/MYB123 (TT2/MYB123), a MYB member of the MBW complex, in turn decreasing the transcription of the downstream regulator *GLABRA2* (*GL2*) (Li et al., 2018; Figure 3). The consequence for seeds is promotion of lipid storage in the embryo at the expense of mucilage and flavonoid pigment production in the seed coat (Li et al., 2018; Figure 3). This differential regulation of *GL2* is probably responsible for the differential balance between seed lipid and pigment/mucilage contents in two natural *Medicago* species, which correlates with *GL2* expression level (Song et al., 2017; Figure 3). However, additional regulation mechanisms through interactions, competition, post-translational or epigenetic modifications may also occur (Nguyen, Tran, & Nguyen, 2019; Xu, Dubos, & Lepiniec, 2015). Altogether, this probably explains how a few molecular actors control several functions.

A phylogenetic analysis of the WD40 domain repeats (WDR) transcription factor family reveals that TTG1 is absent from non-seed plants and appeared in the common ancestor of angiosperms and gymnosperms (extand seed plants). TTG1 has undergone a duplication at this node and, in addition to its ancestral circadian clock function; it acquired new functions for control of epidermal cell differentiation in essentially all organs of plants (Airoldi et al., 2019, Figure 4). Interestingly, the serine 215 that can be phosphorylated by SK11/12 is conserved across seed plant TTG1 orthologs, suggesting that this ancestral regulation mechanism for this master regulator, allowing switches in carbon flow between the seed coat and the embryo, is independent of the presence of seed mucilage (Airoldi et al., 2019; Li et al., 2018).

The bHLH and MYB proteins have been subjected to numerous and recent duplication events (Doroshkov et al., 2019; Sullivan et al., 2019). Their different association modalities within several MBW complexes have probably been co-opted to control the emergence of new biological process often linked to epidermal cells such as seed mucilage establishment (Figure 4). Indeed, the *A. thaliana* MBW regulatory complex controlling seed mucilage and seed coat pigments involves at least two bHLHs (TT8-EGL3), one MYB (TT2) together with one WDR (TTG1) (reviewed in Golz et al., 2018). For the proper establishment of myxospermy, there is a need of at least

three additional MYBs (MYB5-MYB23-MYB61) (Penfield, 2001; Matsui, Hiratsu, Koyama, Tanaka, & Ohme-Takagi, 2005; Li et al., 2009, Figure 3). The functional conservation of TTG1 in mucilage production and release has been demonstrated in two Brassicaceae species *Matthiola incana* (Dressel & Hemleben, 2009) and *Arabis alpina* (Chopra et al., 2014), and in *M. truncatula* (Pang et al., 2009). Orthologs of TTG1 from *Camellia sinensis* (Liu, Hou, et al., 2018) or even from the monocotyledonous species *Setaria italica* can restore mucilage wild type phenotype of *A. thaliana ttg1* mutant through the recovery of *GL2* and *MUCILAGE MODIFIED4/RHAMNOSE BIOSYNTHESIS 2* (*MUM4*) gene expression (Liu et al., 2017). This functional conservation together with clear orthologous relationships supports the hypothesis that TTG1 was present in the common ancestor of seed plants (Figure 4). Using a similar trans-complementation approach, the functional conservation of the two bHLH proteins EGL3 and TT8 has also been demonstrated in all tested angiosperms for all MBW complex-associated traits (root hair, trichome, anthocyanin/proanthocyanin pigments and mucilage) (Zhang & Hülskamp, 2019). Additionally, the phylogeny of these two genes is well resolved at the angiosperm level, with the *Amborella trichopoda* ortholog gene branching early for each gene (Doroshkov et al., 2019, Figure 4). In *A. thaliana*, the three bHLH genes, *EGL3*, *GL3* and *MYC1* (root hair regulators) originated from an ancestral bHLH gene after a triplication within the Brassicaceae (Doroshkov et al., 2019, Figure 4). Brassicaceae, as many other plant clades, have undergone a whole-genome duplication early in their evolutionary history (Mabry et al., 2019). Thus, it will be not surprising to find several duplication events at the Brassicaceae node for other genes belonging to the MSC toolbox, especially for multigene family such as bHLHs and MYBs. Then, these duplicated genes functionally diverged as shown by the unexpected partial rescue of seed mucilage with the transcomplementation of the *A. thaliana gl3/egl3/tt8* triple mutant by *A. alpina GL3* (Zhang & Hülskamp, 2019). Since *A. thaliana* GL3 is not able to complement the mucilage phenotype in *gl3/egl3/tt8*, this indicates an intra-Brassicaceae functional divergence between both *GL3* genes since their last common ancestor. The putative new involvement of an important bHLH regulatory gene such as *GL3* in *A. alpine*, from root hair regulation to mucilage establishment, may partially explain the contrasted morphology of MSCs observed between *A. alpina* and *A. thaliana*, both belonging to Brassicaceae (Chopra et al., 2014). In addition to the rosid intra-family functional divergence, *TTG1* and *bHLH* genes show sequence and functional conservation across angiosperms (Figure 4).

Concerning *MYB* genes, *MYB5* and *TT2* ortholog genes in *M. truncatula* likely have a conserved function because they positively regulate seed coat pigment and mucilage (Liu et al., 2014). This suggests that MBW complexes dedicated to seed mucilage and seed coat pigments are conserved among the rosid species (Figure 4). However, due to the small size and the great sequence variability of *MYB* genes, in its current form, the MYB family phylogeny is not sufficiently well resolved to confirm this hypothesis (Doroshkov et al., 2019). Since the phylogeny of this multigene family is difficult to resolve and since bHLH and MYB association in MBW complexes depends on non-binary competitive interaction (Zhang et al., 2019) and on post translational modifications (Li et al., 2018), more studies will be helpful to fully characterize the evolution of the MBW complexes. Additionally, five genes highlighted by GWAS on flax mucilage content share orthologous relationship with five genes of the *A. thaliana* MSC toolbox (Soto-Cerda et al., 2018). They are *A. thaliana TT8*, confirming the previously shown ancestrality, *NAC-REGULATED SEED MORPHOLOGY 1* (*NARS1*) another non-MBW upstream regulator (Kunieda et al., 2008), and three direct actors implicated in pectin synthesis or modifications, namely *MUCILAGE-MODIFIED 4/RHAMNOSE BIOSYNTHESIS 2* (*MUM4/RMH2*), *SUBTILISIN-LIKE SERINE PROTEASE 1.7* (*SBT1.7/ARA12*) and *GALACTURONOSYLTRANSFERASE-LIKE 5* (*GATL5*) (Usadel, 2004; Rautengarten et al., 2008; Basu et al., 2015, Figure 4). These genes are implicated in myxospermy in flax and *A. thaliana* suggesting a functional conservation since their last common ancestor, the common ancestor of rosids (Figure 4).

The above summary suggests that during seed plant evolution, TTG1 first appeared to balance carbon flow in seed tissues. TTG1 progressively interacted with novel bHLH members, allowing more regulatory functions through modular complex(es) formations, with numerous and versatile recruitment of MYB members for deeper specialization to control different traits in different seed zones, such as seed mucilage in the MSCs. The myxospermy-related MBW complex and some downstream direct actors are conserved in the entire rosid clade.

Altogether, the *A. thaliana* MSC toolbox genes seem to have undergone an asymmetrical selection pressure, that is, the relative conservation of the MBW complex during angiosperm evolution contrasts with the strong and rapid changes in downstream actors observed in *A. thaliana* intra-species variability. Therefore, the more downstream is the actor in the MSC toolbox gene network, the more it should diverge between species and within species, probably explaining the astonishing morphological diversity of MSCs observed in Brassicaceae (Vaughan & Whitehouse, 1971). This could be confirmed by comparative studies of MSC toolbox genes with closely myxospermous model species such as *Camelina sativa* (Brassicaceae). More information on the ancestral state of the MSC toolbox genes in the rosid node may be obtained by the characterization of mucilage genetics in flax. In the following part, we present the newly identified putative ecological functions of myxodiaspory that can be helpful to understand why it can confer a selective advantage and to which kind of constraints.

4 | NEW INSIGHTS IN ECOLOGICAL FUNCTION OF SEED MUCILAGE

Considering that seed mucilage establishment is a costly metabolic investment for the mother plant, its presence implies that it probably displays major functions and that this trait is under a positive selection pressure in the myxodiasporous species. According to mucilage adherent and hydrophilic properties, the scientific community first investigated its influence on seed dispersal and germination (reviewed in

Western, 2012; Yang, Baskin, Baskin, & Huang, 2012), and more recently looked for potential interactions between seed mucilage and the abiotic and biotic constraints.

4.1 | Influence on seed dispersal and germination

A seed adaptation such as myxospermy is expected to influence seed dispersal and germination (Figure 5a). However, these roles can be completely different between closely related species, making it difficult to draw general conclusions valid in all myxodiasporous species. Counter-intuitively, mucilage can be a negative regulator of seed germination in *Leptocereus scopulophilus* (Barrios, Flores, González-Torres, & Palmarola, 2015) and also for the achene germination of *Artemisia monosperma* (Huang & Gutterman, 1999a). In *Blepharis persica*, the seed mucilage could block oxygen transfer under water excess and therefore prevent germination (Witztum, Gutterman, & Evenari, 1969). This function was regularly re-emphasized (last time in Gorai et al., 2014) though never fully demonstrated. However, mucilage can also improve germination (Figure 5a-1). The most cited seed mucilage function in *A. thaliana* was a positive role during germination under osmotic stress conditions, considering the polyethylene glycol (PEG)-dependent decrease of germination rate observed for *myb61*, *gl2* and *ttg1* mutants (Penfield, 2001). In another MSC toolbox downstream gene mutant, a defect in seed mucilage extrusion induces a delayed germination, suggesting a positive effect of mucilage on germination efficiency rather than on germination rate (Arsovski et al., 2009). However, more recently no such phenotypes were

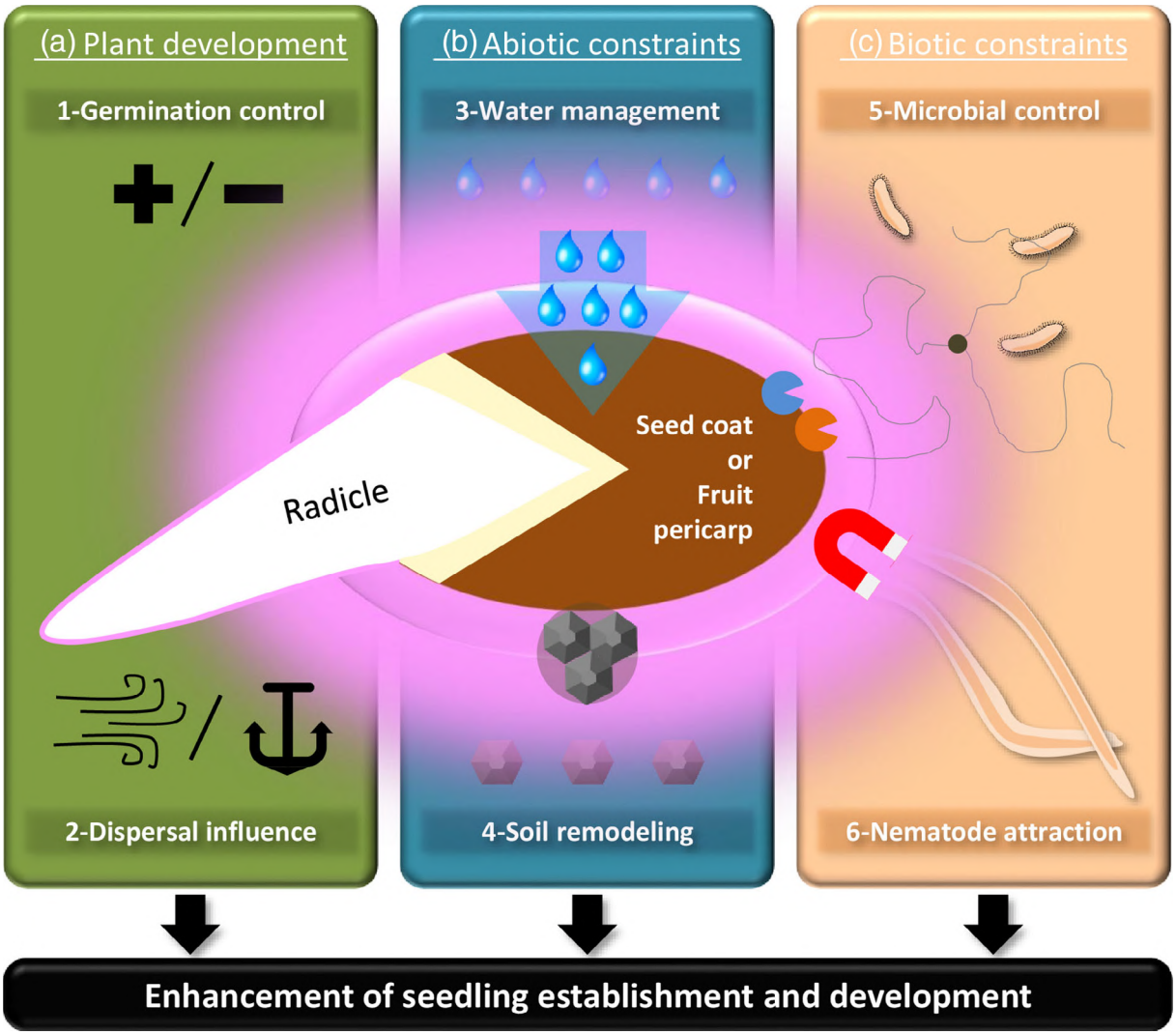


FIGURE 5 Global overview of seed (or fruit) mucilage major ecological functions facing environmental constraints. (a) Seed mucilage may influence plant development through (1) positive or negative impact on germination upon inappropriate condition, depending on the species, or (2) seed dispersal in relation to conferred seed physical properties such as sinking ability or soil anchoring. (b) Seed mucilage is modified by abiotic conditions in close environment through (3) regulation of water flux and water availability as well as (4) soil rheological remodelling properties. (c) Seed mucilage is involved in biotic interactions though (5) direct or indirect influence on microbial community establishment around the seed and the future plant and (6) attraction ability of nematodes

obtained in the *mum2* or *myb61* mutants (Saez-Aguayo et al., 2014). Thus, the role of seed mucilage in *A. thaliana* germination deserves to be deeply explored. For *S. hispanica*, the intact myxocarpous achenes germinate better than the achenes without mucilage (Geneve et al., 2017). Interestingly, seed mucilage-dependent better germination phenotypes are obtained with PEG application and not with salt at equal osmotic potential (Geneve et al., 2017). Upon five desert dwelling species (*Lavandula subnuda*, *Lepidium aucheri*, *Boerhavia elegans*, *Plantago ciliata* and *Plantago amplexicaulis*), seed mucilage presence increased water uptake but its removal led to contrasted germination effects (Bhatt, Santo, & Gallacher, 2016). Therefore, mucilage function in germination seems to be related to water uptake and/or seed permeability to water and possibly to gazes.

Some studies took advantage of the fact that two seed morphotypes occur within the same species (myxospermous and non-myxospermous seeds). These species use these dimorphic seeds to improve species persistence and dispersion (Liu, Wang, Tanveer, & Song, 2018). For three Brassicaceae species having characterized dimorphic seeds for myxospermy, namely *Diptyocharpus strictus*, *C. bursa-pastorisi and Aethionema arabicum*, the seed morphotype with the higher dormancy is not myxospermous (Arshad et al., 2019; Lu, Tan, Baskin, & Baskin, 2010; Toorop et al., 2012) suggesting that myxospermous seeds should germinate without delay. The co-occurrence of seed mucilage and wings on seeds of *D. strictus* (Lu et al., 2010) and *H. deserti* (Gorai et al., 2014) questions whether antitelochory and anemochory are opposite or can be complementary. Antitelochory prevents seed dispersion far from the mother plant while anemochory favours wind-driven dissemination under dry conditions until the seed encounters water and stops its dispersion. The combination of both traits can provide a powerful dispersal advantage until an optimal place regarding hydric conditions is reached (Figure 5a-2). Interestingly, *Lunaria annua* shows a surprising use of mucilage by secreting it from the inner surface of the fruit, keeping the four seeds attached, even after its dehiscence, this allows a differential dispersion between the two halves of the fruit and their two attached seeds by the wind (Leins, Fligge, & Erbar, 2018).

4.2 | Influence facing abiotic constraints

As seed mucilage constitutes highly hydrophilic gels, it is tempting to propose that it may provide water for the embryo. For *A. thaliana*, the seed mucilage takes a large amount of water from the environment and sequesters it through ionic linkages with galacturonic acid residues (Figure 5b-3; Saez-Aguayo et al., 2014). Indeed, mutants lacking mucilage, imbibed their seed faster than the wild type or mutants with un-released mucilage (Saez-Aguayo et al., 2014). However, mucilage allows seeds to sink faster in water as compared to non-myxospermous seeds that can float on the water surface for longer times and even germinate on it. Achenes of *Artemisia sphaerocephala* germinate and float better when their mucilage was removed (Huang & Gutterman, 1999b). Interestingly, adhesive and frictional properties of seed mucilage can change according to its hydration level influencing at least mucilage impact on dispersal properties (Kreitschitz, Kovalev, & Gorb, 2015, 2016). Upon imbibition in *A. arabicum*, thick fibres unfold from each MSCs and are able to conserve their structure and their size upon dehydration (Lenser et al., 2016). This re-dried mucilage allows a dispersal efficiency compromise, for wind, water run-off and buoyancy dissemination ways, as compared to never imbibed seeds and fully imbibed seeds (Arshad et al., 2019). A more complete understanding of the complex roles of seed mucilage in water management would necessitate investigating its role in more natural situations such as succession of wetting and drying cycles, or flooding following a long drying period.

Soil physical properties can have a major impact on water availability and root penetration (Figure 5b-4). The myxocarpy of *A. sphaerocephala* enhances seedling emergence in its sandy environment and reduces plant mortality (Yang, Baskin, Baskin, Liu, & Huang, 2012). By adding seed mucilage extract from *C. bursa pastoris*, soil rheological properties are modified particularly for hydraulic conductivity retaining water for longer time (Deng et al., 2014). A similar effect is stimulated by *S. hispanica* seed mucilage addition, which links soil particles to increase aggregate stability for at least 30 days in diverse kinds of soil (Figure 5b-4; di Marsico et al., 2018). Therefore, seed mucilage could locally improve the soil rheology in agreement with the non-disseminating lifestyle of species such as *A. thaliana* excreting non-adherent mucilage.

4.3 | Biotic interactions

Similar to root border cell mucilage (Knee et al., 2001), non-adherent seed mucilage constitutes a significant amount of polysaccharides released in the environment. Taking into consideration the importance of microorganisms in plant physiology along their development and their omnipresence around and in plant organs, this metabolic investment could indicate an involvement of seed mucilage in dealing with biotic constraints (Figure 5c). Seed mucilage influence on microbial community for the plant was shown in *A. thaliana* using the bacteria *Streptomyces lividans* that inhibits spore germination and growth of the pathogenic fungus *Verticillium dahlia*, the causal agent of verticillium wilt. When both microorganisms are co-inoculated on seeds, *S. lividans* has a better proliferation within the seed mucilage in comparison to *V. dahlia* and considerably alleviates disease symptoms (Meschke & Schrempf, 2010). Such "selective media effect" (Figure 5c-5) promoting microbial hyphae development was further illustrated with *S. hispanica* achene mucilage and *Colletotrichum graminicola* fungi (Geneve et al., 2017). The desert plant *A. sphaerocephala* achene mucilage was shown to be degraded by microorganisms, providing CO_2 and soluble sugars to promote seedling establishment (Yang, Baskin, Baskin, Zhang, & Huang, 2012). This beneficial effect was recently explained in the same species through the positive effect of mucilage on soil microbial community composition and diversity to favour fungal-bacterial interaction and soil enzyme activities, protecting young seedlings from drought and pathogens (Hu, Zhang, et al., 2019). Within the identified microbial community,

Glomeromycota are a group of fungal species well known for this positive impact on plant development through symbiotic interactions. Surprisingly, the positive fungal effect on seedling growth does not synergize with seed mucilage benefits, suggesting that mucilage and the effect of Glomeromycota on plant development act independently to enhance seedling establishment (Hu, Baskin, et al., 2019). Active stress-associated enzymes, such as nucleases, proteases and chitinases, are secreted from the seed coat of several species even in seeds several decades old (Raviv et al., 2017). However, this ability to secrete proteins is conserved in the non-myxospermous species *Raphanus sativus* or in the *A. thaliana gl2* mutant deprived of seed mucilage, suggesting that protein secretion and seed mucilage are independent (Raviv et al., 2017).

Nematodes are in close interaction with plants and can have a major pathogenic impact on plant development. *A. thaliana* seed mucilage contributes to attract root-knot nematodes (Figure 5c-6) with the additional requirement of seed-surface carbohydrates and proteins (Tsai et al., 2019). Considering the parasitic nature of nematodes, this attraction is probably due to a nematode adaptation rather than a plant adaptation. *C. bursa-pastoris*, a closely relative species of *A. thaliana*, has myxospermous seeds that are also able to attract nematodes (Roberts, Warren, & Provan, 2018). Surprisingly, it seems to be a case of protocarnivory because the massive death of trapped nematodes in seed mucilage increases plant development from germination to young seedling establishment, especially under low nutrient level (Roberts et al., 2018). Thus, seed mucilage involvement in biotic interactions is revealing astonishing functions that could affect plant development in unexpected manners as compared to the previously characterized functions.

5 | CONCLUSIONS AND PERSPECTIVES

Myxodiaspory is a macroscopic seed trait that results from a surprising diversity of microscopic features among angiosperms, at both the intra-species and inter-species levels. This striking microscopic morphological diversity makes it difficult to trace back the evolutionary origin of seed mucilage based only on morphology. The *A. thaliana* MSC toolbox genes have been increasingly characterized in recent years. The comparison of these molecular actors within a species or between species has started to shed light on this intriguing evolutionary story. On the one hand, the intense selection pressure for mucilage establishment appears to be mainly applied on a few downstream genes of the MSC toolbox as illustrated in non-myxospermous *A. thaliana* natural populations. On the other hand, the seed mucilage-related MBW upstream master regulatory complex appeared sequentially during seed plant evolution and is notably conserved across rosids. From this regulatory complex, mucilage may have evolved several times independently as a combination of highly diverse traits including two MSC epidermal origins, various patterns of MSC cell wall dynamics, various mucilage polysaccharidic composition and sublayering patterns or several modes of mucilage release. In turn, this diversity of traits probably contributes to the wide range of ecological functions observed in each species that face contrasting environments. The biotic and abiotic constraints are the least studied areas in this field and probably the most promising tracks to uncover new seed mucilage functions related to particular environments.

ACKNOWLEDGMENTS

The authors are thankful to Université Paul Sabatier-Toulouse III (France) and CNRS for supporting their research work. S. V. benefited from a PhD scholarship funded by the University Paul Sabatier-Toulouse III. This work was also supported by the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02) and the French National Research Agency project "MicroWall" (ANR-18-CE20-0007). We would like to thank Charles Uhlmann and Ali Eljebbawi for improving the English language.

CONFLICT OF INTEREST

The authors declare that the work was conducted in the absence of any commercial or financial relationship that could be constructed as a potential conflict of interest.

ORCID

Vincent Burlat https://orcid.org/0000-0002-0897-6011
Christophe Dunand https://orcid.org/0000-0003-1637-4042

REFERENCES

Airoldi, C. A., Hearn, T. J., Brockington, S. F., Webb, A. A. R., & Glover, B. J. (2019). TTG1 proteins regulate circadian activity as well as epidermal cell fate and pigmentation. *Nature Plants*, *5*, 1145–1153.

Arshad, W., Sperber, K., Steinbrecher, T., Nichols, B., Jansen, V. A. A., Leubner-Metzger, G., & Mummenhoff, K. (2019). Dispersal biophysics and adaptive significance of dimorphic diaspores in the annual *Aethionema arabicum* (Brassicaceae). *New Phytologist*, *221*, 1434–1446.

Arsovski, A. a., Haughn, G. W., & Western, T. L. (2010). As a model for plant cell wall research. *Plant Signaling & Behavior*, *5*, 796–801.

Arsovski, A. A., Popma, T. M., Haughn, G. W., Carpita, N. C., McCann, M. C., & Western, T. L. (2009). AtBXL1 encodes a bifunctional -D-xylosidase/-L-arabinofuranosidase required for pectic arabinan modification in *Arabidopsis* mucilage secretory cells. *Plant Physiology*, *150*, 1219–1234.

Baroux, C., & Grossniklaus, U. (2019). Seeds—An evolutionary innovation underlying reproductive success in flowering plants. *Current Topics in Developmental Biology*, *131*, 605–642.

Barrios, D., Flores, J., González-Torres, L. R., & Palmarola, A. (2015). The role of mucilage in the germination of *Leptocereus scopulophilus* (Cactaceae) seeds from Pan de Matanzas, Cuba. *Botany*, *93*, 251–255.

Basu, D., Wang, W., Ma, S., DeBrosse, T., Poirier, E., Emch, K., … Showalter, A. M. (2015). Two hydroxyproline galactosyltransferases, GALT5 and GALT2, function in arabinogalactan-protein glycosylation, growth and development in *Arabidopsis*. *PLoS One*, *10*, e0125624.

Bhatt, A., Santo, A., & Gallacher, D. (2016). Seed mucilage effect on water uptake and germination in five species from the hyper-arid Arabian desert. *Journal of Arid Environments*, *128*, 73–79.

Chen, S., & Wang, S. (2019). GLABRA2, a common regulator for epidermal cell fate determination and anthocyanin biosynthesis in *Arabidopsis*. *International Journal of Molecular Sciences*, *20*, 4997.

Chopra, D., Wolff, H., Span, J., Schellmann, S., Coupland, G., Albani, M. C., … Hülskamp, M. (2014). Analysis of TTG1 function in *Arabis alpina*. *BMC Plant Biology*, *14*, 16.

Dean, G. H., Zheng, H., Tewari, J., Huang, J., Young, D. S., Hwang, Y. T., … Haughn, G. W. (2007). The *Arabidopsis* MUM2 gene encodes a β-galactosidase required for the production of seed coat mucilage with correct hydration properties. *The Plant Cell*, *19*, 4007–4021.

Deng, W., Hallett, P. D., Jeng, D. S., Squire, G. R., Toorop, P. E., & Iannetta, P. P. M. (2014). The effect of natural seed coatings of *Capsella bursa-pastoris* L. Medik. (shepherd's purse) on soil-water retention, stability and hydraulic conductivity. *Plant and Soil*, *387*, 167–176.

Deng, W., Jeng, D. S., Toorop, P. E., Squire, G. R., & Iannetta, P. P. M. (2012). A mathematical model of mucilage expansion in myxospermous seeds of *Capsella bursa-pastoris* (shepherds purse). *Annals of Botany*, *109*, 419–427.

di Marsico, A., Scrano, L., Labella, R., Lanzotti, V., Rossi, R., Cox, L., … Amato, M. (2018). Mucilage from fruits/seeds of chia (*Salvia hispanica* L.) improves soil aggregate stability. *Plant and Soil*, *425*, 57–69.

Doroshkov, A. V., Konstantinov, D. K., Afonnikov, D. A., & Gunbin, K. V. (2019). The evolution of gene regulatory networks controlling *Arabidopsis thaliana* L. trichome development. *BMC Plant Biology*, *19*, 53.

Dressel, A., & Hemleben, V. (2009). Transparent Testa Glabra 1 (TTG1) and TTG1-like genes in *Matthiola incana* R. Br. and related Brassicaceae and mutation in the WD-40 motif. *Plant Biology*, *11*, 204–212.

Fabrissin, I., Cueff, G., Berger, A., Granier, F., Sallé, C., Poulain, D., … North, H. M. (2019). Natural variation reveals a key role for rhamnogalacturonan I in seed outer mucilage and underlying genes. *Plant Physiology*, *181*, 1498–1518.

Francoz, E., Ranocha, P., Burlat, V., & Dunand, C. (2015). *Arabidopsis* seed mucilage secretory cells: Regulation and dynamics. *Trends in Plant Science*, *20*, 515–524.

Francoz, E., Ranocha, P., le Ru, A., Martinez, Y., Fourquaux, I., Jauneau, A., … Burlat, V. (2019). Pectin demethylesterification generates platforms that anchor peroxidases to remodel plant cell wall domains. *Developmental Cell*, *48*, 261–276.

Galloway, A. F., Knox, P., & Krause, K. (2020). Sticky mucilages and exudates of plants: Putative microenvironmental design elements with biotechnological value. *New Phytologist*, *225*, 1461–1469.

Galway, M. E., Masucci, J. D., Lloyd, A. M., Walbot, V., Davis, R. W., & Schiefelbein, J. W. (1994). The TTG gene is required to specify epidermal cell fate and cell patterning in the *Arabidopsis* root. *Developmental Biology*, *166*, 740–754.

Geneve, R. L., Hildebrand, D. F., Phillips, T. D., Al-Amery, M., & Kester, S. T. (2017). Stress influences seed germination in mucilage-producing chia. *Crop Science*, *57*, 2160–2169.

Golz, J. F., Allen, P. J., Li, S. F., Parish, R. W., Jayawardana, N. U., Bacic, A., & Doblin, M. S. (2018). Layers of regulation—Insights into the role of transcription factors controlling mucilage production in the *Arabidopsis* seed coat. *Plant Science*, *272*, 179–192.

Gorai, M., el Aloui, W., Yang, X., & Neffati, M. (2014). Toward understanding the ecological role of mucilage in seed germination of a desert shrub *Henophyton deserti*: Interactive effects of temperature, salinity and osmotic stress. *Plant and Soil*, *374*, 727–738.

Griffiths, J. S., & North, H. M. (2017). Sticking to cellulose: Exploiting *Arabidopsis* seed coat mucilage to understand cellulose biosynthesis and cell wall polysaccharide interactions. *New Phytologist*, *214*, 959–966.

Hu, D., Baskin, J. M., Baskin, C. C., Wang, Z., Zhang, S., Yang, X., & Huang, Z. (2019). Arbuscular mycorrhizal symbiosis and achene mucilage have independent functions in seedling growth of a desert shrub. *Journal of Plant Physiology*, *232*, 1–11.

Hu, D., Zhang, S., Baskin, J. M., Baskin, C. C., Wang, Z., Liu, R., … Huang, Z. (2019). Seed mucilage interacts with soil microbial community and physiochemical processes to affect seedling emergence on desert sand dunes. *Plant Cell and Environment*, *42*, 591–605.

Huang, D., Wang, C., Yuan, J., Cao, J., & Lan, H. (2015). Differentiation of the seed coat and composition of the mucilage of *Lepidium perfoliatum* L.: A desert annual with typical myxospermy. *Acta Biochimica et Biophysica Sinica*, *47*, 775–787.

Huang, Z., & Gutterman, Y. (1999a). Water absorption by mucilaginous achenes of *Artemisia monosperma*: Floating and germination as affected by salt concentrations. *Israel Journal of Plant Sciences*, *47*, 27–34.

Huang, Z., & Gutterman, Y. (1999b). Germination of *Artemisia sphaerocephala* (Asteraceae), occurring in the sandy desert areas of Northwest China. *South African Journal of Botany*, *65*, 187–196.

Jones, V. A. S., & Dolan, L. (2012). The evolution of root hairs and rhizoids. *Annals of Botany*, *110*, 205–212.

Knee, E. M., Gong, F. C., Gao, M., Teplitski, M., Jones, A. R., Foxworthy, A., … Bauer, W. D. (2001). Root mucilage from pea and its utilization by rhizosphere bacteria as a sole carbon source. *Molecular Plant-Microbe Interactions*, *14*, 775–784.

Kreitschitz, A., & Gorb, S. N. (2017). How does the cell wall 'stick' in the mucilage? A detailed microstructural analysis of the seed coat mucilaginous cell wall. *Flora: Morphology, Distribution. Functional Ecology of Plants*, *229*, 9–22.

Kreitschitz, A., & Gorb, S. N. (2018). The micro- and nanoscale spatial architecture of the seed mucilage—Comparative study of selected plant species. *PLoS One*, *13*, e0200522.

Kreitschitz, A., Kovalev, A., & Gorb, S. N. (2015). Slipping vs sticking: Water-dependent adhesive and frictional properties of *Linum usitatissimum* L. seed mucilaginous envelope and its biological significance. *Acta Biomaterialia*, *17*, 152–159.

Kreitschitz, A., Kovalev, A., & Gorb, S. N. (2016). "Sticky invasion"—The physical properties of *Plantago lanceolata* L. seed mucilage. *Beilstein Journal of Nanotechnology*, *7*, 1918–1927.

Kunieda, T., Hara-Nishimura, I., Demura, T., & Haughn, G. W. (2019). *Arabidopsis* FLYING SAUCER 2 functions redundantly with FLY1 to establish normal seed coat mucilage. *Plant and Cell Physiology*, *61*, 308–317.

Kunieda, T., Mitsuda, N., Ohme-Takagi, M., Takeda, S., Aida, M., Tasaka, M., … Hara-Nishimura, I. (2008). NAC family proteins NARS1/NAC2 and NARS2/NAM in the outer integument regulate embryogenesis in *Arabidopsis*. *The Plant Cell*, *20*, 2631–2642.

Kunieda, T., Shimada, T., Kondo, M., Nishimura, M., Nishitani, K., & Hara-Nishimura, I. (2013). Spatiotemporal secretion of PEROXIDASE36 is required for seed coat mucilage extrusion in *Arabidopsis*. *The Plant Cell*, *25*, 1355–1367.

Leins, P., Fligge, K., & Erbar, C. (2018). Silique valves as sails in anemochory of Lunaria (Brassicaceae). *Plant Biology*, *20*, 238–243.

Lenser, T., Graeber, K., Cevik, Ö. S., Adigüzel, N., Dönmez, A. A., Grosche, C., … Leubner-Metzger, G. (2016). Developmental control and plasticity of fruit and seed dimorphism in *Aethionema arabicum*. *Plant Physiology*, *172*, 1691–1707.

Li, C., Zhang, B., Chen, B., Ji, L., & Yu, H. (2018). Site-specific phosphorylation of TRANSPARENT TESTA GLABRA1 mediates carbon partitioning in *Arabidopsis* seeds. *Nature Communications*, *9*, 571.

Li, S. F., Milliken, O. N., Pham, H., Seyit, R., Napoli, R., Preston, J., … Parish, R. W. (2009). The *Arabidopsis* MYB5 transcription factor regulates mucilage synthesis, seed coat development, and trichome morphogenesis. *The Plant Cell*, *21*, 72–89.

Liu, C., Jun, J. H., & Dixon, R. A. (2014). MYB5 and MYB14 play pivotal roles in seed coat polymer biosynthesis in *Medicago truncatula*. *Plant Physiology*, *165*, 1424–1439.

Liu, J., Shim, Y. Y., Shen, J., Wang, Y., Ghosh, S., & Reaney, M. J. T. (2016). Variation of composition and functional properties of gum from six Canadian flaxseed (*Linum usitatissimum* L.) cultivars. *International Journal of Food Science and Technology*, *51*, 2313–2326.

Liu, K., Qi, S., Li, D., Jin, C., Gao, C., Duan, S., … Chen, M. (2017). TRANSPARENT TESTA GLABRA 1 ubiquitously regulates plant growth and development from *Arabidopsis* to foxtail millet (*Setaria italica*). *Plant Science*, *254*, 60–69.

Liu, R., Wang, L., Tanveer, M., & Song, J. (2018). Seed heteromorphism: An important adaptation of halophytes for habitat heterogeneity. *Frontiers in Plant Science*, *9*, 1515.

Liu, Y., Hou, H., Jiang, X., Wang, P., Dai, X., Chen, W., … Xia, T. (2018). A WD40 repeat protein from *Camellia sinensis* regulates anthocyanin and proanthocyanidin accumulation through the formation of MYB–bHLH–WD40 ternary complexes. *International Journal of Molecular Sciences*, *19*, 1686.

Lloyd, A., Brockman, A., Aguirre, L., Campbell, A., Bean, A., Cantero, A., & Gonzalez, A. (2017). Advances in the MYB-bHLH-WD repeat (MBW) pigment regulatory model: Addition of a WRKY factor and co-option of an anthocyanin MYB for betalain regulation. *Plant and Cell Physiology*, *58*, 1431–1441.

Lu, J., Tan, D., Baskin, J. M., & Baskin, C. C. (2010). Fruit and seed heteromorphism in the cold desert annual ephemeral *Diptychocarpus strictus* (Brassicaceae) and possible adaptive significance. *Annals of Botany*, *105*, 999–1014.

Mabry, M., Brose, J., Blischak, P., Sutherland, B., Dismukes, W., Bottoms, C., … Pires, C. (2019). Phylogeny and multiple independent whole-genome duplication events in the Brassicales. *bioRxiv Preprint*.

Macquet, A., Ralet, M. C., Kronenberger, J., Marion-Poll, A., & North, H. M. (2007). In situ, chemical and macromolecular study of the composition of *Arabidopsis thaliana* seed coat mucilage. *Plant and Cell Physiology*, *48*, 984–999.

Macquet, A., Ralet, M.-C., Loudet, O., Kronenberger, J., Mouille, G., Marion-Poll, A., & North, H. M. (2007). A naturally occurring mutation in an *Arabidopsis* accession affects a β-d-galactosidase that increases the hydrophilic potential of rhamnogalacturonan I in seed mucilage. *The Plant Cell*, *19*, 3990–4006.

Matsui, K., Hiratsu, K., Koyama, T., Tanaka, H., & Ohme-Takagi, M. (2005). A chimeric AtMYB23 repressor induces hairy roots, elongation of leaves and stems, and inhibition of the deposition of mucilage on seed coats in *Arabidopsis*. *Plant and Cell Physiology*, *46*, 147–155.

Meschke, H., & Schrempf, H. (2010). *Streptomyces lividans* inhibits the proliferation of the fungus *Verticillium dahliae* on seeds and roots of *Arabidopsis thaliana*. *Microbial Biotechnology*, *3*, 428–443.

Miart, F., Fournet, F., Dubrulle, N., Petit, E., Demailly, H., Dupont, L., … Pageau, K. (2019). Cytological approaches combined with chemical analysis reveals the layered nature of flax mucilage. *Frontiers in Plant Science*, *10*, 684.

Mirhosseini, H., & Amid, B. T. (2012). A review study on chemical composition and molecular structure of newly plant gum exudates and seed gums. *Food Research International*, *46*, 387–398.

Nguyen, C. T., Tran, G. B., & Nguyen, N. H. (2019). The MYB–bHLH–WDR interferers (MBWi) epigenetically suppress the MBW's targets. *Biology of the Cell*, *111*, 284–291.

North, H. M., Berger, A., Saez-Aguayo, S., & Ralet, M. C. (2014). Understanding polysaccharide production and properties using seed coat mutants: Future perspectives for the exploitation of natural variants. *Annals of Botany*, *114*, 1251–1263.

Pang, Y., Wenger, J. P., Saathoff, K., Peel, G. J., Wen, J., Huhman, D., … Dixon, R. A. (2009). A WD40 repeat protein from *Medicago truncatula* is necessary for tissue-specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development. *Plant Physiology*, *151*, 1114–1129.

Penfield, S. (2001). MYB61 is required for mucilage deposition and extrusion in the *Arabidopsis* seed coat. *The Plant Cell*, *13*, 2777–2791.

Phan, J. L., & Burton, R. A. (2018). New insights into the composition and structure of seed mucilage. In *Annual plant reviews* (pp. 1–41). Chichester, England: John Wiley & Sons.

Poulain, D., Botran, L., North, H. M., & Ralet, M.-C. (2019). Composition and physicochemical properties of outer mucilage from seeds of *Arabidopsis* natural accessions. *AoB Plants*, *11*, plz031.

Rautengarten, C., Usadel, B., Neumetzler, L., Hartmann, J., Büssis, D., & Altmann, T. (2008). A subtilisin-like serine protease essential for mucilage release from *Arabidopsis* seed coats. *Plant Journal*, *54*, 466–480.

Raviv, B., Aghajanyan, L., Granot, G., Makover, V., Frenkel, O., Gutterman, Y., & Grafi, G. (2017). The dead seed coat functions as a long-term storage for active hydrolytic enzymes. *PLoS One*, *12*, e0181102.

Renzaglia, K. S., Duff, R. J., Nickrent, D. L., & Garbary, D. J. (2000). Vegetative and reproductive innovations of early land plants: Implications for a unified phylogeny. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *355*, 769–793.

Roberts, H. R., Warren, J. M., & Provan, J. (2018). Evidence for facultative protocarnivory in *Capsella bursa-pastoris* seeds. *Scientific Reports*, *8*, 10120.

Ryding, O. (2001). Myxocarpy in the nepetoideae (Lamiaceae) with notes on myxodiaspory in general. *Systematics and Geography of Plants*, *71*, 503–514.

Saez-Aguayo, S., Ralet, M.-C., Berger, A., Botran, L., Ropartz, D., Marion-Poll, A., & North, H. M. (2013). PECTIN METHYLESTERASE INHIBITOR6 promotes *Arabidopsis* mucilage release by limiting methylesterification of homogalacturonan in seed coat epidermal cells. *The Plant Cell*, *25*, 308–323.

Saez-Aguayo, S., Rondeau-Mouro, C., Macquet, A., Kronholm, I., Ralet, M. C., Berger, A., … North, H. M. (2014). Local evolution of seed flotation in *Arabidopsis*. *PLoS Genetics*, *10*, e1004221.

Shimada, T., Kunieda, T., Sumi, S., Koumoto, Y., Tamura, K., Hatano, K., … Hara-Nishimura, I. (2018). The AP-1 complex is required for proper mucilage formation in *Arabidopsis* seeds. *Plant & Cell Physiology*, *59*, 2331–2338.

Šola, K., Dean, G. H., & Haughn, G. W. (2019). *Arabidopsis* seed mucilage: A specialised extracellular matrix that demonstrates the structure–function versatility of cell wall polysaccharides. In *Annual plant reviews online*, 2, (pp. 1085–1116). New Jersey, USA: Wiley Online Library.

Šola, K., Gilchrist, E. J., Ropartz, D., Wang, L., Feussner, I., Mansfield, S. D., … Haughn, G. W. (2019). RUBY, a putative galactose oxidase, influences pectin properties and promotes cell-to-cell adhesion in the seed coat epidermis of *Arabidopsis*. *The Plant Cell*, *31*, 809–831.

Song, Y., He, L., Wang, X.-D., Smith, N., Wheeler, S., Garg, M. L., & Rose, R. J. (2017). Regulation of carbon partitioning in the seed of the model legume Medicago truncatula and *Medicago orbicularis*: A comparative approach. *Frontiers in Plant Science*, *8*, 2070.

Soto-Cerda, B. J., Cloutier, S., Quian, R., Gajardo, H. A., Olivos, M., & You, F. M. (2018). Genome-wide association analysis of mucilage and hull content in flax (*Linum usitatissimum* l.) seeds. *International Journal of Molecular Sciences*, *19*, 2870.

Soukoulis, C., Gaiani, C., & Hoffmann, L. (2018). Plant seed mucilage as emerging biopolymer in food industry applications. *Current Opinion in Food Science*, *22*, 28–42.

Sullivan, A. M., Arsovski, A. A., Thompson, A., Sandstrom, R., Thurman, R. E., Neph, S., … Queitsch, C. (2019). Mapping and dynamics of regulatory DNA in maturing *Arabidopsis thaliana* siliques. *Frontiers in Plant Science*, *10*, 1434.

Takenaka, Y., Kato, K., Ogawa-Ohnishi, M., Tsuruhama, K., Kajiura, H., Yagyu, K., … Ishimizu, T. (2018). Pectin RG-I rhamnosyltransferases represent a novel plant-specific glycosyltransferase family. *Nature Plants*, *4*, 669–676.

Toorop, P. E., Campos Cuerva, R., Begg, G. S., Locardi, B., Squire, G. R., & Iannetta, P. P. M. (2012). Co-adaptation of seed dormancy and flowering time in the arable weed *Capsella bursa-pastoris* (shepherds purse). *Annals of Botany*, *109*, 481–489.

Tsai, A. Y. L., Higaki, T., Nguyen, C. N., Perfus-Barbeoch, L., Favery, B., & Sawa, S. (2019). Regulation of root-knot nematode behavior by seed-coat mucilage-derived attractants. *Molecular Plant*, *12*, 99–112.

Tucker, M. R., Ma, C., Phan, J., Neumann, K., Shirley, N. J., Hahn, M. G., … Burton, R. A. (2017). Dissecting the genetic basis for seed coat

mucilage heteroxylan biosynthesis in *Plantago ovata* using gamma irradiation and infrared spectroscopy. *Frontiers in Plant Science*, *8*, 328.

Usadel, B. (2004). RHM2 is involved in mucilage pectin synthesis and is required for the development of the seed coat in *Arabidopsis*. *Plant Physiology*, *134*, 286–295.

van Wijk, R., Zhang, Q., Zarza, X., Lamers, M., Marquez, F. R., Guardia, A., … Munnik, T. (2018). Role for *Arabidopsis* PLC7 in stomatal movement, seed mucilage attachment, and leaf serration. *Frontiers in Plant Science*, *9*, 1721.

Vaughan, J. G., & Whitehouse, J. M. (1971). Seed structure and the taxonomy of the Cruciferae. *Botanical Journal of the Linnean Society*, *64*, 383–409.

Voiniciuc, C., Engle, K. A., Günl, M., Dieluweit, S., Schmidt, M. H.-W., Yang, J.-Y., … Usadel, B. (2018). Identification of key enzymes for pectin synthesis in seed mucilage. *Plant Physiology*, *178*, 1045–1064.

Voiniciuc, C., Yang, B., Schmidt, M. H. W., Günl, M., & Usadel, B. (2015). Starting to gel: How *Arabidopsis* seed coat epidermal cells produce specialized secondary cell walls. *International Journal of Molecular Sciences*, *16*, 3452–3473.

Voiniciuc, C., Zimmermann, E., Schmidt, M. H.-W., Günl, M., Fu, L., North, H. M., & Usadel, B. (2016). Extensive natural variation in *Arabidopsis* seed mucilage structure. *Frontiers in Plant Science*, *7*, 803.

Walker, A. R., Davison, P. A., Bolognesi-Winfield, A. C., James, C. M., Srinivasan, N., Blundell, T. L., … Gray, J. C. (1999). The TRANSPARENT TESTA GLABRA1 locus, which regulates trichome differentiation and anthocyanin biosynthesis in *Arabidopsis*, encodes a WD40 repeat protein. *The Plant Cell*, *11*, 1337–1349.

Wang, M., Xu, Z., Ahmed, R. I., Wang, Y., Hu, R., Zhou, G., & Kong, Y. (2019). Tubby-like protein 2 regulates homogalacturonan biosynthesis in *Arabidopsis* seed coat mucilage. *Plant Molecular Biology*, *99*, 421–436.

Weitbrecht, K., Müller, K., & Leubner-Metzger, G. (2011). First off the mark: Early seed germination. *Journal of Experimental Botany*, *62*, 3289–3309.

Western, T. L. (2001). Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in *Arabidopsis*. *Plant Physiology*, *127*, 998–1011.

Western, T. L. (2012). The sticky tale of seed coat mucilages: Production, genetics, and role in seed germination and dispersal. *Seed Science Research*, *22*, 1–25.

Witztum, A., Gutterman, Y., & Evenari, M. (1969). Integumentary mucilage as an oxygen barrier during germination of *Blepharis persica*. *Botanical Gazette*, *130*, 238–241.

Xu, W., Dubos, C., & Lepiniec, L. (2015). Transcriptional control of flavonoid biosynthesis by MYB-bHLH-WDR complexes. *Trends in Plant Science*, *20*, 176–185.

Yang, B., Voiniciuc, C., Fu, L., Dieluweit, S., Klose, H., & Usadel, B. (2019). TRM4 is essential for cellulose deposition in *Arabidopsis* seed mucilage by maintaining cortical microtubule organization and interacting with CESA3. *New Phytologist*, *221*, 881–895.

Yang, X., Baskin, C. C., Baskin, J. M., Liu, G., & Huang, Z. (2012). Seed mucilage improves seedling emergence of a sand desert shrub. *PLoS One*, *7*, e34897.

Yang, X., Baskin, C. C., Baskin, J. M., Zhang, W., & Huang, Z. (2012). Degradation of seed mucilage by soil microflora promotes early seedling growth of a desert sand dune plant. *Plant, Cell and Environment*, *35*, 872–883.

Yang, X., Baskin, J. M., Baskin, C. C., & Huang, Z. (2012). More than just a coating: Ecological importance, taxonomic occurrence and phylogenetic relationships of seed coat mucilage. *Perspectives in Plant Ecology, Evolution and Systematics*, *14*, 434–442.

Yu, L., Lyczakowski, J. J., Pereira, C. S., Kotake, T., Yu, X., Li, A., … Dupree, P. (2018). The patterned structure of galactoglucomannan suggests it may bind to cellulose in seed mucilage. *Plant Physiology*, *178*, 1011–1026.

Yu, L., Yakubov, G. E., Zeng, W., Xing, X., Stenson, J., Bulone, V., & Stokes, J. R. (2017). Multi-layer mucilage of *Plantago ovata* seeds: Rheological differences arise from variations in arabinoxylan side chains. *Carbohydrate Polymers*, *165*, 132–141.

Zhang, B., Chopra, D., Schrader, A., & Hülskamp, M. (2019). Evolutionary comparison of competitive protein-complex formation of MYB, bHLH, and WDR proteins in plants. *Journal of Experimental Botany*, *70*, 3197–3209.

Zhang, B., & Hülskamp, M. (2019). Evolutionary analysis of MBW function by phenotypic Rescue in *Arabidopsis thaliana*. *Frontiers in Plant Science*, *10*, 375.

How to cite this article: Viudes S, Burlat V, Dunand C. Seed mucilage evolution: Diverse molecular mechanisms generate versatile ecological functions for particular environments. *Plant Cell Environ*. 2020;43:2857–2870. https://doi.org/10.1111/pce.13827

1.3.3. La structuration du mucilage des graines

Chez la plupart des espèces myxospermiques, le mucilage (ou au moins une partie) reste solidement accroché à la surface de la graine. Chez *A. thaliana* la structuration du mucilage des graines est principalement assurée par les microfibrilles de cellulose malgré leur faible proportion relative en comparaison des pectines. Ceci est appuyé par la perte d'adhérence du mucilage chez les mutants pertes de fonctions des celluloses synthases qui sont responsables de la synthèse de cellulose au niveau du mucilage (Mendu et al., 2011). Dans les parois cellulaires végétales il existe également des protéines qui participent à la structuration des polysaccharides (cf. partie 1.2.3), et il pourrait en être de même dans le mucilage des graines. C'est un champ de recherche assez récent où le peu de données disponibles sont concentrées chez *A. thaliana*. Un séquençage des protéines présentes dans son mucilage en a dénombré 28 dont 4 identifiées précédemment par des approches génétiques (Tsai et al., 2017). Une protéine qui n'a pas été détectée dans l'étude précédemment citée, une arabinogalactane protéine (AGP) nommé SALT-OVERLY SENSITIVE 5 (SOS5), participe à l'adhérence du mucilage mais d'une manière indépendante des celluloses synthases (Griffiths et al., 2014). Elle participerait à l'interaction entre polysaccharides, probablement via les chaînes latérales de galactanes des RGI (Basu et al., 2015), pour permettre le bon déploiement des microfibrilles de cellulose, et ainsi consolider la structuration du mucilage adhérent (Griffiths et al., 2016). A noter que SOS5 fonctionne dans le mucilage de pair avec FEI2, un recepteur-like kinase qui participe à sa localisation (Griffiths et al., 2016), et avec au moins deux des galactosyltransferases capables de glycosyler SOS5 (GALT2 et GALT5) (Basu et al., 2016). Chez d'autres espèces, le mucilage des graines présente une grande diversité de micro et nanostructures (Kreitschitz et Gorb, 2018), on peut donc également s'attendre chez ces mêmes espèces à des mécanismes moléculaires différents de ceux d'*A. thaliana*.

1.3.4. Les fonctions écologiques non abordées

Depuis la publication de la revue contenue dans la partie précédente, quelques nouvelles avancées ont été faites concernant les rôles écologiques des mucilages au niveau de la dispersion des graines. Le mucilage des graines permet aux graines d'adhérer au sol et ainsi d'empêcher leur récolte par les fourmis pour une majorité d'espèce myxospermiques (Pan et al., 2021). Sans surprise, plus le mucilage est abondant plus il faut de force pour déloger la graine mais étonnamment sa masse ne rentre pas en compte. La diminution de la prédation des graines par les fourmis aurait pu être due à une sorte de camouflage procuré par les particules de sol collé au mucilage tout autour de la graine, mais c'est bien l'adhésion au sol qui en est responsable (LoPresti et al., 2019). La myxospermie peut aussi faciliter le passage à travers le système digestif des oiseaux (pigeons en l'occurrence) et permettre de conserver une viabilité partielle provocant ainsi une dissémination par leur intermédiaire (Kreitschitz et al., 2021). La composition du mucilage semble être liée à cette fonction car les graines de Plantago (mucilage composé à majorité d'hemicelluloses) ont gardé bien plus de viabilité que celle de Lin (mucilage majoritairement pectique) et aucune de celles du Basilic (majoritairement

cellulosique) n'a survécu à la digestion. Ce genre d'approches, large échelle pour les fourmis (53 espèces réparties dans 13 familles) et comparaison entre espèces au mucilage bien caractérisé, sont prometteuses car elles permettent de voir un peu plus clair à travers l'incroyable diversité de mucilages. Des effets composition-dépendant comme celui montré pour les pigeons pourrait être la clef du mystère des rôles écologiques parfois espèces-dépendant décrit plus tôt (§ 1.3.2.).

1.3.5. Les traces fossiles de la myxospermie

La découverte de fossiles permet d'avoir un aperçu de la morphologie d'un représentant d'une espèce aujourd'hui disparue, soit car elle s'est complètement éteinte ou bien que son évolution et les divergences l'ont progressivement transformée en une espèce actuelle ou une multitude d'entre elles. Les fossiles permettent donc de connaitre précisément la morphologie d'un ancêtre commun s'il peut être relié à ses descendants actuels. Ce sont donc des arguments forts en faveur de l'ancestralité d'un trait morphologique qui permettent en plus de le dater relativement précisément. Malheureusement le mucilage des graines ne comporte pas de parties dures, il est donc peu propice à la fossilisation contrairement aux os, carapaces ou bois. Une graine sèche est en revanche suffisamment dure et déshydratée pour être fossilisée mais si le mucilage n'a pas été imbibé il est compliqué d'être certain de sa nature. Néanmoins dans les analyses histologiques de graines fossiles présentes dans la littérature scientifique, quelques études anciennes font mention de cellules mucilagineuses dans le testa (**Figure 12**). Deux d'entre elles identifient des cellules à mucilage au niveau de la couche cellulaire épidermique de graines datant du Paléozoïque (**Figure 12**, (Arber, 1910; Grove et Rothwell, 1980)). La plus ancienne appuie leur nature mucilagineuse par des observations de changement de tailles des cellules (**Figure 12B et 12C**). La plus récente note leur remplissage par une matière de couleur ambrée (**Figure 12D**). Dans les deux études ce sont des graines de deux espèces affiliées au genre *Mitrospermum* qui serait des Cordaitales, un groupe aujourd'hui éteint, frère des Conifères. Les graines fossiles du groupe Conostoma montrent des cellules épidermiques qui semblent provoquer un « souffle » décollant ce qui se trouve au-dessus d'elles attribué à leur potentiel contenu en mucilage (Olivier et Salisbury, 1911). Les graines fossiles de *Physostoma stellatum* montrent un phénomène similaire où l'on voit relativement bien la couche cellulaire épidermique se scinder en deux au niveau des parois radiales, laissant la totalité des parois externes flotter au-dessus du reste de la graine (**Figure 13**, (Holden, 1954)). Conostoma, et Physostoma sont des genres appartenant au Pteridospermes (fougères à graines), groupe présent au carbonifère mais aujourd'hui disparu. Pour des graines fossiles du genre Diplotesta, de grandes cellules de la couche externe du testa semblant être remplies de mucilage sont décrites comme dépourvues de noyau et de cytoplasme (Bertrand, 1907). Tous ces arguments (localisations épidermique, contenu cellulaire particulier, rupture pariétale accompagnée de l'éloignement de la partie externe, cellule complètement différenciée en paroi modifiée) attestent que des cellules à/sécrétrices de mucilage ont existé très tôt dans l'histoire des graines.

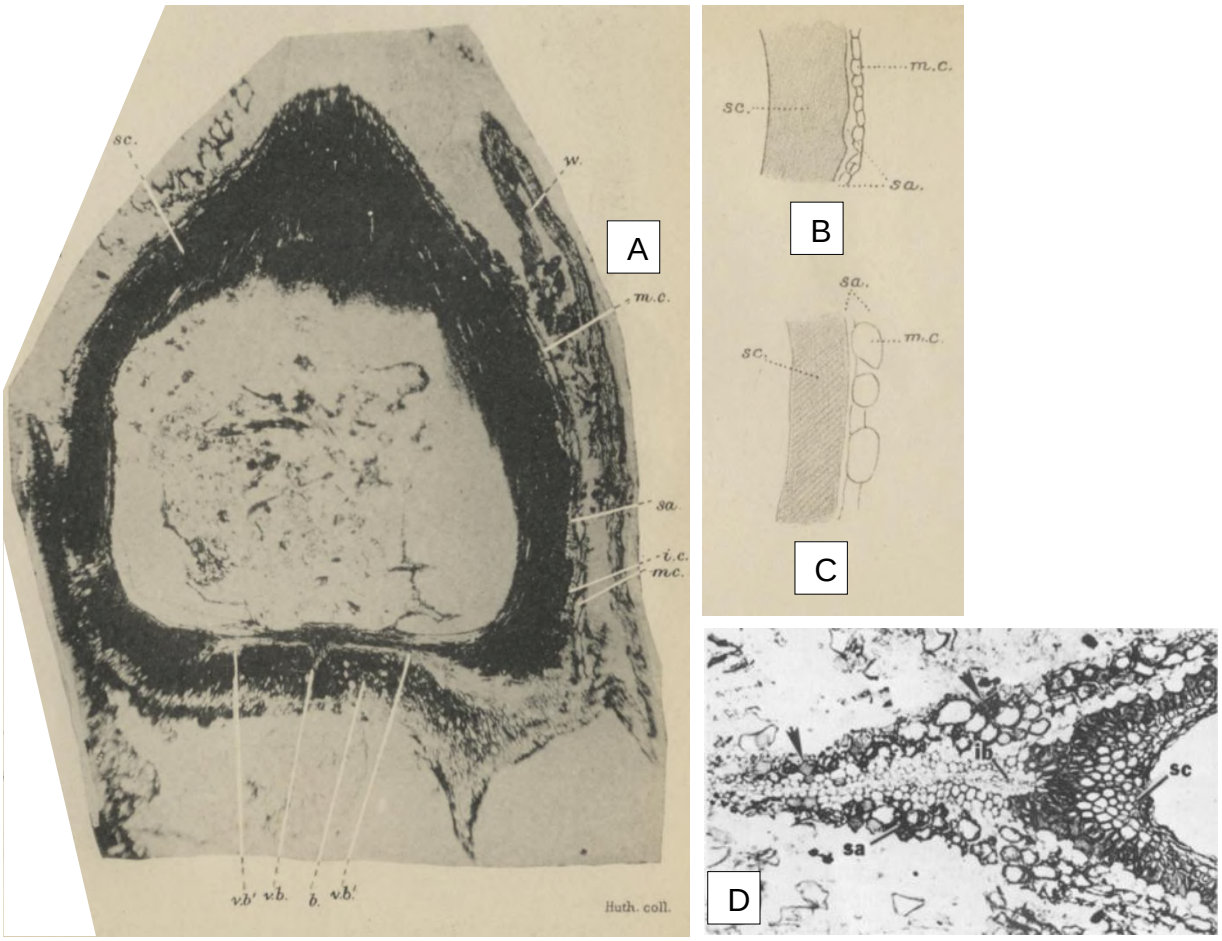


Figure 12 : Coupes de graine fossiles du genre Mitrospermum extraites de Arber, 1910 (A, B, C) et de Grove et Rothwell, 1980 (D). **(A)** Coupe longitudinale dans le plan de l'aplatissement de la graine. Les cellules à mucilage de l'épiderme du sarcotesta sont dessinées dans leur état non gonflé **(B)** ou gonflé à l'extrême **(C)** pour montrer leur taille relative au sclerotesta. Abréviations : (mc) Cellules a mucilage, (ie) plus petites cellules internes, (sa) sarcotesta, (sc) sclerotesta, (vb) faisceau vasculaire, (w) aile. **(D)** Coupe transversale du sclerotesta, du sarcotesta, et des faisceaux tégumentaires (ib). Les cellules à mucilage remplies d'une substance ambrée sont indiquées par des flèches.

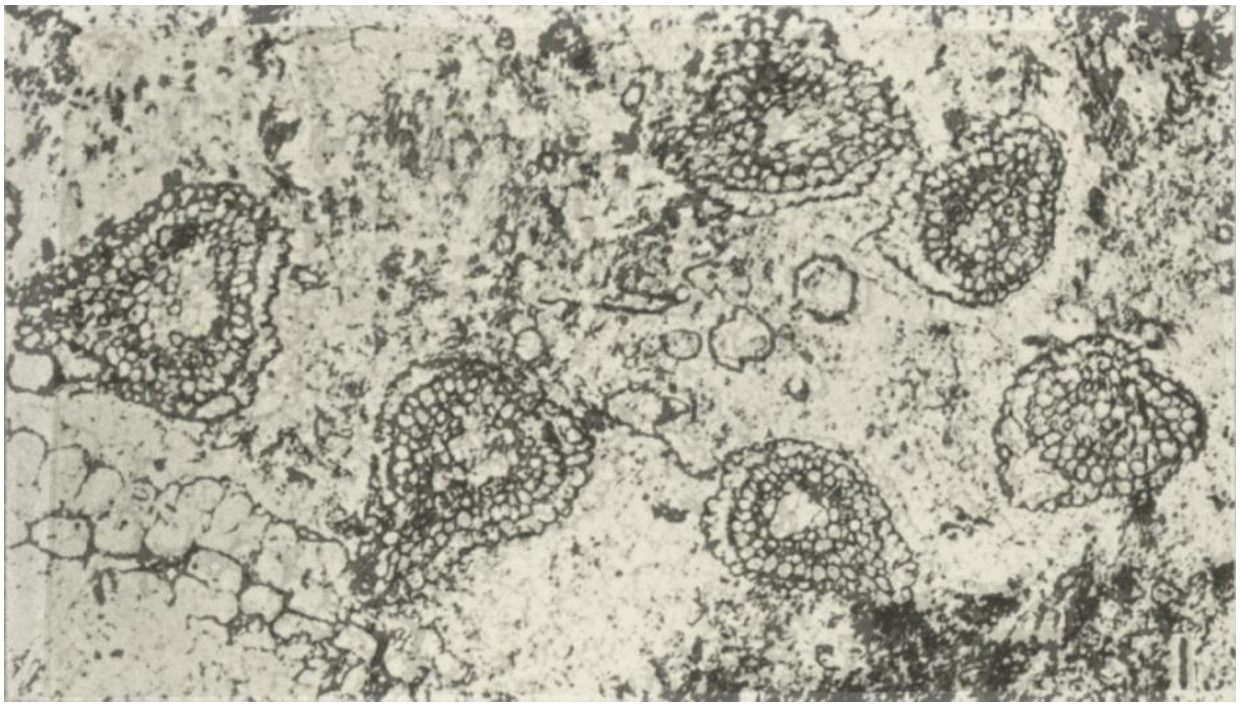


Figure 13 : Coupe proche de l'apex d'une graine fossile de *Physostoma stellatum* qui ne montre que les « tentacules » qui l'entoure (extrait de Holden, 1954). La couche cellulaire épidermique montre un décollement de la paroi externe par la rupture des parois radiales qui fait penser à des cellules sécrétrices de mucilage.

1.3.6. Utilisations par l'Homme

Le mucilage est utilisé depuis longtemps dans la pharmacologie Ces dernières années, les mucilages des graines trouvent de nombreuses applications dans l'industrie alimentaire et pharmaceutique grâce à leur diversité de structures et de compositions et au regain d'intérêt pour les produits d'origines naturelles. Ils sont utilisés pour leurs propriétés physiques (pour structurer, texturer, émulsifier, épaissir) mais aussi dans un cadre diététique car ils peuvent de par leur composition influer la digestion directement ou indirectement (Soukoulis et al., 2018).

La non-toxicité, l'absence de goût, d'odeur, ou de couleur, et la biodégradabilité des mucilages en font des substances très polyvalentes pouvant être utilisées pour l'emballage d'aliments ou de médicaments de l'échelle macroscopique au microscopique (film comestible, formation de comprimé, encapsulation, nanohydrogel) (Tosif et al., 2021)

Objectifs de la thèse

L'objectif de cette thèse est de mieux comprendre la variabilité naturelle de la myxospermie. Quelle histoire évolutive se cache derrière la diversité de myxospermie ou son absence entre les différentes espèces ? Quelle sont les contraintes environnementales qui ont provoqué l'apparition de cette adaptation ? Sur quelle base génétique repose-t-elle ? Ces questions sont très larges et nécessitent donc un angle d'approche plus resserré. L'espèce modèle myxospermique la mieux caractérisée à ce jour est *Arabidopsis thaliana*, elle est donc devenue naturellement le point central autour duquel les recherches se sont organisées. Pour tenter de répondre aux problématiques posées, cette thèse s'est ramifiée en deux types d'approches comparatives qui ont été menées en parallèle.

L'approche qui fait l'objet du chapitre 1 compare différentes espèces myxospermiques ou non myxospermiques au niveau morphologique et génétique (à l'échelle du gène) pour tenter de retracer le chemin évolutif qui les sépare. Les espèces ont été choisies au sein d'une famille en particulier (la famille d'*A. thaliana*) pour espérer réduire leurs potentielles divergences et ainsi obtenir des résultats interprétables, mais aussi pour des raisons de faisabilité qui seront expliquées en introduction du chapitre.

L'approche qui fait l'objet du chapitre 2 compare les différences morphologiques, génétiques (à l'échelle du nucléotide), et environnementales des populations naturelles d'une même espèce (*A. thaliana*) pour identifier de nouvelles base génétiques et de nouveaux rôles écologiques du mucilage des graines.

La première approche vise donc plus à étendre et comparer les connaissances acquises sur *A. thaliana* à d'autres espèces de sa famille pour comprendre leur évolution, alors que la deuxième approche prétend arriver à de nouvelles avancées sur les bases génétiques et les rôles écologique de la myxospermie chez *A. thaliana*.

Chapitre 1 : l'évolution de la myxospermie chez les Brassicaceae

Préface à l'article de recherche

La précédente thèse encadrée par le même duo de chercheur que celle-ci, menée par Edith Francoz, portait entre autre sur le développement des cellules sécrétrices de mucilage et sur les mécanismes d'ouverture de celles-ci chez l'espèce modèle *Arabidopsis thaliana* (Francoz et al., 2015, Francoz et al., 2019). Cette thèse a donc apporté des connaissances et des méthodologies pour étudier le mucilage extrudé à l'échelle macroscopique et les cellules sécrétrices en coupe à l'échelle microscopique (Francoz et al., 2019). C'est fort de ces connaissances, et face à la diversité de mucilages des graines décrite dans la littérature scientifique ainsi qu'au manque de compréhension du chemin évolutif qui les relient (cf article de revue § 1.3.2), que l'idée d'une étude comparative entre espèces est née. Ce genre d'approche comparative pour en déduire un scénario évolutif a été inspiré, entre autre, par les travaux au Laboratoire de Pierre-Marc Delaux et Jean Keller (Radhakrishnan et al., 2020). Ils s'intéressent à un trait, certes très différent (les symbioses intracellulaires), mais avec une méthodologie qui fait une utilisation conjointe de la morphologie et des données génomiques. Les nombreux gènes caractérisés chez *Arabidopsis thaliana* pour être impliqués dans la myxospermie constituaient un point de départ génomique conséquent pour une approche ciblée recherchant les bases génétiques de l'évolution de la myxospermie. De plus, les Brassicaceae, la famille d'*Arabidopsis thaliana*, contiennent de nombreuses espèces séquencées et des descriptions morphologiques du mucilage (Western, 2012, North et al., 2014) et des cellules épidermiques (Vaughan et Witheouse, 1971) qui assurait d'y trouver une grande diversité morphologique. Je me suis donc formé à la microscopie pour étudier divers traits liés à la myxospermie chez d'autres espèces de Brassicaceae qu'*Arabidopsis thaliana*, et à la biologie computationnelle pour regarder chez ces mêmes espèces ce qu'il advient des gènes apparentés à ceux impliqués dans le mucilage chez *Arabidopsis thaliana.* Ce chapitre a été publié dans *Cells* à la toute fin de ma thèse.

cells

MDPI

Article

Myxospermy Evolution in Brassicaceae: A Highly Complex and Diverse Trait with *Arabidopsis* as an Uncommon Model

Sébastien Viudes, Christophe Dunand * and Vincent Burlat *

Laboratoire de Recherche en Sciences Végétales, Université de Toulouse, CNRS, UPS, Toulouse INP, 31320 Auzeville-Tolosane, France; sebastien.viudes@lrsv.ups-tlse.fr

* Correspondence: dunand@lrsv.ups-tlse.fr (C.D.); burlat@lrsv.ups-tlse.fr (V.B.); Tel.: +33-534-32-38-58 (C.D.); +33-534-32-38-55 (V.B.)

Abstract: The ability to extrude mucilage upon seed imbibition (myxospermy) occurs in several Angiosperm taxonomic groups, but its ancestral nature or evolutionary convergence origin remains misunderstood. We investigated seed mucilage evolution in the Brassicaceae family with comparison to the knowledge accumulated in *Arabidopsis thaliana*. The myxospermy occurrence was evaluated in 27 Brassicaceae species. Phenotyping included mucilage secretory cell morphology and topochemistry to highlight subtle myxospermy traits. In parallel, computational biology was driven on the one hundred genes constituting the so-called *A. thaliana* mucilage secretory cell toolbox to confront their sequence conservation to the observed phenotypes. Mucilage secretory cells show high morphology diversity; the three studied *Arabidopsis* species had a specific extrusion modality compared to the other studied Brassicaceae species. Orthologous genes from the *A. thaliana* mucilage secretory cell toolbox were mostly found in all studied species without correlation with the occurrence of myxospermy or even more sub-cellular traits. Seed mucilage may be an ancestral feature of the Brassicaceae family. It consists of highly diverse subtle traits, probably underlined by several genes not yet characterized in *A. thaliana* or by species-specific genes. Therefore, *A. thaliana* is probably not a sufficient reference for future myxospermy evo–devo studies.

Keywords: *Arabidopsis thaliana* MSC toolbox gene orthologs; Brassicaceae species; cell wall microdomains; diversity; evolution; mucilage secretory cells (MSCs); seed mucilage (SM)

Citation: Viudes, S.; Dunand, C.; Burlat, V. Myxospermy Evolution in Brassicaceae: A Highly Complex and Diverse Trait with *Arabidopsis* as an Uncommon Model. *Cells* **2021**, *10*, 2470. https://doi.org/10.3390/cells10092470

Academic Editor: Bruce Kohorn

Received: 31 August 2021
Accepted: 16 September 2021
Published: 18 September 2021

1. Introduction

During land plant evolution, secondary growth and seeds were major innovations leading to the impressive species radiation currently observed in spermatophytes (seed plants) [1]. The adaptive capacity of seeds comes from their layered compartments surrounding the embryo with endosperm and seed coat, providing nutrients and protection, respectively, with both participating to seed physiology and developmental regulation [2]. The seed coat is a maternal tissue originating from the ovule integument that shows a high variability in architecture and thickness among seed plants [3]. In *Arabidopsis thaliana*, five outer cell layers undergo specialization during seed development, to finally be dehydrated and form a compacted seed coat in the mature dry seed. In this dead tissue, the two main layers are the inner integument enriched in pigments and the outer integument accumulating polysaccharides [4]. In *A. thaliana*, the outermost cell layer of the outer integument (oi2), also called mucilage secretory cells (MSCs), is composed of polarized epidermal cells, with a volcano-shaped central secondary cell wall structure called columella, surrounded by a hexagonal prominent primary cell wall viewed from the top of seed surface [5,6]. The dehydrated mucilage is compacted between the columella and the outer periclinal and radial primary cell walls [7]. Upon imbibition, the polysaccharidic mucilage swelling induces pressure towards the primary wall that simultaneously breaks at the outer periclinal/radial primary wall interface constituted by a prefragilized cell wall microdomain [8], allowing for mucilage extrusion to take place all around the seed [9]. The released seed

mucilage (SM) is composed of two main layers: the adherent mucilage (AM) strongly attached to the seed coat, and the water soluble non-adherent mucilage (NAM) [10]. The AM remains attached to the top of the columella, which is also the attachment point of cellulose microfibrils that provide an AM structure skeleton [11]. *A. thaliana* SM represents 3% of seed dry mass and is mainly composed of sparsely branched rhamnogalacturonan I (RG-I) pectins [6,12]. This composition allows the use of ruthenium red to easily assess SM occurrence, and immunocytolabeling with cell-wall-specific probes for more accurate SM phenotyping [13,14]. There are about 100 known MSC toolbox genes that contain all molecular actors known to be involved in *A. thaliana* SM synthesis, regulation and release [7,15,16] (Table S1). One third of them are involved in the transcriptional regulatory pathway, half are direct actors of mucilage synthesis, secretion, modification and structure, and the remaining part regroups genes implicated in cell wall dynamics and hormone synthesis and perception (Table S1) [7].

In Angiosperms, several so-called myxospermous species spread among a majority of orders were reported to release SM with highly diverse compositions and contrasted MSC morphology [15,17]. Great variation in mucilage amount and/or compositions also occurs at the genus level, such as in the *Arabidopsis* species [18] or *Plantago* species [19,20] and even among populations, such as in *A. thaliana* [21,22] and flax (*Linum usitatissimum*) [23,24]. Those large variations make it difficult to trace back the evolutionary scenario of myxospermy only based on the global biochemical or morphological description. Thus, for inter-species comparison, the analysis of molecular actors may provide a useful additional level of information to better understand morphological changes and the myxospermy evolution. On the one hand, the highly conserved MYB-bHLH-WD40 repeat (MBW) regulatory complex in Angiosperms is positioned at the basis of the *A. thaliana* MSC toolbox transcriptional pathway. Interestingly, some actors of this complex such as TT2 and MYB5 orthologs seem to only be detected in the Rosid clade, suggesting their apparition in the common ancestor of this clade [16]. On the other hand, the molecular actors underlying intra-species variability in *A. thaliana* correspond to a few enzymes responsible for polysaccharide production and modification [18,21,25]. Beside the members of the aforementioned conserved regulatory complex, no more *A. thaliana* MSC toolbox genes were studied with such an evolutionary perspective. The aim of this article is to fulfill the remaining gap occurring between *A. thaliana* myxospermy deep knowledge and the potential ancestral myxospermous species present in the Rosid common ancestor. The Brassicaceae family constitutes an excellent taxonomic framework to initiate such a study, since it displays great MSC and SM morphological variations [18,26,27], as well as a large number of available qualitative genomic data and a well-characterized phylogeny [28–30].

We selected 33 Brassicaceae species with worldwide distribution that covered the phylogenetic tree. We performed a systematic phenotyping analysis of 27 out of these 33 Brassicaceae species using (i) a ruthenium red assay to evaluate the putative mucilage release, (ii) a histochemical morphological analysis of the dry seed and imbibed seed epidermis to provide insight on the morphology of MSC and of the SM release mode and (iii) an immunofluorescence analysis of selected cell wall epitopes within the MSC cell wall domains. This phenotyping analysis was compared to the *A. thaliana* MSC toolbox gene orthologs that we found while studying 32 out of these 33 Brassicaceae species. This genomic survey was completed by data mining of published seed development transcriptomic data for *A. thaliana*, *Camelina sativa*, *Brassica napus* and *Aethionema arabicum*.

2. Materials and Methods

2.1. Plant Material, Seeds and Genomes

We selected Brassicaceae species according to three criteria: (i) the species distribution among 3 lineages, containing 15 tribes, (ii) the genomic/seed transcriptomic data availability and (iii) the seed availability for phenotyping. A pool of 33 Brassicaceae species was first chosen according to the phylogenetic distribution. Seeds were available for phenotyping for 27 of them, to which we attributed a number code for the sake of clarity

(Tables S2 and S3). Genomic data were available for 32 species (except species 26) and seed specific transcriptomic data were available for 4 species (species 1, 6, 13, 27) among these 27 species. Seeds were not available for 6 of the 33 species (species A to F). Two tribes were more represented than the others: the Camelineae tribe (including *A. thaliana* for fundamental research interest) and the Brassiceae tribe (including nearly all Brassicaceae domesticated crop species). Two non-Brassicaceae species, *Medicago truncatula* and *Linum usatissimum*, were added to the genome list to serve as an outgroup for phylogeny analysis. Seeds used during this study originated from commercial resources, scientific exchanges or personal harvests, as detailed in Table S2. The information on genome origin and relative quality is compiled in Table S3.

2.2. Dry Seed Phenotyping

The gross morphology of mature dry seed (color, size, and seed coat surface appearance) was screened for 27 species using an Epson perfection V100 photo scanner at a 6400 dpi resolution in opaque mode.

2.3. Phenotyping of Adherent Mucilage

In order to standardize the release and the staining of adherent SM among the 27 species, we used the ruthenium red staining protocol previously described [8] with a unique modification (seeds were first imbibed for 5 min in a 0.1% aqueous solution of calcofluor to also visualize cellulose microfibrils). The seeds were observed with a Leica DMIRB/E inverted microscope using a Leica MC190HD camera in the bright field mode and fluorescent mode using a UV filter set (excitation: 387/11 nm; dichroic mirror: 405 nm; emission: 440/40 nm) for ruthenium red and calcofluor staining, respectively.

2.4. Cross Section of Mucilage Secretory Cells, Histochemistry and Immunofluorescence

Dry seeds were first individually punctured with an ultrathin needle under a dissecting microscope to facilitate further fixation. Dry seeds from 27 species were fixed in a fixative solution (1.25% glutaraldehyde/2% paraformaldehyde in 0.05 M PIPES, 5 mM EGTA, 5 mM $MgSO_4$, pH 6.9 buffer complemented with 0.1% Triton X-100 and 50% ethanol to avoid SM release and have a view of intact epidermal cells/MSCs. Further infiltration with LRW resin used a EM AMW (Automatic Microwave Tissue Processor for Electron Microscopy, Leica Microsystèmes SAS, Nanterre, France), as previously described [8]. Imbibed seeds were similarly processed to study the SM releasing mode, except that ethanol was removed from the fixative solution. Semi-thin cross sections (1 µm-thick) were cut using a histo diamond knife (Diatome) and an Ultracut E Reicher Ultramicrotome. Serial sections were spread on silane-coated slides to constitute series of species-specific slides comprising numerous serial sections, as well as two series of tissue arrays encompassing the whole panel of dry seeds or of imbibed seeds, respectively. Tissue arrays and species-specific slides were simultaneously stained for 1 min at 60 °C in 0.05% Toluidine blue in 0.1 M acetate buffer pH 4.6 for morphology. Tissue arrays were also processed for immunofluorescence as previously described [8] using selected monoclonal antibodies, allowing one to visualize the cell wall domain important for MSC opening in *A. thaliana* (LM20) [8] and the main component of the *A. thaliana* mucilage (INRA-RU2) [13], respectively. The LM20 rat monoclonal antibody specific for partially demethylesterified homogalacturonans [31,32] and the INRA-RU2 mouse monoclonal [13] specific for the RGI backbone were used at 1:10 dilution and indirectly detected using anti-rat Ig-A488 and anti-mouse Ig-A488 (Invitrogen, Thermofisher Scientific, Walthman, MA, USA), respectively. Slides were mounted under a coverslip inc prolong gold anti-fade (Invitrogen, Thermofisher Scientific, Walthman, MA, USA) and scanned at high resolution (20× or 40×) with a nano 2.0RS Hamamatsu slide scanner in the bright field mode and fluorescent mode (FITC filter set) for Toluidine blue and immunofluorescence slides, respectively. Scanned images were analyzed with NDP view (Hamamatsu), and figures were mounted using Corel Photopaint (CorelDraw graphics suite X6, Ottawa, ON, Canada), using drawings made with Microsoft PowerPoint

(Microsoft office 2007, Redmond, WA, USA). Grey panels corresponded to unavailable or unexploitable samples.

2.5. Genomic Data Mining of A. thaliana MSC Toolbox Gene Orthologs

Coding DNA Sequences (CDS) of 32 Brassicaceae species and 2 outgroup species were used directly from available annotated genomes or predicted with fgenesh software for species without annotation (Table S3) [33]. Every CDS belonging to each Brassicaceae species was concatenated in a single multifasta file for further ortholog data mining. Each MSC toolbox protein sequence of *A. thaliana* collected from Phythozome was blasted against the concatened Brassicaceae multi-CDS file using tblastn (ncbi-blast, version 2.7.1+, https://www.ncbi.nlm.nih.gov/, (accessed on 30 August 2021)). Only the 1000 first hits, or all hits with an e-value lower than 10-10, were recovered. Each blast output was aligned using mafft (version 7.313, https://mafft.cbrc.jp/alignment/software/, (accessed on 30 August 2021)) and gap deletion was applied for a gap shared in 10% or more of the sequences using trimal (version 1.4.1, http://trimal.cgenomics.org/, (accessed on 30 August 2021)). Phylogenetic trees were built using iqtree (version 1.6.7, http://www.iqtree.org/, (accessed on 30 August 2021)) with 10,000 bootstraps from these alignments, and the clade containing the initial *A. thaliana* MSC toolbox gene up to *M. truncatula* and *L. usitatissimum* sequence(s) was extracted. The obtained tree was rooted on the branch containing these two species, because they do not belong to the Brassicaceae family and consequently constitute an outgroup for the other species. As Brassicaceae species genes are often duplicated from their common ancestor, several phylogeny trees may contain paralog gene(s) of the *A. thaliana* gene of interest and their corresponding orthologs. In such cases, only genes of the sub clade containing the *A. thaliana* MSC toolbox gene and delimited by the closest *A. arabicum* sequence, as it is the early divergent species of the studied Brassicaceaes species, were considered as orthologs.

2.6. Transcriptomic Data Mining of A. thaliana MSC Toolbox Gene Orthologs from A. thaliana, C. sativa, B. napus and A. arabicum Public Seed Development Transcriptomes

The whole seed development kinetic expression values were retrieved: for *A. thaliana*, 6 seed developmental stages [34], for *C. sativa*, 6 cumulated seed developmental stages [35,36], for *B. napus*, 4 seed developmental stages [37] and for *A. arabicum*, 3 developmental stages of myxospermous vs. non-myxospermous dimorphic diaspores [38]. For each gene in each species, the expression values of whole seed development kinetics samples were summed and ranked using Microsoft Excel. The mean SUM expression values for the top 10 genes in each species were used as references to calculate the relative sum expression percentage of orthologous genes of *A. thaliana* MSC toolbox genes (Table S1) in each species. For *C. sativa*, we used the previously defined correlation table [39] to identify, in the hybridized genomes, the homeologous genes corresponding to *A. thaliana* genes, and the expression values of each homeologous gene corresponding to an *A. thaliana* gene were first summed (details of homeologs are shown on the right part of Table S10). For *B. napus*, the same addition was originally made [37]. To highlight the differences in the columns "SUM%", a common red (40)-to yellow (1)-to blue (0) heatmap was drawn using Microsoft Excel.

3. Results

3.1. Phylogenic Relationship between 27 Selected Species and Their Mature Dry Seed Whole Morphology

In total, 27 Brassicaceae species were chosen in order to obtain available seeds and genomic data for 26 of them, and to provide a good taxonomic coverage among the family (Figure 1). The numbering in brackets following each species is used in all the figures and text to simplify their descriptions.

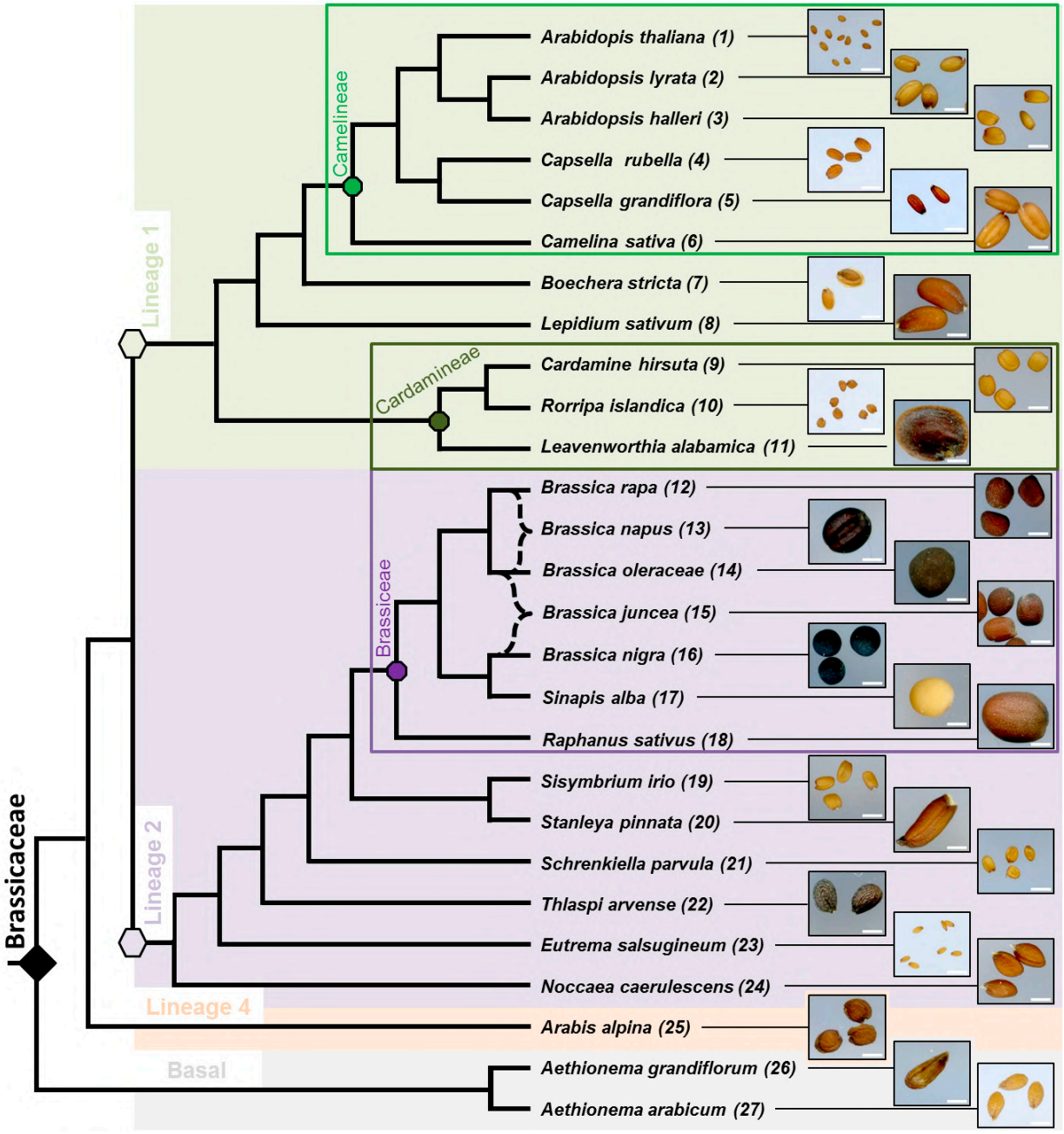


Figure 1. Simplified phylogenetic tree of the 27 phenotyped Brassicaceae species and illustration of their dry seed phenotype. Phylogenetic relationships between species are based on [28–30]. For sake of clarity, the species names were each associated with a number used to facilitate their identification along the manuscript. Lineages are highlighted by different background colors and the three tribes containing more than one species are labeled by colored rectangles. Note the high diversity of dry seed size and morphology observed with high resolution scanning. Scale bar: 1 mm.

Lineage 1 (11 species) and 2 (13 species) are the most represented because they contain the well-studied tribes Camelineae (six species) and Brassiceae (seven species). Camelineae is the tribe of *A. thaliana* (1) and Brassiceae contain a lot of cultivated species, such as rapeseed (*Brassica napus* (13)). Note that the unusual branch style used for *B. napus* (13) and *Brassica juncea* (15) symbolizes their peculiar origins. Both species evolved from a common ancestor coming from the hybridization between the two connected species indicated within the tree (Figure 1) to form the current species. Within the family, seed color, size and shape show wide variation between species. All Camelineae seeds are oblong with

different sizes (from 0.5 to 5 mm). Among the family, some seeds belonging to different tribes are flattened and winged, such as in *Boechera stricta* (7), *Leavenworthia alabamica* (11) and *Arabis alpina* (25). Then, these morphological modifications are probably due to evolutionary convergence of the two related traits in the Brassicaceae. Spherical seeds are only found in the Brassiceae tribe despite their contrasted colors, as exemplified by the black and white mustards (*Brassica nigra* (16) and *Sinapis alba* (17)). Then, this seed shape change should have occurred in the common ancestor of the Brassiceae. *A. arabicum* (27) has dimorphic seeds [40]; only the myxospermous morphotype is shown here.

3.2. Myxospermy Occurrence and Morphology at the Whole Seed Level

Within Camelineae, three *Arabidopsis* species (1–3), two *Capsella* species (4, 5) and *Camelina sativa* (6) extruded a ruthenium red-stained pectinaceous adherent mucilage, (AM) similarly to *A. thaliana* (1) used as a control (Figure 2).

Despite a common cohesion with the seed surface, the thickness and the flatness of the AM periphery can change between species. The other species of lineage 1 have cohesive AM similar to Camelineae, such as *Lepidium sativum* (8), *Cardamine hirsuta* (9), and *Rorripa islandica* (10) (Figure 2). However, *Boechera stricta* (7) does not extrude AM (the faint pink ruthenium red staining corresponds to light transmission through the thin cell wall pectinaceous material of the wing rather than AM extrusion per se), and *Leavenworthia alabamica* (11) shows extruded filamentous mucilage structures (Figure 2). Interestingly, these two exceptions also correspond to the only flattened and winged seed in the lineage 1 studied species (Figure 1). In lineage 2, most studied Brassiceae species do not release mucilage (Figure 2) except for the two mustards, *B. nigra* (16) and *S. alba* (17). These two species have a cohesive layer surrounding the seed, as in Camelineae. However, they do not seem to have cellulosic content in their mucilage, contrary to all the other myxospermous species studied (Figure S1). The mucilage trait of the other studied species belonging to lineage 2 is highly diverse (Figure 2). Some do not display any AM, such as *Thlaspi arvense* (22) and *Noccaea caerulescens* (24), or show a very thin layer of AM that requires a higher resolution to be observed, such as for *Stanleya pinnata* (20). Other species of lineage 2 show the extrusion of several conical structures at the seed surface, such as *Sisymbrium irio* (19) and *Eutrema salsugineum* (23), or display large pieces of poorly cohesive mucilage for *Schrenkiela parvula* (21) (Figure 2). Finally, in the species of lineage 4, *Arabis alpina* (25) and the two early divergent species studied *Aethionema grandiflorum* (26) and *Aethionema arabicum* (27) show the extrusion of independent mucilage structures (Figure 2), similar but longer to those observed in *Sisymbrium irio* (19) and *Eutrema salsugineum* (23) (Figure 2). These structures mostly look like spines all around the seed and, in *A. alpina* (25), these occur in both the spherical part of the seed and the wing. *A. arabicum* (27) displays the longest and thickest ones that ended by a spherical shape (Figure 2).

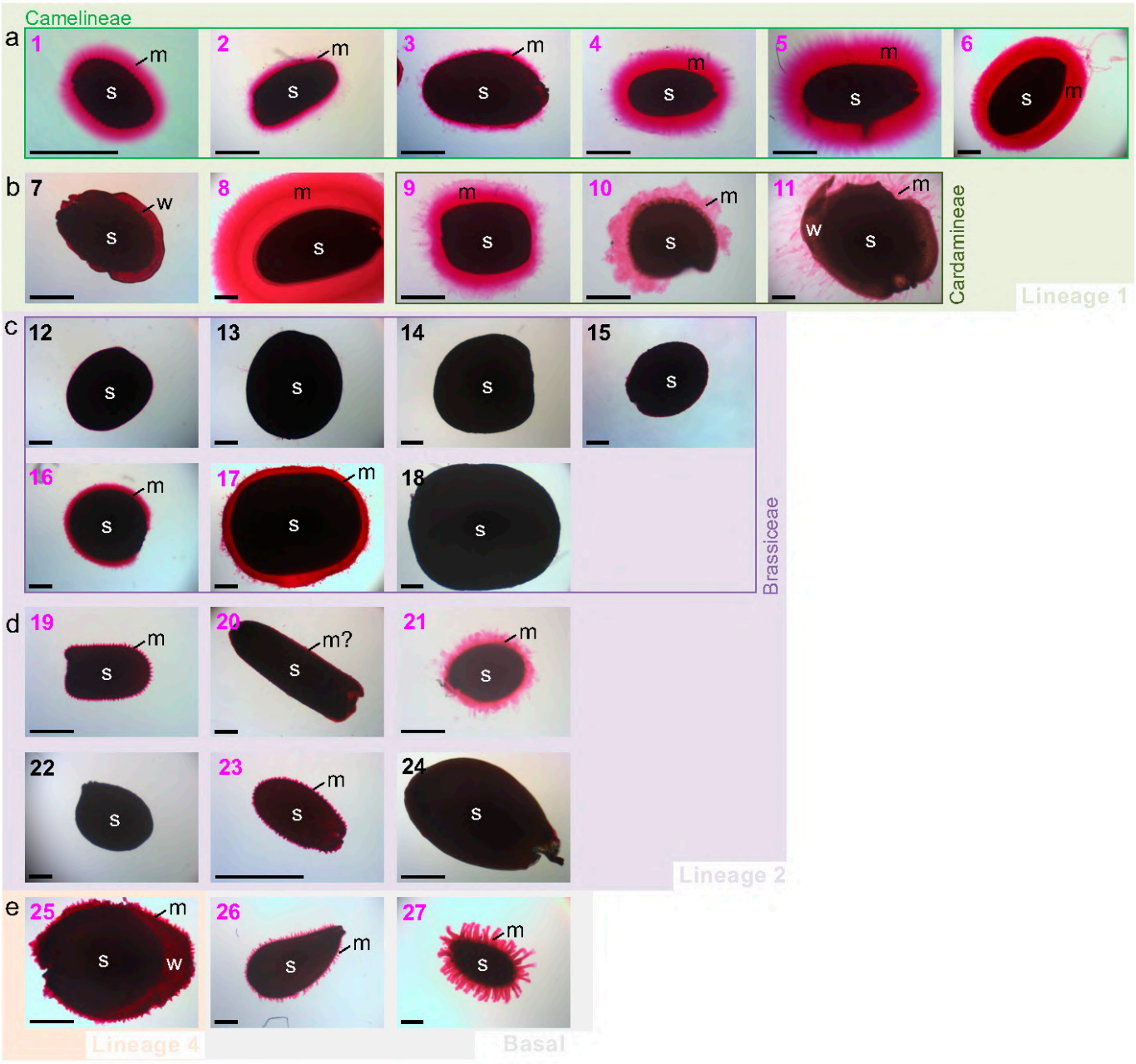


Figure 2. Among the 27 Brassicaceae species studied, 19 appear as myxospermous and 8 as non-myxospermous, regardless of their phylogenetic position. Camelineae lineage 1 species (**a**), other lineage 1 species (**b**), Brassiceae lineage 2 species (**c**), other lineage 2 species (**d**), lineage 4 and basal species (**e**). Following a strong shaking of imbibed seeds to facilitate the extrusion of putative mucilage, the adherent pectinaceous mucilage was stained in pink by ruthenium red in a standardized manner among the 27 species. Species numbers presented in Figure 1 were used to identify the species and colored in pink and black for the 19 myxospermous and 8 non-myxospermous species, respectively. Most of myxospermous species display a clearly visible polymorphic pink mucilage halo, while four myxospermous species (19, 20, 23, 26) show a faint mucilage halo (see Figure S1 for additional view) as compared to the non-myxospermous species (7, 12–15, 18, 22, 24). Note that the polarized pink staining in *B. stricta* (7) corresponds to staining of the pectinaceous cell wall material from the thin wing allowing light transmission. The color background, image framing and image spaces reflects the phylogenetic distribution of the species (see Figure 1 for more detail). Scale bar: 500 µm.

3.3. Dry Seed Epidermis/MSC Morphology and MSC Opening Mode in Imbibed Seeds

Increased resolution and additional information on the mucilage extrusion mode were obtained through histochemical study of epidermal cells before (dry seed) and after (imbibed seed) potential mucilage release, followed by simplified summary drawings for each species (Figures 3–5). The metachromatic Toluidine blue O stains the primary cell wall in dark blue and the mucilage and columella in fainter blue-pinkish, whenever present. These three structures were represented with more contrasted blue, pink and green color codes, respectively, in the deduced and simplified drawings, while unidentified internal staining was represented in grey (Figures 3–5). In *A. thaliana* (1) dry seed, SM is stored under the MSC primary cell wall. It forms a donut shape around the columella, a volcano-shaped secondary cell wall. Upon seed imbibition, the mucilage inflates, breaking a primary cell wall domain located at the top of radial cell walls separating each cell and allowing mucilage extrusion all around the seed, while the ruptured primary wall remains attached to the top of the columella (Figure 3). The six studied Camelineae species (1–6) display a MSC morphology very similar to *A. thaliana* before imbibition (Figure 3). They have rectangular cells delimitated by a radial cell wall with a central columella. However, after imbibition, the breaking point of the radial cell wall domain commonly observed in the three *Arabidopsis* species (1–3) seems to be different from the rupture mode observed in the two *Capsella* species (4, 5) and *C. sativa* (6) despite their phylogentic proximity (Figure 3). Contrary to the three *Arabidopsis* species, *C. sativa* has a primary wall-breaking domain localized at the top of the columella, and following imbibition, the periclinal primary CW remains attached to the radial CW. *Capsella* species show an intermediate situation between the *Arabidopsis* species and *C. sativa*, since they have a part of periclinal CW that remains attached to both the top of columella and of radial CW (Figure 3). Within lineage 1, the epidermis/MSCs of non-Camelineae species are much more diverse (Figure 3). They are all myxospermous except for *Boechera stricta* (7), which has no primary CW rupture and whose seed epidermis is also extremely reduced in size (Figure 3). An *Arabidopsis*-like columella is observed in *Boechera stricta* (7) and *Cardamine hirsuta* (9), but it can also be drastically changed in shape, as for *Lepidium sativum* (8) or simply disappear, as for *Rorippa islandica* (10) and *Leavenworthia alabamica* (11) (Figure 3). The primary CW rupture zone of *L. sativum* (8) and *C. hirsuta* (9) is central, as for *C. sativa*, while for *R. islandica* (10) and *L. alabamica* (11), the rupture occurs at the top of the radial CW as for the *Arabidopsis* species, but this only happens within a restricted number of cells (Figure 3).

Within lineage 2, the studied Brassiceae species can have more or less flattened epidermal cells, but they never possess a columella, per se, or a columella-like structure, even if *S. alba* (17) MSCs contain intracellular structures of unknown function (Figure 4). As expected with the previously described ruthenium red staining, both mustards (*B. nigra* (16) and *S. alba* (17)) are the only Brassiceae species displaying a clear extrusion with a rupture of primary CW occurring at the middle of MSC periclinal CW (Figure 4). Among the five studied species of lineage 2 that do not belong to the Brassiceae tribe, there is a great diversity of epidermal cell morphology (Figure 4). *Schrenkiella parvula* (21) MSCs are similar to mustard MSCs (no intracellular structure and central opening), but the border of released mucilage patches is more reactive to Toluidine blue (Figure 4). The released mucilage forms even more delimitated patches with spine shapes above each MSC in *Sisymbrium irio* (19) and *Eutrema salsugineum* (23), in agreement with the observed ruthenium red staining (Figures 2 and 4). *Stanleya pinnata* (20) shows intracellular layered structures that increase in volume but are not extruded following imbibition, while *Noccaea caerulescens* (24) shows a similar pattern without an apparent cell volume increase (Figure 4). *Thlaspi arvense* (22) seed's epidermis does not display particular sign of MSC differentiation, since these cells appear empty, without any change following imbibition (Figure 4).

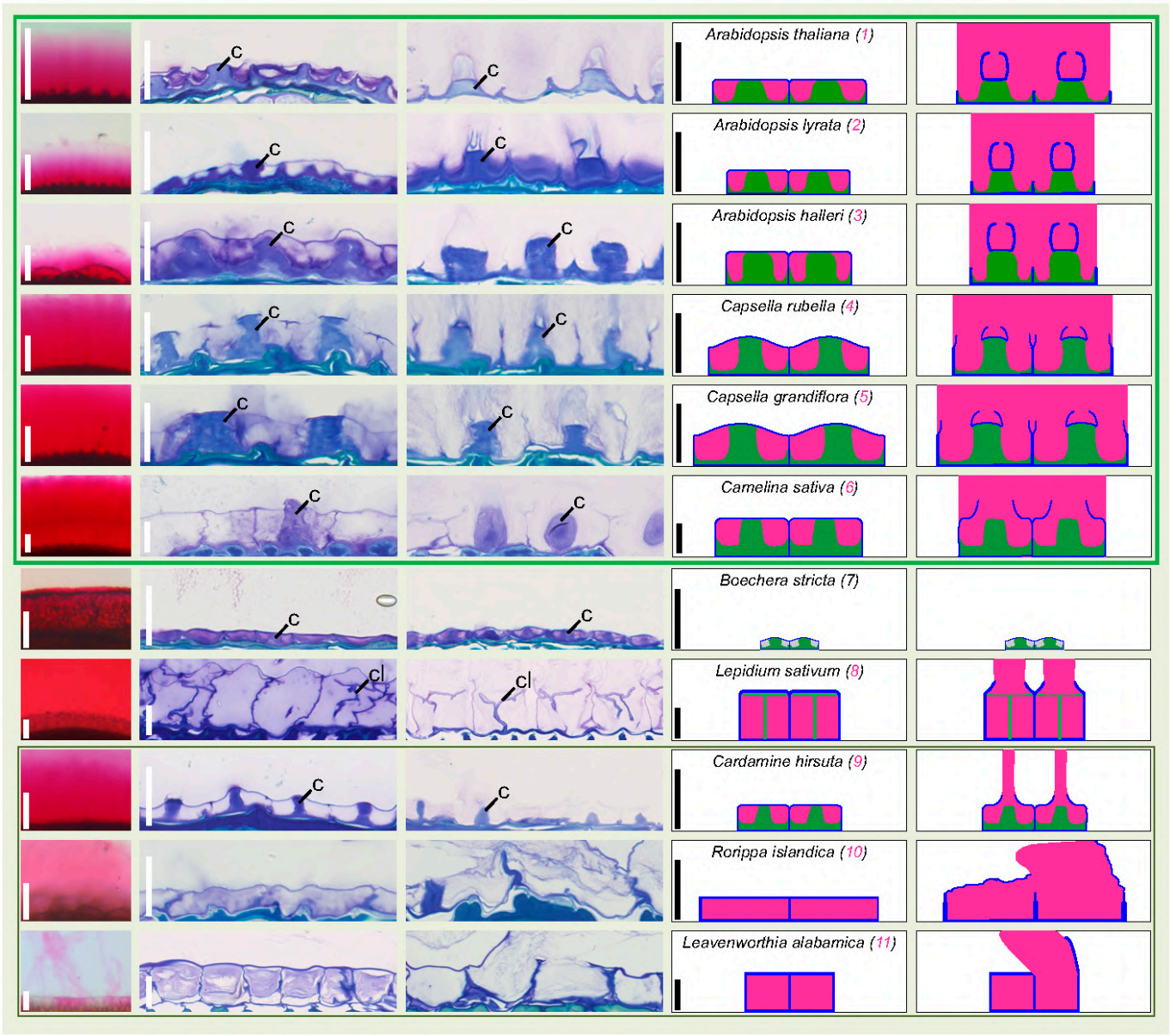


Figure 3. Lineage 1 dry seed MSC morphology is similar to the well-studied *A. thaliana* MSC morphology in all studied Camelineae species and highly diverse in non-Camelinae species, while the *A. thaliana* MSC opening mode is restricted to *Arabidopsis* species. Column 1: magnified view of imbibed seeds stained with ruthenium red (full images are shown in Figure 2) reminds the ability to release mucilage or not. Scale bar: 100 µm. Columns 2 and 3: Toluidine blue histochemical staining of dry seeds and imbibed seeds, respectively, focusing on the outermost cell layer morphology and of mucilage release mode for myxospermous species. Note that the inner part of the seed is located below the imaged cells. Columella (c) and columella-like structure (cl) are labeled. Scale bar: 25 µm. Columns 4 and 5: simplified drawings of two epidermal cells at the same scaling shown in Toluidine blue images. For each line the species name, the species number (pink for myxospermous and black for non-myxospermous species), as well as the scale of cross sections is reminded in column 4. A color code is applied for mucilage (pink), primary wall (blue), columella or columella-like structure (green), intracellular unidentified structures or sublayering (grey).

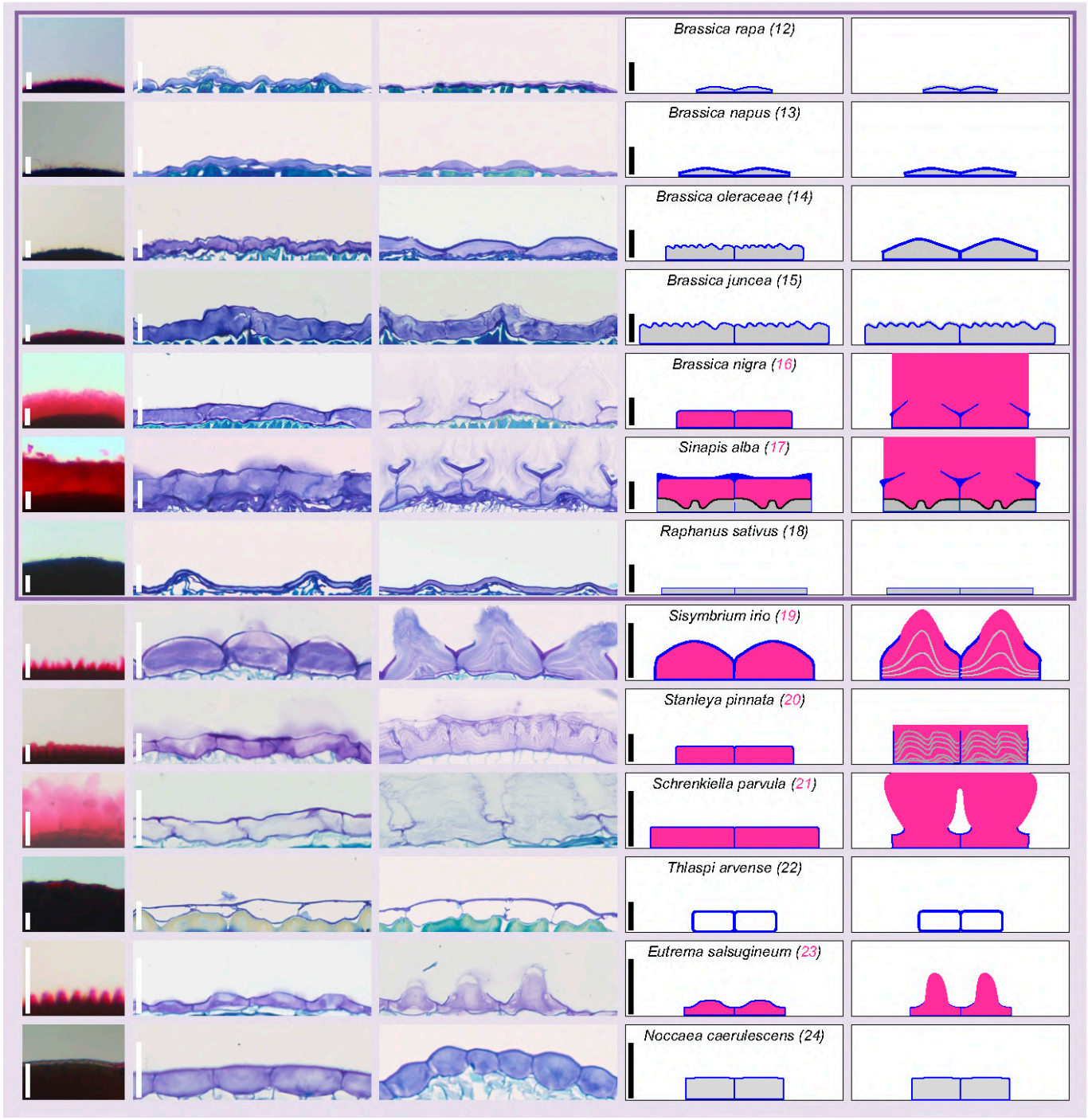


Figure 4. Brassiceae and other lineage 2 species seed epidermis/MSC morphology is highly diverse and contrasted as compared to lineage 1 species. Column 1: magnified view of imbibed seeds stained with ruthenium red (full images are shown in Figure 2) reminds the ability to release mucilage or not. Scale bar: 100 µm. Columns 2 and 3: Toluidine blue histochemical staining of dry seeds and imbibed seeds, respectively, focusing on the outermost cell layer morphology and of mucilage release mode for myxospermous species. Note that the inner part of the seed is located below the imaged cells. Scale bar: 25 µm. Columns 4 and 5: simplified drawings of two epidermal cells at the same scaling shown in Toluidine blue images. For each line the species name, the species number (pink for myxospermous and black for non-myxospermous species), as well as the scale of cross sections is reminded in column 4. A color code is applied for mucilage (pink), primary wall (blue), columella or columella-like structure (green), intracellular unidentified structures or sublayering (grey).

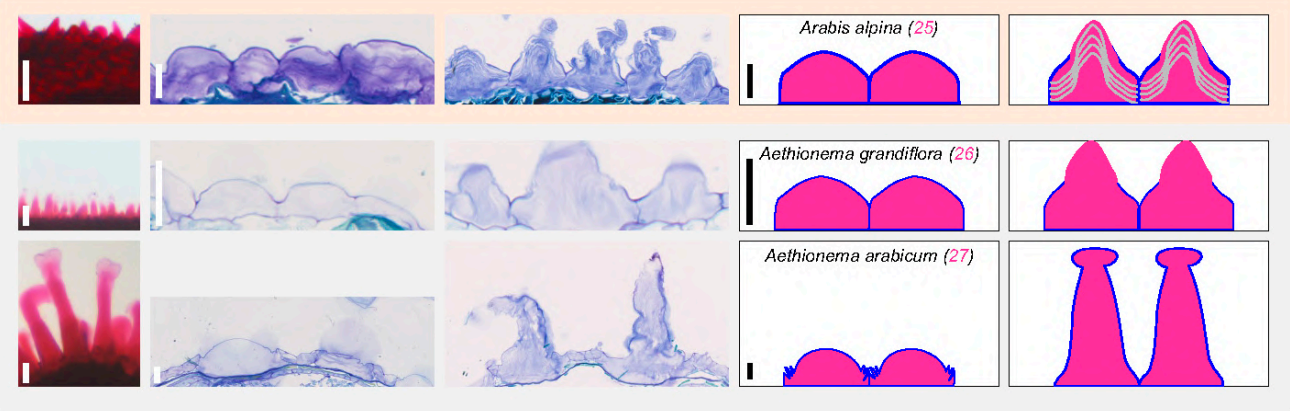


Figure 5. Lineage 4 and early diverging studied species display large MSCs forming a spiny-shaped mucilage upon imbibition. Column 1: magnified view of imbibed seeds stained with ruthenium red (full images are shown in Figure 2) reminds the ability to release mucilage or not. Scale bar: 100 µm. Columns 2 and 3: Toluidine blue histochemical staining of dry seeds and imbibed seeds, respectively, focusing on the MSC morphology and of mucilage release. Note that the inner part of the seed is located below the imaged cells. Scale bar: 25 µm. Columns 4 and 5: simplified drawings of two epidermal cells at the same scaling shown in Toluidine blue images. For each line the species name, the species number (pink for myxospermous species), as well as the scale of cross sections is reminded in column 4. A color code is applied for mucilage (pink), primary wall (blue), columella or columella-like structure (green), intracellular unidentified structures or sublayering (grey).

No lineage 3 species could be observed and the only lineage 4 studied species, *Arabis alpina* (25), displays central mucilage extrusion in independent layered spine-shaped structures, similar to the lineage 2 species *S. irio* (19) and *E. salsugineum* (23) (Figures 4 and 5).

Finally, the two studied early diverging species from the Brassicaceae family, *A. grandiflora* (26) and *A. arabicum* (27), release independent mucilage structures after imbibition (Figure 5). The larger size of *A. arabicum* (27) MSCs explains the larger size of the mucilage spine-shaped structure released upon imbibition. It seems that contrary to the other species releasing delimitated mucilage structure above each MSC (*A. grandiflora* (26), *A. alpina* (25), *E. salsugineum* (23), and *S. irio* (19)), *A. arabicum* (27) MSCs do not undergo CW rupture at the central position of the periclinal CW. It rather seems that the periclinal CW unfolds in an accordion during imbibition (Figures 4 and 5).

3.4. In Situ Immunolabeling of SM Content and MSC Primary Cell Wall Microdomains

For some species that do not extrude mucilage, it is difficult to decipher whether the seed epidermal cells actually contain unreleased mucilage by a lack of primary CW rupture, or if these cells do not differentiate as MSCs. *A. thaliana* mucilage is enriched in sparsely branched RGI pectins labeled with INRA-RU2 monoclonal antibody [13]. RU2 can be used as a marker of mucilage presence on serial seed cross sections in parallel of sections used for Toluidine blue staining. Upon imbibition, Camelineae species maintain a strong AM signal close to the newly revealed seed surface, especially for the three *Arabidopsis* species (1–3). As expected, RU2 staining is located within MSCs of almost every myxospermous species, regardless of their position within the phylogenetic tree (Figure 6). Surprisingly, RU2 labels the cell layer just below the MSCs and not the MSCs themselves for the two *Aethionema* species (26, 27). Interestingly, *L. sativum* (8), *R. islandica* (10) and *L. alabamica* (11), which belong to lineage 1 but do not present a columella, show weak-to-absent RU2 labeling, suggesting a mucilage compositional change. *Stanleya pinata* (20) MSCs are labeled only in the upper part while they are imbibed, suggesting that they extrude mucilage and not only inflate during imbibition. *C. hirsuta* (9) probably has an enrichment of sparsely branched RGI, as the labeling of its MSC content by RU2 is the strongest in the family. The eight non-myxospermous species (*B. stricta* (7), five non-mustard Brassiceae species (12–15, 18),

Thlaspi arvense (22), and *N. caerulescens* (24)) are not labeled with RU2, supporting the lack of pectinaceous mucilage (Figure 6).

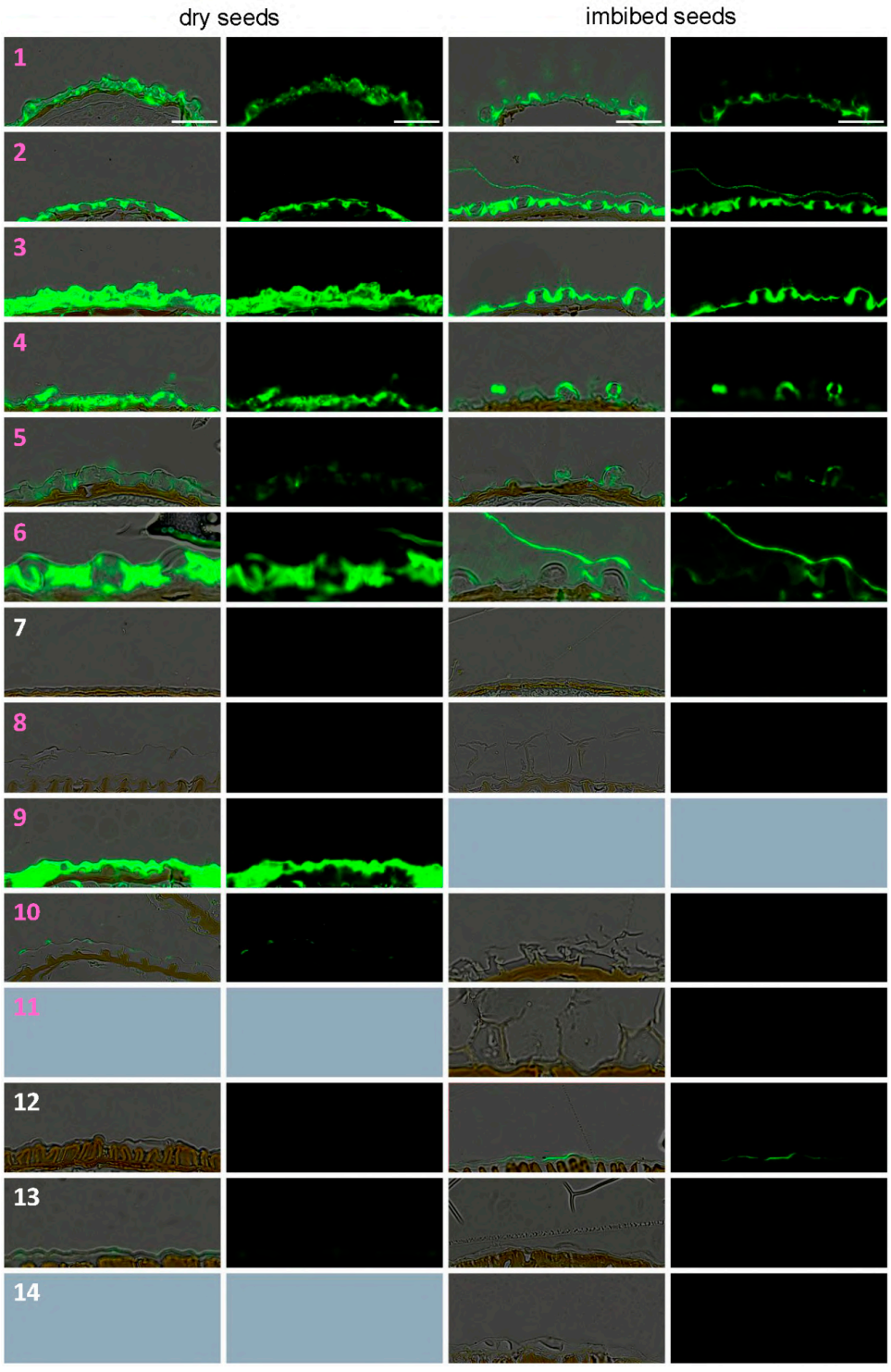


Figure 6. *Cont.*

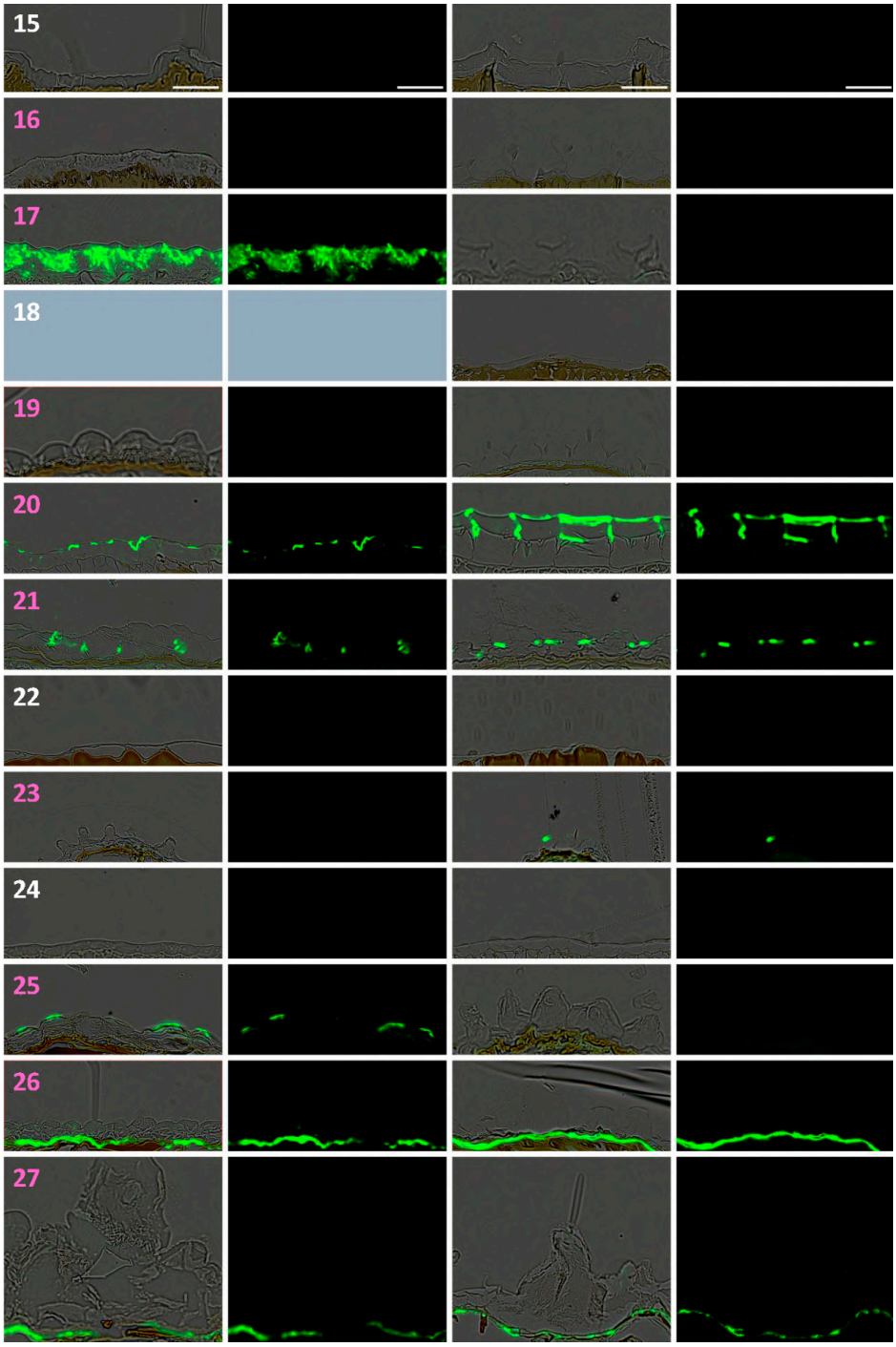


Figure 6. Immunofluorescence survey of mucilage in dry and imbibed seeds using the INRA-RU2 anti-RGI backbone antibody (2 pages). Serial sections of tissue arrays used for Toluidine blue staining were labeled with INRA-RU2 and anti-mouse Alexa488 as a secondary antibody. Note the recurrent occurrence of labeling in most myxospermous species regardless their position in the phylogenetic tree. Species numbers defined in Figure 1 were colored in pink and white for myxospermous and non-myxospermous species, respectively. Grey panels correspond to unavailable or unexploitable samples. Bars: 50 µm.

In *A. thaliana* dry seeds, LM20 antibody labels the MSC radial CW microdomain corresponding to the rupture zone for a proper extrusion during imbibition [8] (Figure S2). Here, the studied species were screened with this antibody to evaluate two hypotheses. First, whether the epidermal cells that do not undergo rupture during imbibition lack

this specific domain. Second, whether the shift in the rupture zone location from the periphery to the center could be correlated with a shift of the LM20 labeling zone. The periclinal primary CW of almost every studied myxospermous and non-myxospermous species displays a signal for LM20 (Figure S2). Only two species do not have labeling: the non-myxospermous species *R. sativus* (18) and *T. arvense* (22), which interestingly have very atypical epidermal cells, namely the thinnest but wider ones and the only empty ones, respectively. The labeling can be continuous as for *L. alabamica* (11) and all studied Brassiceae species (12–17) or discontinuous, as in *A. thaliana* (1). For a discontinuous signal it may be difficult to decipher the localization of the labeling, since the perfect cutting plan is not always available on tissue arrays. The myxospermous species with a spiny morphology of MSCs (19, 23, 25–27) show an interesting central labeling in dry seed MSCs, and from either side of the opened area following imbibition. For *S. pinnata* (20), the LM20 labeling of imbibed MSCs is located just below the labeling zone observed for RU2, highlighting again the potential extrusion of this species through the lateral zone of MSCs.

3.5. Bioinformatics Insight on *A. thaliana* MSC Toolbox Gene Orthologs in Brassicaceae Species

With the aim to highlight potential change at the genetic level between myxospermous and non-myxospermous species and to establish relationships with the subtle traits revealed by the phenotyping, we evaluated a targeted approach, identifying in 32 species, including 26 of the 27 species phenotyped in this study, the putative orthologous genes of the *A. thaliana* MSC toolbox genes characterized so far (Table S1). This approach was based on two hypotheses. First, orthologous genes of *A. thaliana* MSC toolbox genes in other Brassicaceae species may have the same functions as the ones characterized for *A. thaliana*. Second, gene(s) involved in myxospermy could have undergone pseudogenization in species that have lost myxospermy and, as a result, become dissimilar in their sequences compared to those which still have their function. Thorough data mining allowed us to find at least one orthologous gene for a large majority of *A. thaliana* MSC toolbox genes for each Brassicaceae species analyzed, including the eight studied non-myxospermous species (Figure 7, Table S4).

There are a few interesting exceptions to the general tendency of gene conservation with a single copy in each studied species. *CESA9* undergo a massive pseudogenization, as an ortholog was only found in the *Arabidopsis* species (1, 2, 3) and in the earliest diverging species, such as *A. arabicum* (27) (Table S4). *PMEI14* is an example of the rare duplication event that occurs in the common ancestor of a sub-clade of Brassicaeceae, here lineage 1, as there is more co-orthologs in it compared to the rest of the family (Table S4). This may also be the case for *DOF4.2*, *SK11-SK12*, *TBA1-TBA2-TBA3*, and *MYB75*; however, their corresponding phylogenetical trees were not sufficiently resolved to be properly interpreted. This is probably due to the short size of the sequences (Table S4). Therefore, unfortunately, no simple direct relationship could be established between the absence of *A. thaliana* MSC toolbox gene orthologs and the lack of the myxospermy trait (Figure 7). Indeed, species showing the less orthologous gene presence are myxospermous, such as *Brassica nigra* (16) and *Arabis alpina* (25), and conversely, non-myxospermous species, such as *Boechera stricta* (7), contain most of *A. thaliana* orthologs (Figure 7, Table S5). When sorting the data to look for a putative correlation between gene losses and a more subtle phenotype at the cellular level such as a columella presence (Table S6) or the position of the primary cell wall rupture zone (Table S7), no clear correspondence appears as well. In the last case, PMEI6 and PRX36 are shown to be directly linked to the fragilization of a microdomain at the top of the radial CW in *A. thaliana* [8], leading to a phenotype of defect of mucilage release, also observed in in mutant for *SBT1.7/ARA12* [41]. The presence of orthologs of these genes among species is not correlated to radial CW rupture, nor is their absence correlated to other CW rupture zones (Table S7). When grouping investigated genes by their characterized function categories, in *A. thaliana*, any decrease in the presence of orthologous genes related to mucilage polysaccharides synthesis for non-myxospermous species was observed (Table S8). As the mutations of the MSC toolbox

genes do not have the same phenotypic impact on *A. thaliana*, these genes were also sorted according to their impact, but again, no enrichment of losses was found for these categories (Tables S1 and S9).

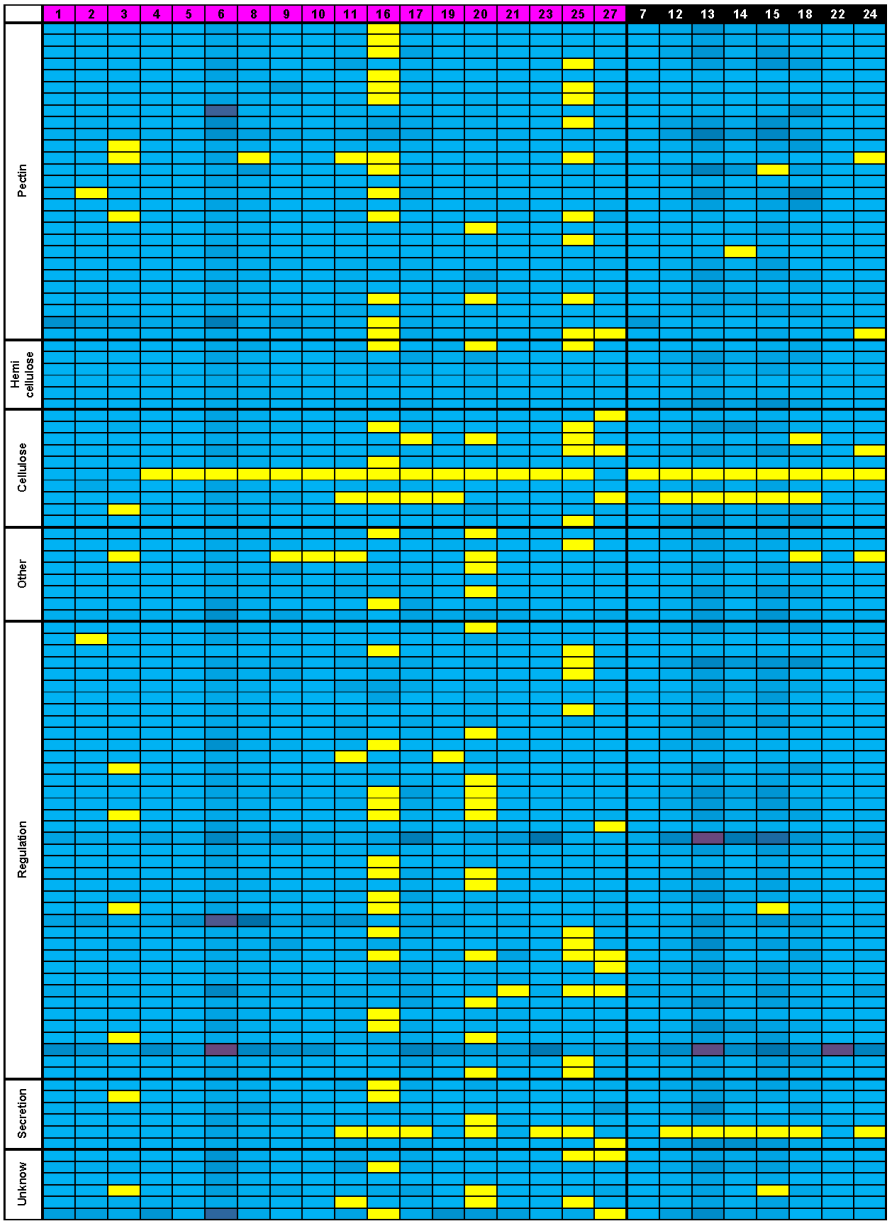


Figure 7. The presence of the *A. thaliana* MSC toolbox gene orthologs throughout the Brassicaceae family is unrelated to their myxospermous ability or lack thereof. Orthologous genes presence (blue) or absence (yellow) is highlighted for each *A. thaliana* MSC toolbox genes (in line) within each studied species (in column). Genes are grouped in functional categories and their numbers for each species due to duplications are symbolized by a colored gradient from light blue (1 gene) to dark purple (23 genes). Species are identified with the number attributed in Figure 1, with myxospermous ones on the left side (in pink) and non-myxospermous ones on the right (in black). See Tables S4–S9 for more details.

Since no clear relationship could be established between the myxospermous traits and the presence of *A. thaliana* MSC toolbox gene orthologs, we analyzed published seed development transcriptomic data from *A. thaliana*, *C. sativa*, *B. napus* and *A. arabicum*. Although these species only represent a fraction of the species targeted in this study, they had the double advantage of being spread among the Brassicaceae phylogenetic

tree (Figure 1) and being either myxospermous (*A. thaliana* (1), *C. sativa* (6), *A. arabicum* (27) dehiscent diasporas) or not (*B. napus* (13), *A. arabicum* (27) non-dehiscent diasporas) (Figure 2). In an effort to unify the heterogeneity of orthologous gene expression data sources from different species, we summed the expression values from all available whole seed development kinetic samples for each gene in each species (Table S10). For each species, the summed expression values of orthologous genes of *A. thaliana* MSC toolbox genes (Table S1) were expressed as a percentage of the mean summed value from the 10 most expressed genes used as a reference (Figure 8, Table S10). Interestingly, the common top 10 genes partially occurred among species. The relative expression profile of MSC toolbox genes was clearly higher in *A. thaliana* as compared to the orthologs in other species, regardless of whether they were myxospermous or not. Despite their relatively weak expression profile, a slightly but significantly higher expression profile occurred in the myxospermous *C. sativa* as compared to the non-myxospermous *B. napus*. However, using the recent elegant transcriptomic data from the dimorphic *A. arabicum* seeds [38] (Figure 8, Table S10), the opposite trends occurred.

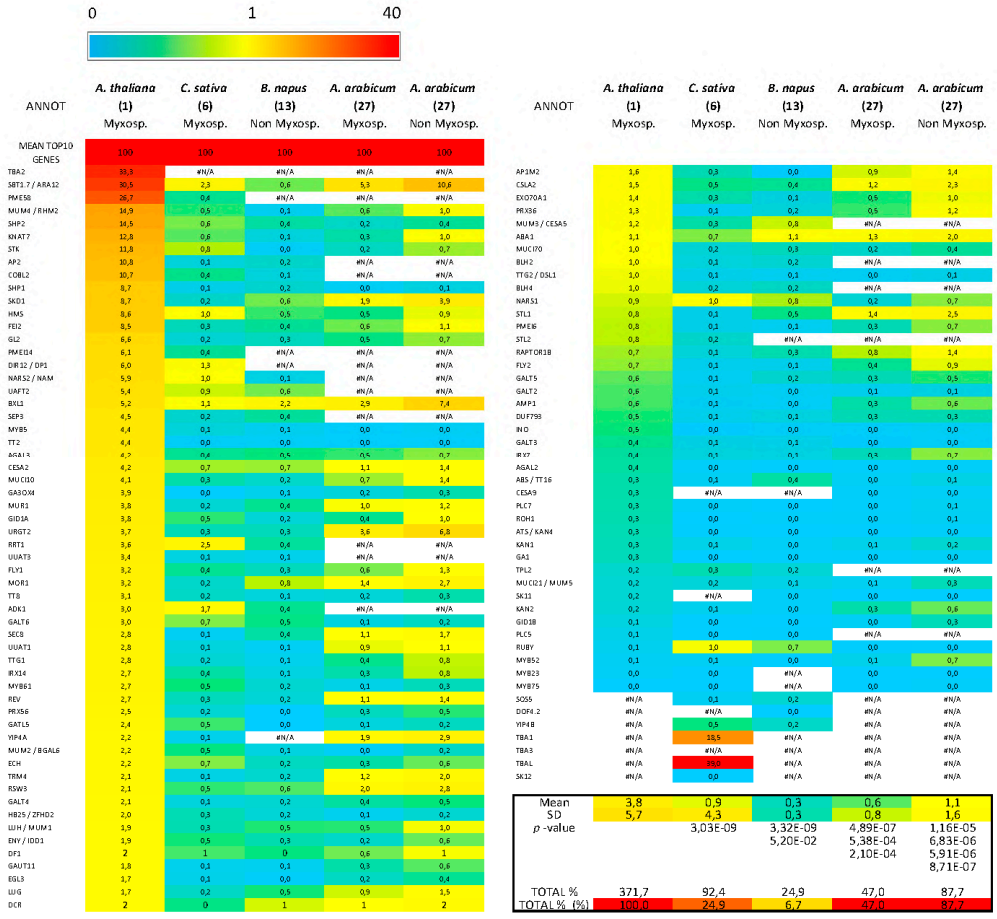


Figure 8. Comparison of the relative expression level of the *A. thaliana* MSC toolbox genes and of their orthologous genes from *C. sativa*, *B. napus*, and *A. arabicum* (myxospermous and non-myxospermous dimorphic diaspores) during whole seed development. The expression value of each *A. thaliana* MSC toolbox gene ortholog was summed for the whole seed development kinetics samples and were displayed in % of the similarly processed 10 most expressed genes in the whole genome for each species. Note that the relative expression values are significantly higher in *A. thaliana* as compared to the three other species. The relative expression values are slightly higher in the myxospermous species *C. sativa* than in the non-myxospermous species *B. napus* but are globally slightly higher in the non-myxospermous seeds than in the myxospermous seeds of the dimorphic species *A. arabicum*. To highlight the differences, a common red (40)-to yellow (1)-to blue (0) heatmap was used. See Table S10 and methods for detailed information and legend.

While considering the significant fold change ratio, 1, 22 and 14 of 105 *A. thaliana* MSC toolbox genes were upregulated in myxospermous vs. non-myxospermous seeds at 0, 1 and 30 DAP, respectively. However, 1, 21 and 17 *A. thaliana* MSC toolbox genes were conversely upregulated in non-myxospermous seeds at 0, 1 and 30 DAP, respectively [38]. Moreover, none of these upregulated MSC toolbox gene orthologs were ranked among the most differentially expressed genes [38]. Therefore, it could be speculated that the *A. thaliana* MSC toolbox genes are rather specific to the specific myxospermous sub-traits observed in *A. thaliana*, and that use of this toolbox to decipher the molecular actors of myxospermy in other species could be tricky.

4. Discussion

4.1. Is Myxospermy an Ancestral Feature of the Brassicaceae Family?

Myxospermy was found in all studied lineages, and importantly, in the two *Aethionema* species belonging to the early divergent group. This could indicate a potential presence of myxospermy in the Brassicaceae family common ancestor. In addition to the two *Aethionema* species studied here, two other *Aethionema* species were also reported to be myxospermous [26], consolidating the myxospermic ability of the early diverging clade of Brassicaceae family. Non-myxospermous species should consequently have undergone independent losses from this ancestral status. At the cellular level, several arguments can also support the hypothesis of multiple independent losses of the ancestral trait. The non-myxospermous lineage 1 species *Boechera stricta* (7) has an epidermis cell layer severely reduced in size, although these cells contain a columella and the species is a close relative to the fully myxospemous Camelineae tribe. Additionally, RU2 antibodies that label the pectinaceous mucilage of every species of lineage 1 do not label *B. stricta* epidermal cells, suggesting a loss of pectinaceous mucilage, while a biochemical change of the polysaccharide composition cannot be excluded. Altogether, it seems that *B. stricta* underwent a de-differentiation of its epidermis cell layer from MSCs to non-MSCs, avoiding the accumulation of hydrophilic polysaccharides and their extrusion, but keeping a columella whose function remains elusive. High diversity of epidermis morphology also occurs within the Brassicaea tribe. The thickest MSCs belong to the two myxospermous mustards, and the non-myxospermous species show variable epidermal cell thickness, up to the thinnest and widest cells observed in *R. sativum* (18). Contrary to *B. stricta* (7), the loss of the myxospermy trait in Brassiceae is linked to a change in epidermal cell shape. However, for Brassiceae species, it is harder to decipher whether a progressive disappearance occurred since their common ancestor, but as they are all cultivated species, it is possible that the myxospermy trait was differentially affected during the selection process, as for the pea seed coat [42]. If myxospermy can be lost through the de-differentiation of the MSC cells, the neo-functionalization of the trait could occur as well thanks to epidermis morphological change. This is the case for Cardamineae species (9–11) that have highly diverse MSC and mucilage morphology, although these are close relative species of Brassiceae, reinforcing the idea that SM can quickly evolve. This is also illustrated in lineage 2 by the highly contrasted phenotype observed between the phylogenetically relatively close myxospermous *S. pinnata* (20) and non-myxospermous *T. arvense* (22). All of these inter-species comparisons highlight how changeable myxospermy is as a trait. Moreover, *A. arabicum* (27) and *Diptychocarpus strictus* (not phenotyped here) were shown to have seeds with dimorphism on myxospermy ability [40,43]. This highlights that seed mucilage can be easily disabled or activated during seed formation without any genetic change and could be rather explained by complex dimorphism-related gene expression [38]. However, we did not observe any correlation between the global expression profile of *A. thaliana* MSC toolbox gene orthologs and dimorphism (Figure 7; Table S10) [38]. The morphological change might be correlated with subtle differential expression between the myxospermous and non-myxospermous seeds for *GL2*, *MYB61* and *SHP2* regulator genes and also more downstream actors, such as *PMEI6*, *CESA2*, *CESA5*, *CSLA2*, and *GALT5* [38]. While gene expression regulation is obviously crucial to allow or deny myxospermy establishment, large-scale analysis of

coding genomic sequences from the *A. thaliana* MSC toolbox gene orthologs may not be appropriate to clear interspecies comparison. When changes in sequence are responsible for myxospermy loss, it can correspond to a deletion in a single gene, as exemplified by *MUM2* polymorphism in Shahdara *A. thaliana* natural population [44]. We had no access to such a resolution in our study but rather searched for stronger pseudogenization that should occur in non-myxospermous species for genes that have no additional function. Consequently, the absence of correlation between the myxospermy traits and the presence/absence of *A. thaliana* MSC toolbox gene orthologs in the studied Brassicaceae species could be due to the pleiotropic functions of the corresponding proteins and/or to their late co-option for mucilage traits in the Brassicaceae ancestor. However, on seven candidate genes highlighted by GWAS on flax seed mucilage content, six genes were orthologous to the *A. thaliana* MSC toolbox genes [24]. This could indicate that some of the toolbox genes can keep their function in mucilage even outside of the Brassicaceae family.

4.2. Arabidopsis Species as an Uncommon Model for Mucilage Secretory Cells

While some myxospermy traits detected with ruthenium red staining or RU2 immunolabeling appear widely distributed along the Brassicaceae phylogeny, other myxospermy traits are more species-specific. Among the studied myxospermous Brassicaceae species, SM extrusion occurs through the center of the MSCs, except for the three studied *Arabidopsis* species. In *A. thaliana*, the weakening of the microdomain localized at the top of radial CW leads to a proper rupture at this location upon seed imbibition [8]. Consequently, for each imbibed MSC, the piece of primary CW that cover MSCs look like an upside-down umbrella attached to the top of a columella [9]. As a columella is located in the center of MSCs, the radial opening could be a consequence of columella apparition for a proper extrusion of SM. However, *C. sativa* (6) and *C. hirsuta* (9) display a central opening as well as a columella presence. Consequently, the columella presence and rupture zone seem to be independent. However, the columella may have appeared first during evolution, as it is present in several species of lineage 1, contrary to the rupture of radial CW, which is restricted to *Arabidopsis* species. Note that for *R. islandica* (10), *L. alabamica* (11) and *S. pinnata* (20), the rupture occurs at the radial position of MSCs but as they have very distinct MSC morphology, it is probably an independent apparition. An ancestor of the Brassicaceae family should probably have MSCs without a columella with a central opening upon imbibition for mucilage extrusion that does not correspond to the *Arabidopsis* phenotype.

4.3. Ecological Perspective

The diversity of the observed myxospermy traits could be linked to ecological roles. The mucilage morphology of *L. alabamica* (11) and its absence in *B. stricta* (7) are interesting. Both species are the only ones of lineage 1 with flattened and winged seeds, and they show a unique mucilage phenotype and an absence of myxospermy, respectively. Both species can indicate the negative effect of cohesive mucilage on wind dispersion for winged seeds. *Dyptiocarpus strictus*, a Brassicaceae species not included in our study, confirms this incompatibility with a seed dimorphism [43]. One morphotype has cohesive mucilage around the seed but a reduced wing, while the other lacks mucilage but has a large wing. *L. alabamica* (11) mucilage could be studied for its physical influence on the seed adherence and floatability because its filamentous shape can be linked to the co-occurrence of wind and another kind of dissemination avoiding staying glued or sinking too fast.

Altogether, the present study documents the high diversity of myxospermy traits and shows the limits of using *A. thaliana* as the unique reference for myxospermy evodevo studies. This might be overcome by considering the increasing number of genomic and transcriptomic studies targeted on non-model species that could bring some of these to the status of model species.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/cells10092470/s1, Figure S1: Ruthenium red and calcofluor double staining for the 19 myxospermous species highlights the cellulosic part contained in seed mucilage, Figure S2: Im-

munofluorescence survey of the putative occurrence of a topochemical signature of the cell wall domain rupture zone in dry and imbibed seeds using the LM20 antibody, Table S1: *A. thaliana* MSC toolbox genes, Table S2: List and seed provider of the 27 species phenotyped in this study, Table S3: List and Genomics data origin of the 32 species used in this study, Table S4: Detailed genomic data mining of *A. thaliana* MSC toolbox gene orthologs among Brassicaceae species, Table S5: Sorting of the studied species according to the myxospermy occurrence highlights the absence of correlation with the presence of *A. thaliana* MSC tool box gene orthologs related to mucilage synthesis in *Arabidopsis*, Table S6: Sorting of the studied species according to the columella occurrence highlights the absence of correlation with the presence of any *A. thaliana* MSC tool box gene orthologs, Table S7: Sorting of the studied species according to their opening zone of the primary cell wall that allows the mucilage extrusion highlights the absence of correlation with orthologs from the genes controlling the opening mode in *A. thaliana*, Table S8: Sorting of studied genes according to their polysaccharide-related functions highlights the absence of correlation with the occurrence of myxospermy, Table S9: Sorting of studied genes according to their impact in myxospermy (strength of the phenotype in the corresponding mutant) highlights the absence of correlation with the occurrence of myxospermy, Table S10 (Related to Figure 8): Comparison of the relative expression values of *A. thaliana* MSC tool box gene orthologs between *A. thaliana*, *C. sativa*, *B. napus* and *A. arabicum* whole seed development illustrates their higher relative expression values in *A. thaliana*.

Author Contributions: Conceptualization, C.D. and V.B.; methodology, S.V., C.D. and V.B.; validation, S.V., C.D. and V.B.; formal analysis, S.V.; investigation, S.V.; resources, C.D. and V.B.; data curation, S.V., C.D. and V.B.; writing—original draft preparation, S.V.; writing—review and editing, S.V., C.D. and V.B.; visualization, S.V.; supervision, C.D. and V.B.; project administration, C.D. and V.B.; funding acquisition, C.D. and V.B. All authors have read and agreed to the published version of the manuscript.

Funding: This work was granted by Paul Sabatier Toulouse 3 University and Centre National de la Recherche Scientifique (CNRS). The PhD scholarship of S.V. was granted by Paul Sabatier Toulouse 3 University. This work was also supported by the French Laboratory of Excellence project "TULIP" (ANR-10-LABX-41; ANR-11-IDEX-0002-02) and funded by the French National Research Agency project "MicroWall" (ANR-18-CE20-0007).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: The microscopy slide scanning was performed at the FRAIB imaging platform of TRI (https://trigenotoul.com/en/ (accessed on 30 August 2021)). We also thank Jean Keller and Helen San-Clement for their help in the computational biology and Nathalie Séjalon-Delmas, Outi Savolainen, Hélène Frérot, Tanja Slotte, Kemal Melik Taskin, Jeremiah Bush, Maheshi Dassanayake and Klaus Mummenhoff for providing seeds.

Conflicts of Interest: The authors declare no conflict of interest.

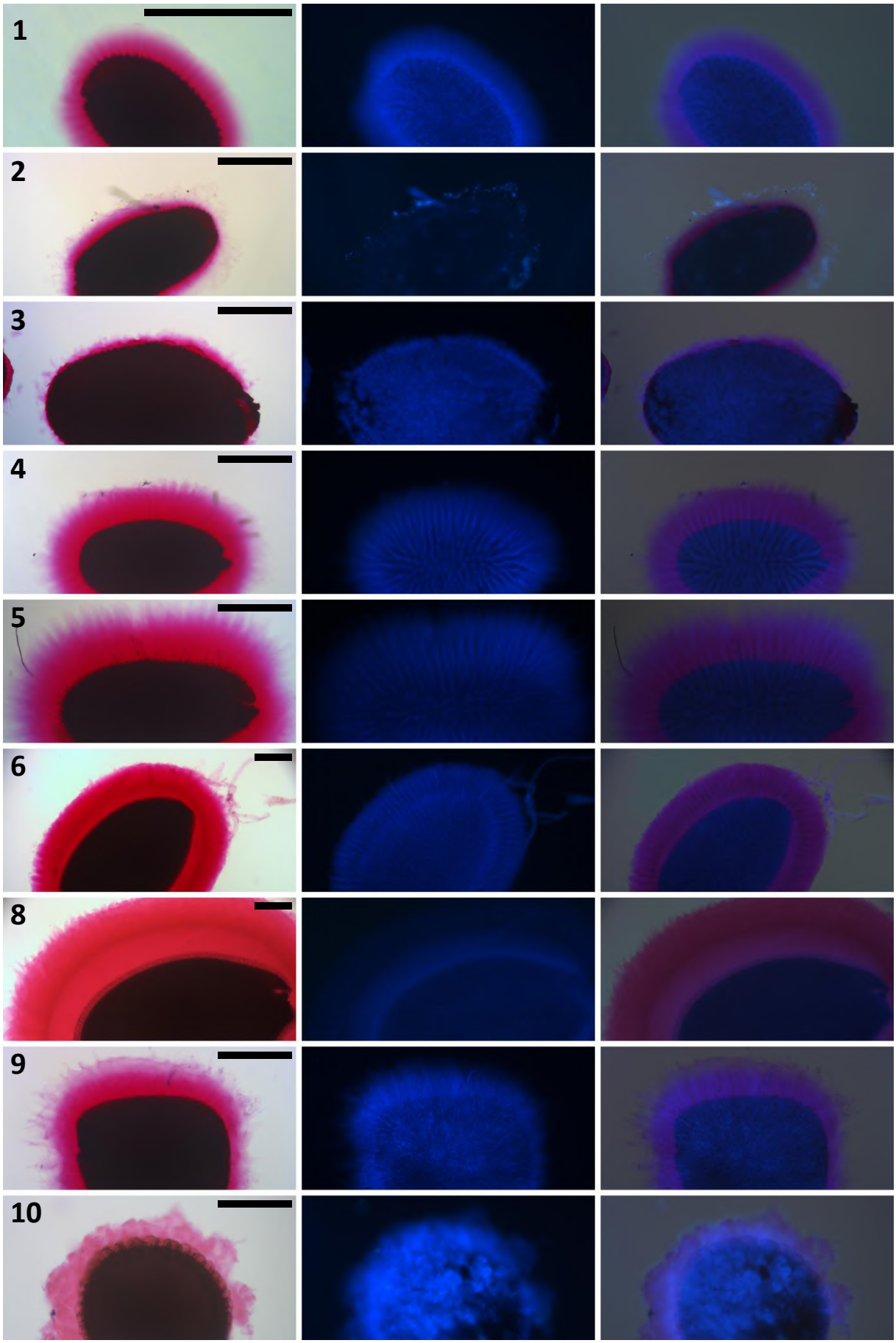
References

1. Doyle, J.A. Phylogenetic analyses and morphological innovations in land plants. *Annu. Plant Rev.* **2017**, *45*, 1–50.
2. Baroux, C.; Grossniklaus, U. Seeds—An evolutionary innovation underlying reproductive success in flowering plants. *Curr. Top. Dev. Biol.* **2019**, *131*, 605–642.
3. Coen, O.; Magnani, E. Seed coat thickness in the evolution of Angiosperms. *Cell. Mol. Life Sci.* **2018**, *75*, 2509–2518. [CrossRef] [PubMed]
4. Beeckman, T.; De Rycke, R.; Viane, R.; Inzé, D. Histological study of seed coat development in *Arabidopsis thaliana*. *J. Plant Res.* **2000**, *113*, 139–148. [CrossRef]
5. Western, T.L. Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis. *Plant Physiol.* **2001**, *127*, 998–1011. [PubMed]
6. Voiniciuc, C.; Yang, B.; Schmidt, M.H.W.; Günl, M.; Usadel, B. Starting to gel: How Arabidopsis seed coat epidermal cells produce specialized secondary cell walls. *Int. J. Mol. Sci.* **2015**, *16*, 3452–3473. [CrossRef]
7. Francoz, E.; Ranocha, P.; Burlat, V.; Dunand, C. Arabidopsis seed mucilage secretory cells: Regulation and dynamics. *Trends Plant Sci.* **2015**, *20*, 515–524. [CrossRef]

8. Francoz, E.; Ranocha, P.; Le Ru, A.; Martinez, Y.; Fourquaux, I.; Jauneau, A.; Dunand, C.; Burlat, V. Pectin demethylesterification generates platforms that anchor peroxidases to remodel plant cell wall domains. *Dev. Cell* **2019**, *48*, 261–276. [CrossRef]
9. Šola, K.; Dean, G.H.; Haughn, G.W. Arabidopsis seed mucilage: A specialised extracellular matrix that demonstrates the structure–function versatility of cell wall polysaccharides. *Annu. Plant Rev.* **2019**, *2*, 1085–1116.
10. Tsai, A.Y.-L.; Kunieda, T.; Rogalski, J.; Foster, L.J.; Ellis, B.E.; Haughn, G.W. Identification and characterization of Arabidopsis seed coat mucilage proteins. *Plant Physiol.* **2017**, *173*, 1059–1074. [CrossRef]
11. Griffiths, J.S.; Šola, K.; Kushwaha, R.; Lam, P.; Tateno, M.; Young, R.; Voiniciuc, C.; Dean, G.; Mansfield, S.D.; Debolt, S.; et al. Unidirectional movement of cellulose synthase complexes in Arabidopsis seed coat epidermal cells deposit cellulose involved in mucilage extrusion, adherence, and ray formation. *Plant Physiol.* **2015**, *168*, 502–520. [CrossRef]
12. Macquet, A.; Ralet, M.C.; Kronenberger, J.; Marion-Poll, A.; North, H.M. In situ, chemical and macromolecular study of the composition of *Arabidopsis thaliana* seed coat mucilage. *Plant Cell Physiol.* **2007**, *48*, 984–999. [CrossRef] [PubMed]
13. Ralet, M.C.; Tranquet, O.; Poulain, D.; Moïse, A.; Guillon, F. Monoclonal antibodies to rhamnogalacturonan I backbone. *Planta* **2010**, *231*, 1373–1383. [CrossRef] [PubMed]
14. Saez-Aguayo, S.; Rautengarten, C.; Temple, H.; Sanhueza, D.; Ejsmentewicz, T.; Sandoval-Ibañez, O.; Doñas, D.; Parra-Rojas, J.P.; Ebert, B.; Lehner, A.; et al. UUAT1 is a golgi-localized UDP-uronic acid transporter that modulates the polysaccharide composition of Arabidopsis seed mucilage. *Plant Cell* **2017**, *29*, 129–143. [CrossRef] [PubMed]
15. Phan, J.L.; Burton, R.A. New insights into the composition and structure of seed mucilage. *Annu. Plant Rev.* **2018**, *1*, 63–104.
16. Viudes, S.; Burlat, V.; Dunand, C. Seed mucilage evolution: Diverse molecular mechanisms generate versatile ecological functions for particular environments. *Plant Cell Environ.* **2020**, *43*, 2857–2870. [CrossRef]
17. Yang, X.; Baskin, J.M.; Baskin, C.C.; Huang, Z. More than just a coating: Ecological importance, taxonomic occurrence and phylogenetic relationships of seed coat mucilage. *Perspect. Plant Ecol. Evol. Syst.* **2012**, *14*, 434–442. [CrossRef]
18. North, H.M.; Berger, A.; Saez-Aguayo, S.; Ralet, M.C. Understanding polysaccharide production and properties using seed coat mutants: Future perspectives for the exploitation of natural variants. *Ann. Bot.* **2014**, *114*, 1251–1263. [CrossRef]
19. Phan, J.L.; Tucker, M.R.; Khor, S.F.; Shirley, N.; Lahnstein, J.; Beahan, C.; Bacic, A.; Burton, R.A. Differences in glycosyltransferase family 61 accompany variation in seed coat mucilage composition in Plantago spp. *J. Exp. Bot.* **2016**, *67*, 6481–6495. [CrossRef]
20. Cowley, J.M.; Burton, R.A. The goo-d stuff: Plantago as a myxospermous model with modern utility. *New Phytol.* **2021**, *229*, 1917–1923. [CrossRef]
21. Voiniciuc, C.; Zimmermann, E.; Schmidt, M.H.-W.; Günl, M.; Fu, L.; North, H.M.; Usadel, B. Extensive natural variation in Arabidopsis seed mucilage structure. *Front. Plant Sci.* **2016**, *7*, 803. [CrossRef]
22. Poulain, D.; Botran, L.; North, H.M.; Ralet, M.-C. Composition and physicochemical properties of outer mucilage from seeds of Arabidopsis natural accessions. *AoB Plants* **2019**, *11*, plz031. [CrossRef]
23. Liu, J.; Shim, Y.Y.; Shen, J.; Wang, Y.; Ghosh, S.; Reaney, M.J.T. Variation of composition and functional properties of gum from six Canadian flaxseed (*Linum usitatissimum* L.) cultivars. *Int. J. Food Sci. Technol.* **2016**, *51*, 2313–2326. [CrossRef]
24. Soto-Cerda, B.J.; Cloutier, S.; Quian, R.; Gajardo, H.A.; Olivos, M.; You, F.M. Genome-wide association analysis of mucilage and hull content in flax (*Linum usitatissimum* l.) seeds. *Int. J. Mol. Sci.* **2018**, *19*, 2870. [CrossRef] [PubMed]
25. Fabrissin, I.; Cueff, G.; Berger, A.; Granier, F.; Sallé, C.; Poulain, D.; Ralet, M.-C.; North, H.M. Natural variation reveals a key role for rhamnogalacturonan I in seed outer mucilage and underlying genes. *Plant Physiol.* **2019**, *181*, 1498–1518. [CrossRef] [PubMed]
26. Vaughan, J.G.; Whitehouse, J.M. Seed structure and the taxonomy of the Cruciferae. *Bot. J. Linn. Soc.* **1971**, *64*, 383–409. [CrossRef]
27. Western, T.L. The sticky tale of seed coat mucilages: Production, genetics, and role in seed germination and dispersal. *Seed Sci. Res.* **2012**, *22*, 1–25. [CrossRef]
28. Hohmann, N.; Wolf, E.M.; Lysak, M.A.; Koch, M.A. A time-calibrated road map of Brassicaceae species radiation and evolutionary history. *Plant Cell* **2015**, *27*, 2770–2784. [CrossRef]
29. Huang, C.H.; Sun, R.; Hu, Y.; Zeng, L.; Zhang, N.; Cai, L.; Zhang, Q.; Koch, M.A.; Al-Shehbaz, I.; Edger, P.P.; et al. Resolution of Brassicaceae phylogeny using nuclear genes uncovers nested radiations and supports convergent morphological evolution. *Mol. Biol. Evol.* **2016**, *33*, 394–412. [CrossRef] [PubMed]
30. Nikolov, L.A.; Shushkov, P.; Nevado, B.; Gan, X.; Al-Shehbaz, I.A.; Filatov, D.; Bailey, C.D.; Tsiantis, M. Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. *New Phytol.* **2019**, *222*, 1638–1651. [CrossRef] [PubMed]
31. Clausen, M.H.; Willats, W.G.T.; Knox, J.P. Synthetic methyl hexagalacturonate hapten inhibitors of anti-homogalacturonan monoclonal antibodies LM7, JIM5 and JIM7. *Carbohydr. Res.* **2003**, *338*, 1797–1800. [CrossRef]
32. Verhertbruggen, Y.; Marcus, S.E.; Haeger, A.; Ordaz-Ortiz, J.J.; Knox, J.P. An extended set of monoclonal antibodies to pectic homogalacturonan. *Carbohydr. Res.* **2009**, *344*, 1858–1862. [CrossRef]
33. Solovyev, V.; Kosarev, P.; Seledsov, I.; Vorobyev, D. Automatic annotation of eukaryotic genes, pseudogenes and promoters. *Genome Biol.* **2006**, *7*, S10. [CrossRef]
34. Belmonte, M.F.; Kirkbride, R.C.; Stone, S.L.; Pelletier, J.M.; Bui, A.Q.; Yeung, E.C.; Hashimoto, M.; Fei, J.; Harada, C.M.; Munoz, M.D.; et al. Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E435–E444. [CrossRef]
35. Abdullah, H.M.; Akbari, P.; Paulose, B.; Schnell, D.; Qi, W.; Park, Y.; Pareek, A.; Dhankher, O.P. Transcriptome profiling of *Camelina sativa* to identify genes involved in triacylglycerol biosynthesis and accumulation in the developing seeds. *Biotechnol. Biofuels* **2016**, *9*, 136. [CrossRef] [PubMed]

36. Kagale, S.; Nixon, J.; Khedikar, Y.; Pasha, A.; Provart, N.J.; Clarke, W.E.; Bollina, V.; Robinson, S.J.; Coutu, C.; Hegedus, D.D.; et al. The developmental transcriptome atlas of the biofuel crop *Camelina sativa*. *Plant J.* **2016**, *88*, 879–894. [CrossRef] [PubMed]
37. Wan, H.; Cui, Y.; Ding, Y.; Mei, J.; Dong, H.; Zhang, W.; Wu, S.; Liang, Y.; Zhang, C.; Li, J.; et al. Time-series analyses of transcriptomes and proteomes reveal molecular networks underlying oil accumulation in canola. *Front. Plant Sci.* **2017**, *7*, 2007. [CrossRef] [PubMed]
38. Arshad, W.; Lenser, T.; Wilhelmsson, P.K.I.; Chandler, J.O.; Steinbrecher, T.; Marone, F.; Pérez, M.; Collinson, M.E.; Stuppy, W.; Rensing, S.A.; et al. A tale of two morphs: Developmental patterns and mechanisms of seed coat differentiation in the dimorphic diaspore model *Aethionema arabicum* (Brassicaceae). *Plant J.* **2021**, *107*, 166–181. [CrossRef] [PubMed]
39. Kagale, S.; Koh, C.; Nixon, J.; Bollina, V.; Clarke, W.E.; Tuteja, R.; Spillane, C.; Robinson, S.J.; Links, M.G.; Clarke, C.; et al. The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. *Nat. Commun.* **2014**, *5*, 3706. [CrossRef] [PubMed]
40. Lenser, T.; Graeber, K.; Cevik, Ö.S.; Adigüzel, N.; Dönmez, A.A.; Grosche, C.; Kettermann, M.; Mayland-Quellhorst, S.; Mérai, Z.; Mohammadin, S.; et al. Developmental control and plasticity of fruit and seed dimorphism in *Aethionema arabicum*. *Plant Physiol.* **2016**, *172*, 1691–1707. [CrossRef]
41. Rautengarten, C.; Usadel, B.; Neumetzler, L.; Hartmann, J.; Büssis, D.; Altmann, T. A subtilisin-like serine protease essential for mucilage release from Arabidopsis seed coats. *Plant J.* **2008**, *54*, 466–480. [CrossRef] [PubMed]
42. Hradilová, I.; Trněný, O.; Válková, M.; Cechová, M.; Janská, A.; Prokešová, L.; Aamir, K.; Krezdorn, N.; Rotter, B.; Winter, P.; et al. A combined comparative transcriptomic, metabolomic, and anatomical analyses of two key domestication traits: Pod dehiscence and seed dormancy in Pea (Pisum sp.). *Front. Plant Sci.* **2017**, *8*, 542. [CrossRef] [PubMed]
43. Lu, J.; Tan, D.; Baskin, J.M.; Baskin, C.C. Fruit and seed heteromorphism in the cold desert annual ephemeral *Diptychocarpus strictus* (Brassicaceae) and possible adaptive significance. *Ann. Bot.* **2010**, *105*, 999–1014. [CrossRef] [PubMed]
44. Macquet, A.; Ralet, M.-C.; Loudet, O.; Kronenberger, J.; Mouille, G.; Marion-Poll, A.; North, H.M. A naturally occurring mutation in an Arabidopsis accession affects a β-d-Galactosidase that increases the hydrophilic potential of rhamnogalacturonan I in seed mucilage. *Plant Cell* **2007**, *19*, 3990–4006. [CrossRef] [PubMed]

Viudes et al Fig. S1 (2 pages; legend on page 2)



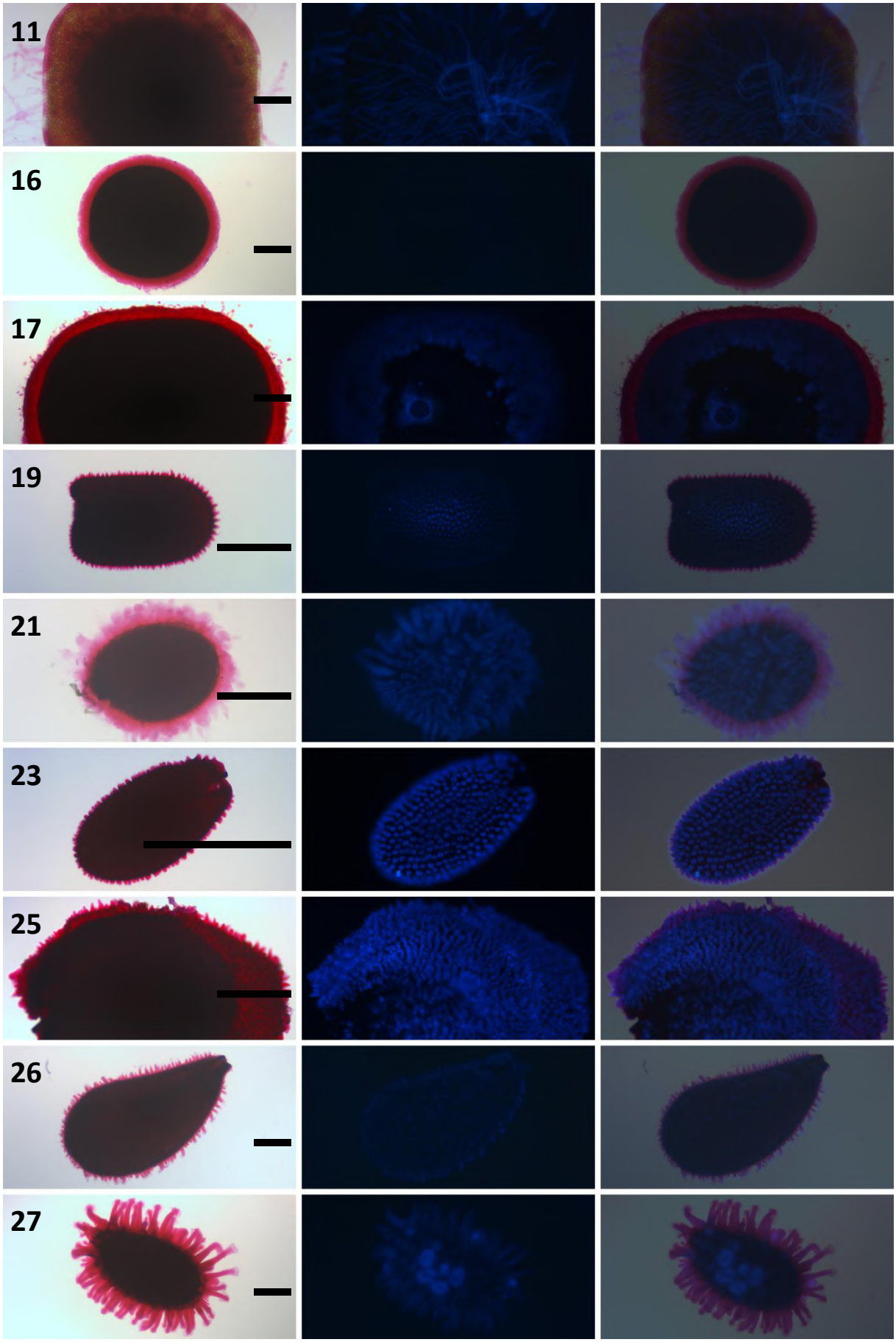
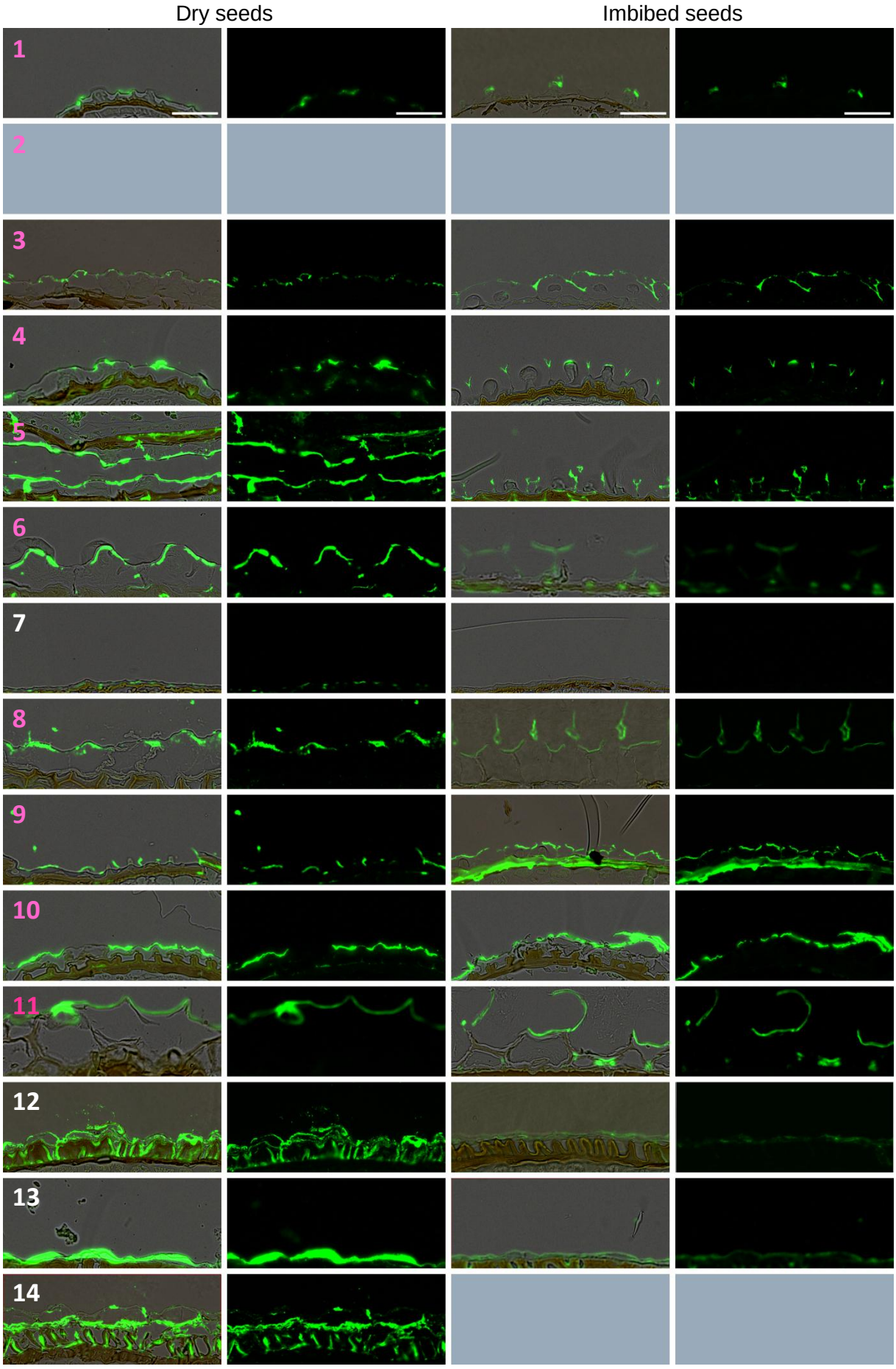


Figure S1: Ruthenium red and calcofluor double staining for the 19 myxospermous species highlights the cellulosic part contained in seed mucilage. Species numbers presented in Fig. **1** were used to label the images (only the myxospermous species are shown here). Column 1 images correspond to the ruthenium red staining after vigorous shaking shown in Fig. **2**. Column 2 shows the same seed under UV light to reveal the calcofluor labelling which is specific to cellulose. Column 3 is the merge of column 1 and 2 with a transparency set up at 70% for black background images. Bars: 500 µm.

Viudes et al Fig. S2 (2 pages; legend on page 2)



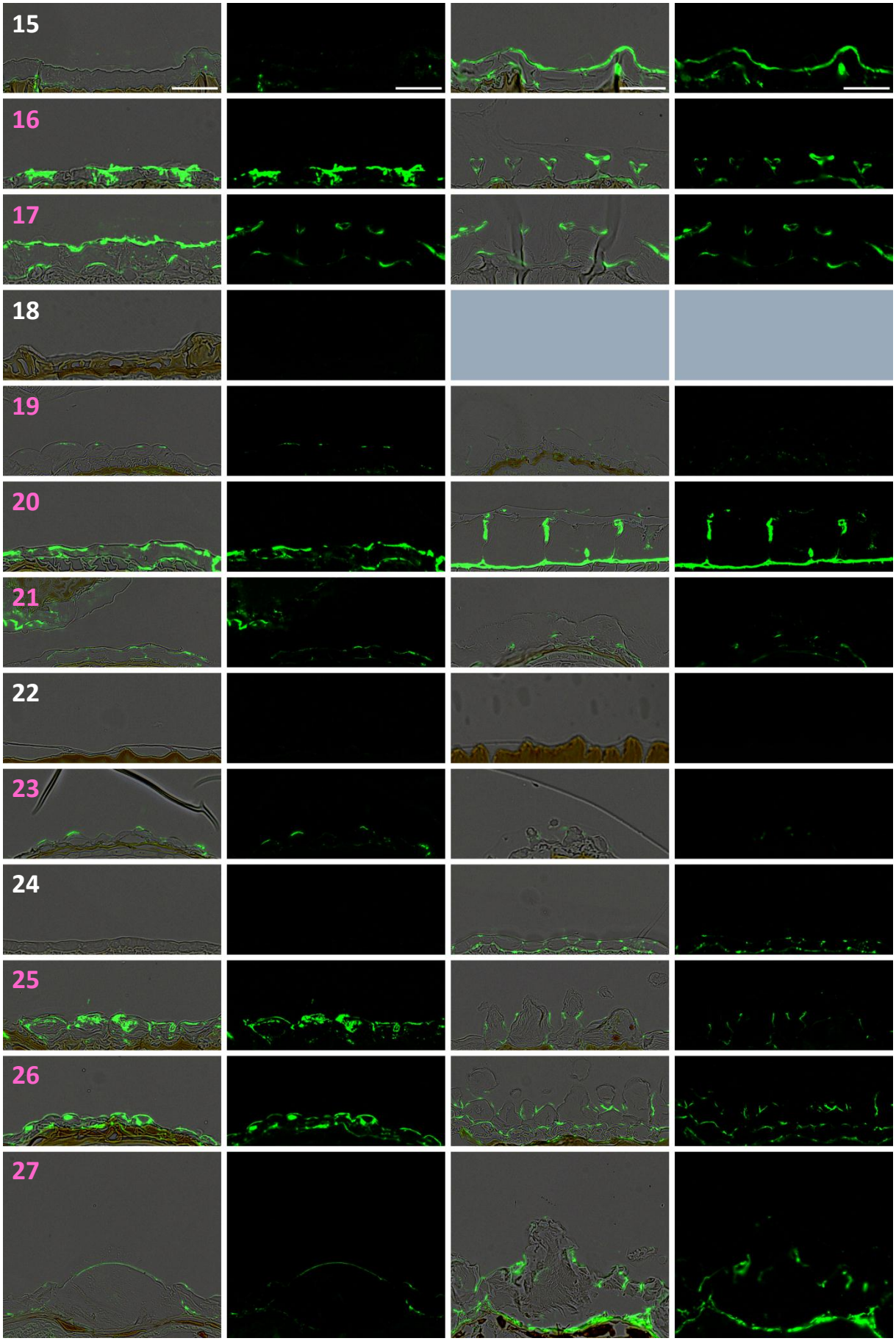


Figure S2: Immunofluorescence survey of the putative occurrence of a topochemical signature of the cell wall domain rupture zone in dry and imbibed seeds using the LM20 antibody. Serial sections of tissue arrays used for toluidine blue staining were used for immunofluorescence labeling with LM20 anti-partially demethylesterified homogalacturonan. This antibody has been previously shown to label the prefragilized cell wall microdomain that is ruptured upon *A. thaliana* seed imbibition (Francoz et al. 2019). Note the recurrent occurrence of punctuated labeling in the epidermal/MSC cell walls. Species numbers defined in Fig. **1** were colored in pink and white for myxospermous and non-myxospermous species, respectively. Grey panels correspond to unavailable or unexploitable samples. Bars: 50 µm.

Transition : de la diversité inter-espèces à la diversité intra-espèce

Malgré l'impressionnante quantité de savoir portant sur le mucilage de la plante modèle *Arabidopsis thaliana* (développement, contrôle génétique, composition biochimique), son rôle écologique reste étonnamment un mystère. Il a pourtant fait l'objet de quelques approches notamment au niveau de la germination des graines, mais les résultats manquent de clarté et de cohérence (cf article de revue § 1.3.2). Cela a donc motivé l'ouverture de cette partie de la thèse. De plus, comme décrit précédemment (chapitre 1), le mucilage des graines est incroyablement variable entre deux espèces même si elles sont très proches phylogénétiquement. Il l'est tellement qu'au sein d'une seule espèce le mucilage des graines peut-être diffèrent entre les populations naturelles qui la composent. Cette variabilité morphologique contient de précieuses informations qui peuvent être associées à la variabilité génétique sous-jacente (identification de nouveaux gènes impliqués dans la fonction) ou être reliées aux conditions environnementales (déduction de quelles conditions ou contraintes cette morphologie est l'adaptation). Ces deux types d'approches menées sur l'espèce modèle *A. thaliana* seront l'objet du chapitre 2 ci-après.

Chapitre 2 : les bases génétiques et écologiques de la variabilité naturelle de mucilage chez *Arabidopsis thaliana*

Préface à l'article de recherche

Dans le contexte d'un projet commun entre Fabrice Roux et Christophe Dunand, les 168 populations naturelles d'*A. thaliana* prélevées par l'équipe de Fabrice Roux (Frachon et al 2018) ont été mises à profit pour investiguer la variabilité intra-espèce du mucilage. La maitrise de techniques de phénotypage haut débit du mucilage des graines ainsi que les données génétiques et écologiques disponibles pour chacune de ces populations rendaient ce projet particulièrement pertinent. La grande plus-value de ces 168 populations, comparée aux populations utilisées jusqu'alors dans la recherche provient de l'aspect unique de certaines données écologiques disponibles : les paramètres pédologiques et microbiologiques du sol (Bartoli et al 2018, Frachon et al 2019). De plus, l'équipe de Fabrice Roux maitrise les outils statistiques nécessaires à la mise en place de génomique d'association à l'échelle du génome entier (GWAS) qui permettent d'identifier les régions génomiques corrélées à la variabilité morphologique observée. Le chapitre qui suit est sous la forme d'un article en anglais même si il n'est pas encore prêt à être soumis à un journal. La méthode utilisé pour la GWAS est innovante et n'a pas encore été publiée. Elle mériterait donc une description exhaustive pour sa valorisation. Je n'ai malheureusement pas pu encore échanger suffisamment avec Fabrice Roux pour en comprendre suffisamment et ainsi écrire par moi-même la partie matériel et méthode correspondante. Par ailleurs cet article pourrait gagner en impact par l'ajout d'une ou deux caractérisations fonctionnelles des gènes candidats les plus prometteurs et/ou de validation d'une ou plusieurs hypothèses provenant des statistiques corrélatives entre le mucilage et les paramètres environnementaux. Dépendant de la stratégie de publication l'article pourrait aussi être publié sans résultats additionnels et les deux ajouts précédemment évoqués faire l'objet de deux publications indépendantes. Contrairement au premier chapitre, je n'ai pas contribué à l'intégralité des travaux présentés ci-après. La production des graines ainsi que la confrontation des phénotypes aux données de polymorphismes nucléotidiques par génétique d'association ont été réalisées par Fabrice Roux et son équipe. De plus Werner Schweyer a contribué au phénotypage des populations lors de son stage de licence professionnelle que j'ai co-encadré.

Genome wide association study of Arabidopsis seed mucilage layers at the micro-geographic scale

Sébastien Viudes[1], Werner Schweyer[1], Vincent Burlat[1], Fabrice Roux[2], Christophe Dunand[1],

[1] Laboratoire de Recherche en Sciences Végétales, CNRS, INP, UPS, Université de Toulouse, 24 chemin de Borde Rouge, 31320, Auzeville-Tolosane, France

[2] Laboratoire des Interactions Plantes-Microorganismes, INRAE, CNRS, Université de Toulouse, Castanet-Tolosan, France

Abstract

The myxospermous species *Arabidopsis thaliana* extrudes a polysaccharidic mucilage from the seed coat epidermis during imbibition. The whole seed mucilage can be divided into a seed-adherent layer and a fully soluble layer. Myxospermy traits were previously shown to be variable for both mucilage layers in the intra-species natural variability. The adherent mucilage can be variable in size and composition while the soluble mucilage had polymorphic composition and physical properties. Here, 166 *A. thaliana* regional populations from south France are phenotyped for their adherent and non-adherent mucilage size. GWAS highlights 55 and 28 candidate genes, corresponding to the 26 and 10 QTLs correlated to adherent and non-adherent mucilage, respectively. One of these gene was previously characterized validating the strategy. Putative or characterized function and expression data available in the literature were used to filter the gene candidates. Previously published climatic, soil composition, and microbial community data related to the present populations were used to investigate statistical correlation with the variation in size of mucilage layers. The identified candidate genes and interesting correlation with micro-geographical parameters are promising basis for future myxospermy-related gene characterisation and to better understand the ecological roles of myxospermy.

Introduction

Arabidopsis thaliana is a model in plant cellular and molecular biology for at least four decades (Meyerowitz, 1987). Its small genome and plant size as well as its short life cycle compared to other plants raised it as the most practical laboratory flowering plant (Krämer, 2015). In non-crop species, the morphology and ecology are among the first explored research fields. However, *A. thaliana* ecological interest only recently increased, way after the deep genomic and developmental knowledge accumulation. *A. thaliana* appears to be a widely spread species across the world, being more reported in occidental Europe, North America and Asia (Krämer, 2015). Thanks to the currently accumulated genetic data of numerous natural populations, it appears that the distribution of *A. thaliana* has been severely concentrated in small geographic areas during the latest glaciation particularly in north Africa and in the Iberian peninsula (Brennan et al., 2014). The subsequent migration of *A. thaliana* throughout Europe coincided with the agricultural propagation, indicating the disseminating role of

human activities (François et al., 2008). The intra-species genetic variability and its reliability to explain morphological traits was highlighted by the 1001 genomes project providing 1135 *A. thaliana* genomes to the scientific community (Weigel and Mott, 2009; Alonso-Blanco et al., 2016). Natural variation is a great resource for the functional characterization of new genes using genome wide association studies (GWAS) (Visscher et al., 2012). This method is particularly adapted to *A. thaliana* thanks to its self-fertilisation (Korte and Farlow, 2013).

Seed mucilage is a polysaccharidic hydrogel extruded around seeds upon imbibition by the outermost epidermal cell layer of the seed coat in myxospermous species (Phan and Burton, 2018). In *A. thaliana*, around one hundred genes are participating to mucilage synthesis and mucilage secretory cell formation (Viudes et al., 2021). A large majority of these genes has been characterized through classical forward and reverse genetic studies. First of these were major regulators or main actors of mucilage establishment that trigger important phenotypes when mutated (Western, 2001). Progressively, discovered genes had more subtle role for myxospermy, displaying faint phenotypes such as specific polysaccharide proportion or decoration changes when mutated. Evidence of seed mucilage variability in natural diversity was first evidenced with the characterisation of an *A. thaliana* natural mutant, originated from Tajikistan, which is not able to extrude mucilage (Macquet et al., 2007a). Within central Asia and Scandinavia among 22 tested *A. thaliana* populations, 7 were reported to be fully defective for their mucilage release, and 2 showed extrusions only for a few accessions constituting the population (Saez-Aguayo et al., 2014). Total seed mucilage can have astonishing polysaccharide composition changes across natural populations, including mannose, galactose, galacturonic acid and rhamnose monosaccharide content (Voiniciuc et al., 2016). *A. thaliana* mucilage has two distinct layers, the inner layer is highly adherent to the seed thanks to cellulosic fibrils, and the outer layer is fully soluble (Tsai et al., 2017). The area of adherent mucilage was shown to be also variable between natural populations probably linked to the cellulose crystalline structure changes observed as well (Voiniciuc et al., 2016). The non-adherent mucilage is also variable for its composition and physical properties (Poulain et al., 2019). Interestingly, the intra-species loss of the myxospermy trait can have evolved several times due to independent genetic mutations into relatively restricted geographical zones (Saez-Aguayo et al., 2014). Two GWAS led on *A. thaliana* adherent and non-adherent seed mucilage highlighted 21 and 11 significant SNPs, respectively, reporting the problem of myxospermy polygenicity (Voiniciuc et al., 2016; Fabrissin et al., 2019). The morphological natural variation can also provide interesting insight to study ecological function of morphological traits through statistical analysis of environmental data of each population harvest site. However, this has not been included yet in myxospermy GWAS. For seed mucilage the most interesting environmental data should be the soil biotic and abiotic parameters as mucilage is the interface between the seed and the underground environment. In the close related species *Capsella bursa-pastoris*, the seed mucilage was shown to be implicated in soil structuration and water retention

capacity (Deng et al., 2014). Additionally, when *A. thaliana* imbibed seeds are co-inoculated with *Streptomyces lividans* (soil bacteria) and *Verticilium dahlia* (plant pathogen fungi), the growth of the former is enhanced as compared to the latter in the adherent mucilage, reducing the disease symptom on the consequently developing plant (Meschke and Schrempf, 2010).

Here, we aimed to use GWAS to further investigate the genetics-underlined natural variability of *A. thaliana* myxospermy. The facts that the natural populations we used were harvested in a narrow geographic area and that abiotic and biotic environmental data were available were two new parameters compared to previous myxospermy-related GWAS. Additionally, we quantitatively phenotyped the adherent mucilage area to increase the precision compared to the semi-quantitative phenotyping performed in previous studies. Non-adherent mucilage abundance was also phenotyped providing the opportunity to study the relationship between the two layers of mucilage. For the presently studied populations, the soil composition and microbiota community composition are already available creating a unique opportunity to further address in the future, the ecological role of myxospermy.

Material and methods

Plant material and experimental setup

A. thaliana natural populations harvested initially in 2015 in the Midi-Pyrénées region (south-west of France) were used (Frachon et al., 2018). The average distance between these populations is 100.6 km. The 168 populations additionally possess already published data on their original biotope climate, leaf and root microbiota, and soil composition (Bartoli et al., 2018). Every population were composed of several individuals (accessions) harvested on the same site. In the present study, 166 out of the 168 populations are used, each one composed of one to three accessions, bringing the total of phenotyped seed lots to 424. Seed lots were obtained from plant cultivated outdoor in the INRA campus of Castanet-Tolosan, sown at the same time. Given the high number of accessions, the experiments were performed once for each seed lots through several experimental batches distributed along a few months. The number of accessions in each batch is dependant of the experiment (see below). For further standardisation of the results, a single seed lot of Col-0 seeds was used as an internal standard systematically included in all experimental batches.

Adherent mucilage and seed area phenotyping

A total of 37 batches of experiments were performed. Each batch of experiment contained eleven different accessions and the Col-0 as a standard. The adherent mucilage is strongly attached to the seed. To be able to statistically compare its area between different accessions in the most standardized manner we used a previously described protocol (Francoz et al., 2019). Briefly, about 100 seeds per accessions were vigorously shaken at 250 rpm in 1.6 ml of 0.01 M Tris-HCl pH7.5 contained in a 2 ml microtube to remove the non-adherent mucilage. The remaining adherent layer of mucilage was then

stained with a 0.02 % (w/v) ruthenium red solution in Tris-HCl. The seeds were rinsed in Tris-HCl, transferred in 24-well microplates in Tris-HCl and imaged using a Epson perfection V100 photo scanner at 6400 dpi resolution. The measurements of seed area including or not the adherent mucilage were realised on imageJ using a previously described macro (Francoz et al., 2019). The non-adherent mucilage area was calculated for each seed by subtracting the two previous measurements.

Non-adherent mucilage area phenotyping

A total of 98 batches of experiments were performed. Each batch of experiment was composed of three or six different accessions plus the Col-0 as standard (about 9 seeds each). The non-adherent mucilage is totally soluble in water and quickly released upon seed imbibition. To estimate its area, we used the MuSeeQ protocol (Miart et al., 2018). Briefly, about ten dry seeds by accession are individually and synchronously disposed on perfectly flat 0.6 % (w/v) agarose media containing 0.00004 % (w/v) of toluidine blue O (Sigma-Aldrich) polychromatic stain. Upon the contact with the aqueous media, the non-adherent mucilage is released, spreading around the seed on the flat surface, and is simultaneously stained in pink contrasting with the blue colour of the media. Top-view pictures were taken 24 h after seed deposition with a camera (Canon DS 126271). The non-adherent mucilage area is measured for each seed using imageJ with a previously described macro (Miart et al., 2018).

GWA mapping with local score analysis

The GWAS was performed either with the adherent or non-adherent mucilage areas. The seed area was used as a cofactor only for the adherent-mucilage GWAS. The GWA method used for this study is adapted from Aoun et al., (2020) using the local score method (Bonhomme et al., 2019). For significant SNPs, the genes located at this position and within an interval of 2-kb downstream and upstream were reported. Full list of QTLs and their contained SNPs are available in Table S1 and S2, for adherent and non-adherent mucilage, respectively.

Transcriptomic profiling of candidate genes

The expression profiles of the candidate genes were extracted from *A. thaliana* tissue-specific seed development transcriptomic data (Belmonte et al., 2013). Whole transcriptomic data set used to decipher the seed coat specific expression of the candidate genes is available in Table S3.

Statistical analysis and figure building

Every figure in the present article is built using RStudio (version 1.1.463) with the ggplot2 and corrplot packages. Correlations between measured parameters are Pearson correlations. The correlations with environmental data were obtained with the rcorr function from the Hmisc package using spearman's rho.

Results and Discussion

Measured parameters and accessions variability after standardization

Three parameters coming from measurement on 2D images were obtained for each of the 424 accessions: the area of the seed, of the adherent and of the non-adherent mucilage. As the seed and the adherent mucilage are 3D object, their measured values correspond to orthogonal projections. These projections are coherent with the reality since the seeds systematically lay down longitudinally in solution due to their oblong shape. Non-adherent mucilage spreads on a flat surface and is consequently already a 2D object before the imaging. Considering the large number of accessions, the phenotyping was divided randomly in different experimental batches. In each batch, the same Columbia (Col-0) seed lots were used as an internal control for further standardization and to reduce a potential batch effect (Fig. 1A & 1B). Col-0 seed measurements repeated in 44 batch show that the adherent mucilage is a relatively more variable feature compared to the seed size (Fig. 1A). The high reproducibility of the Col-0 seed area values confirms the robustness of the imaging measurement and consequently indicates that the adherent mucilage variation observed in Col-0 is linked to technical variability probably corresponding to subtle differences in staining contrast leading to subtle differences in mask adjustments while using the imaging macro (Fig. S1). The non-adherent mucilage is soluble making it more difficult to be quantified in a high throughput manner (Fig. 1B; Fig. S2).

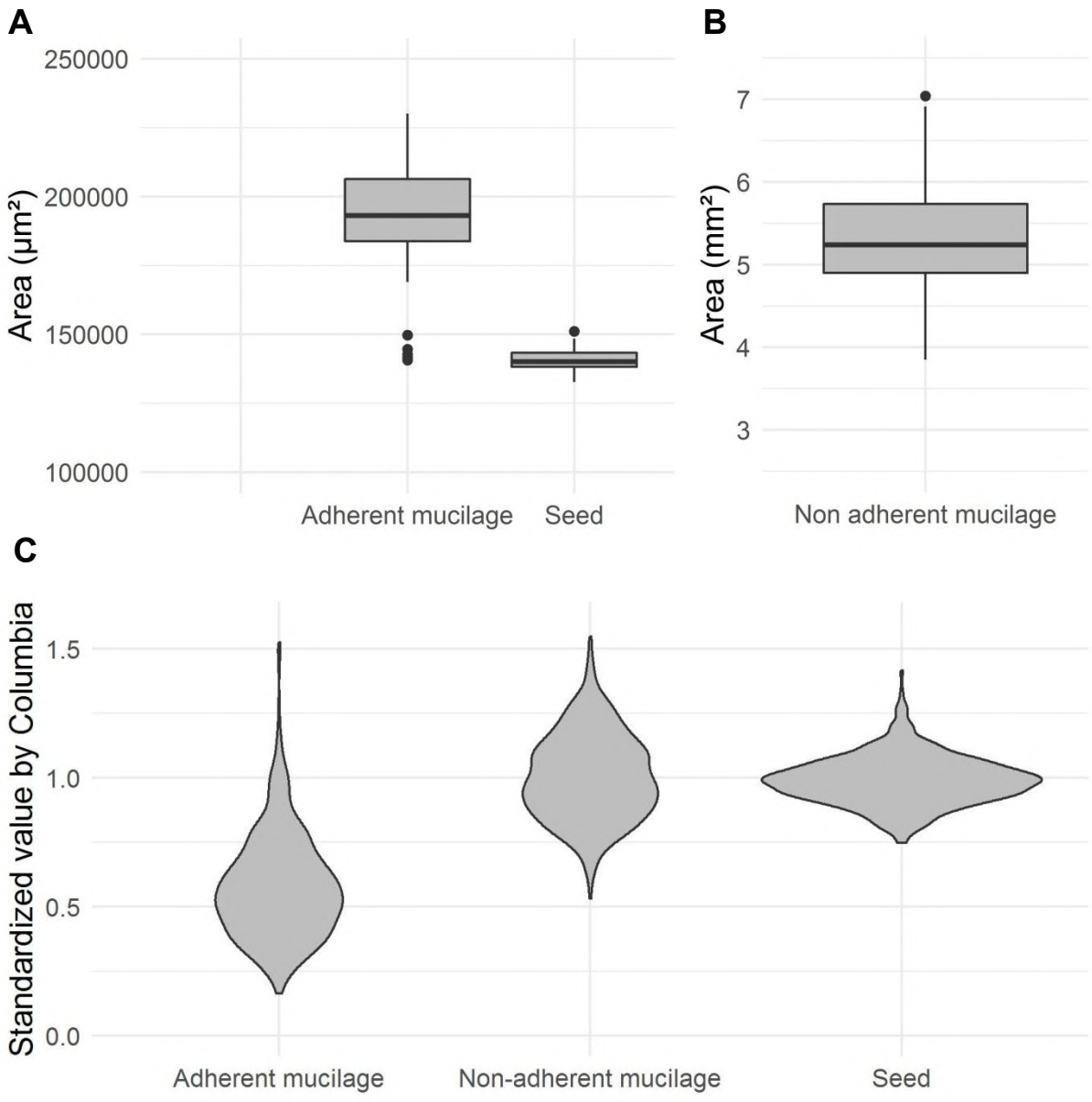


Figure 1: Relatively low variation of the three parameters measured on the Col-0 standard repeated in each batch and repartition of the observed variability of the studied natural populations phenotypes after normalization with the internal Col-0 standard. (A) Boxplot showing mean values of seed area and adherent mucilage area of the Col-0 seeds used as a standard in the 44 individual experimental batches during the phenotyping of studied populations. The area is expressed in µm². **(B)** Boxplot showing mean values of non-adherent mucilage area of the Col-0 seeds used as a standard in the 98 batches of experiment during the phenotyping of studied populations. The area is expressed in mm². **(C)** Violin plot showing mean values of seed, adherent mucilage, and non-adherent mucilage area normalized with the Col-standard (y axis) for each of the 424 accessions. Area values are expressed relatively to the mean area of Col-0 seeds contained in each respective experimental batch. The x axis wideness of violin plot shows the relative frequency of accessions for the corresponding value displayed on the y axis.

This parameter is relatively more variable between the 98 internal Col-0 control repeats probably due to technical variations (Fig. 1B; Fig. S2). Despite these relative variations of the adherent and non-

adherent mucilage areas of Col-0 internal standard, the adherent mucilage area is highly different among the 424 accessions and overall smaller than those of Col-0 (Fig. 1A, C). A previous study has already noticed the larger size of the Col-0 adherent mucilage capsule compared to those of other natural populations (Voiniciuc et al., 2016). The adherent-mucilage area of the accessions is distributed between 0.2 times to more than 1.5 times compared to those of Col-0 (Fig. 1C). In comparison, the non-adherent mucilage area shows less variability between populations and is more homogeneously spread from 0.5 to 1.5 times compared to the Col-0 reference (Fig. 1C). Overall, the variation of the Col-0 repeats was clearly lower than the variation observed among the accessions for both mucilage layers (Fig. S1 and S2). Finally, the seed area of the 424 accession populations is not very different from Col-0 and is symmetrically distributed compared to the reference (Fig. 1C).

Table 1: Global statistics for the used GWAS model shows that variance is surprisingly more explained by morphological variability between the 424 accessions than between the 166 populations.

	Adherent mucilage area			Non adherent mucilage area		
Terms	***F* or LRT**	***P***	***R²***	***F* or LRT**	***P***	***R²***
Population	2,9	3,2E-14	32,3	1,82	1,1E-05	13,8
Accession	30514,8	1,0E-16	42,7	642,3	1,0E-16	34,1
Seed area	13811,8	1,0E-32	-	ne	ne	ne
Residuals			25,0			52,1

Unexpected intra-population variability

Each population is represented by one to three individual accessions originating from a relatively restricted area (Frachon et al., 2018). Consequently, individual accessions of one population were not expected to have important differences for their mucilage phenotypes. In reality, some populations can be very homogeneous for adherent or non-adherent mucilage areas, but other one can contain much contrasted accessions for their seed mucilage layers, as exemplified for 4 populations (Fig. 2). Overall, there is surprisingly more variance between accessions than between populations (Table 1). 32.3 % of the adherent mucilage variance is explained by morphological variability between populations whereas 42.7 % is due to the variability between accessions. For non-adherent mucilage it is even more striking since 13.8 % and 34.1 % of the variance is due to populations and accessions, respectively. As available genomic data for each population comes from pooled DNA of contained accessions (Frachon et al., 2018), this morphological intra population variability cannot be exploited for GWAS. However, as the population effect is highly significant (p-value at 3.2E-14 and 1.1E-05 for adherent and non-adherent mucilage, respectively), GWAS is still pertinent (Table 1).

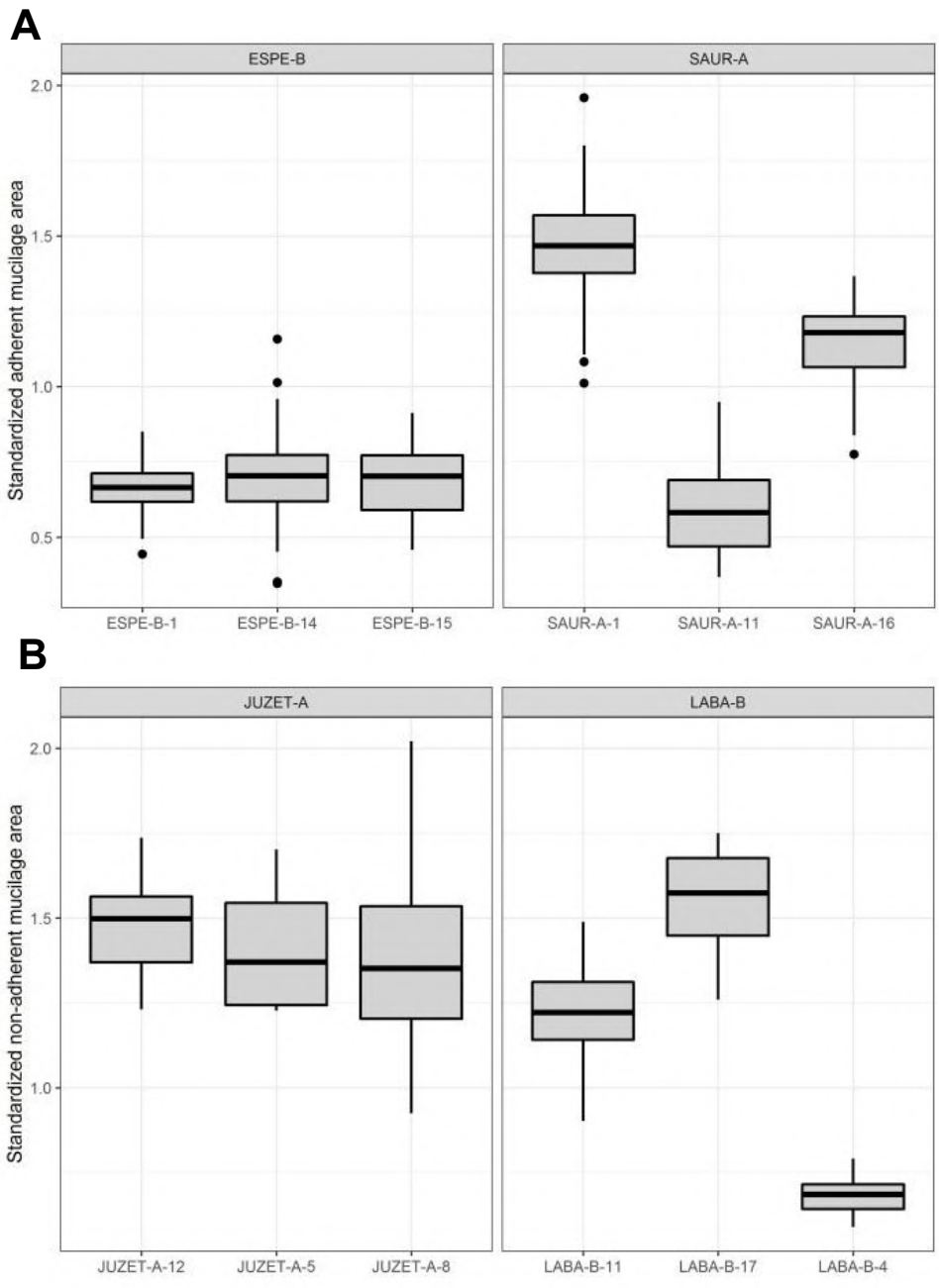


Figure 2: The populations can contain either homogeneous or contrasted accessions for both mucilage layers. Example of adherent **(A)** and non-adherent **(B)** mucilage standardized area values for four population examples containing three homogeneous (left side) and three contrasted (right side) accessions, respectively.

The variation of the areas of seed mucilage layers is highly polygenic

The GWAS highlighted 26 QTLs including 55 genes for the adherent mucilage and 10 QTLs including 28 genes for the non-adherent mucilage, altogether corresponding to 83 genes (Fig. 3; Fig. 4; Table 2). As inter-population variance accounted only at 13.8 % for non-adherent mucilage compared to the 32.3 % for adherent mucilage, it seems coherent to obtain less QTLs for non-adherent mucilage.

In *A. thaliana* there is already about one hundred genes characterized for their implication in seed mucilage (Viudes et al., 2021). Interestingly, only one of the 83 candidate genes highlighted here belongs to this gene list. It corresponds to CuAOα1 (At1g31670) a copper amine oxidase that was identified with GWAS on non-adherent mucilage (Table 2). Insertion mutant for this gene has less rhamnogalacturonan I (RGI) pectin domain in the non-adherent mucilage (Fabrissin et al., 2019). Altogether, the previous GWAS studies led on mucilage did not reveal any previously characterized myxospermy-related genes but were able to characterize three new genes, including CuAOα1 (Voiniciuc et al., 2016; Fabrissin et al., 2019). Some of the 83 genes identified in our study are members of multigenic families previously shown to contain myxospermy-related genes. It is the case of GA2OX4 (At1g47990) and SK19 (At2g03160) for adherent mucilage (Fig. 4A); HB-2 (At4g16780), SK41 (At1g09840), CuAOα2 (At1g31690), and CuAOα3 (At1g31710) for non-adherent mucilage (Fig. 4B) that belong to the same multigenic families as the previously characterized GA3OX4 (At1g80330; Kim et al., 2005), SK11/SK12 (At4g34210/At4g34470; Li et al., 2018), HB-25 (At5g65410; Bueso et al., 2014), and CuAOα1 (At1g31670, Fabrissin et al., 2019). Otherwise, two multigenic families not previously shown to be involved in myxospermy are more represented than the others: cytochrome P450 (CYP450) with six genes (At2g34500, At4g15380, At3g25180, At4g15393, At2g34490, At1g13110) and serine carboxypeptidase like (SCPL) with 4 genes (At2g23000, At2g22980, At2g22970, At2g22990) (Fig. 4A, B). Interestingly, among the 83 candidate genes, 40 are expressed in the seed coat along the seed development kinetics (Belmonte et al., 2013, Fig. 4). Previously characterized functions in other contexts, associated with seed development expression data provide interesting elements for some candidates. SFP2 (At5g27360) is a candidate identified in the most significant adherent mucilage QTL 24 (Fig. 3) with 42 significant SNP (p-value <0.05) into the gene region and is expressed during early seed development (Fig. 4A). It is the tandem paralog of SFP1 (At5g27350) also present in adherent mucilage QTL 24, but not expressed during seed development (Fig. 4A). SFP1 has been characterized as a monosaccharide transporter during leaves senescence (Quirino et al., 2001). As the gene duplication seems to be recent, the authors were surprised not to find similar function for SFP2. Therefore, SFP2 could be rather involved in sugar transport related to seed adherent mucilage polysaccharides. For non-adherent mucilage, an interesting polysaccharide-related candidate is At1g13130, an uncharacterized seed coat specific cellulase found upstream of 22 significant SNPs within the QTL 2 (Table 2). As cellulose structured the adherent mucilage and thus reduced the non-adherent part, this cellulase might be implicated in mucilage cellulose modification. At1g13130 is expressed sharply at heart and linear cotyledon stages (Fig. 4A), co-occurring partially with cellulose fibrils formation since CESA5 (At5g09870) (main cellulose synthase involved in seed mucilage) is highly expressed at the linear cotyledon stage (Sullivan et al., 2011).

Intriguingly, several genes present within the highlighted QTLs are related to the secondary metabolism. IMS2/MAM3 (At5g23020) is located upstream of 10 significant SNPs within the adherent mucilage QTL 21 (Table 2) and expressed only at the mature green stage just before the drying of the seed (Fig. 3; Fig. 4A). It was shown to have a role in glucosinolate biosynthesis (Petersen et al., 2019). CYP71A2 (At2g34490) is the gene containing all the 17 significant SNPs of the non-adherent mucilage QTL 8 and shows a seed coat-specific expression increasing along seed development (Fig. 4B) and extending to endosperm at mature green stage (Table S3). CYP450 genes can be implicated in seed suberin and cutin biosynthesis (Renard et al., 2020) and are often implicated in secondary metabolite decoration (Schuller and Werck-Reichhart, 2003). SCPL11 (At2g22970), SNG1 (At2g22990), and SCPL10 (At2g23000), all localized within non-adherent QTL 6 and 7 (Fig. 3), have seed coat specific expression. These three genes are serine carboxypeptidase-like, a large multigenic family implicated in secondary metabolism. SCPL19/SNG2 (At5g09640) has a role in sinapoylcholine formation in *A. thaliana* seeds (Shirley et al., 2001). Although a direct link between *A. thaliana* seed mucilage and secondary metabolism remains elusive, there is a regulatory network that linked seed mucilage, seed pigments, and seed secondary metabolites (Lepiniec et al., 2006; Salem et al., 2017; Viudes et al., 2020).

While mining the expression data of candidate genes, we noticed that, additionally to its expression in seed restricted to the seed coat, CuAOα1 (At1g31670) is expressed in root specifically at root tip (Klepikova et al., 2016), especially in the lateral root cap that its expression is enhanced by high nitrate concentration (Gifford et al., 2008). During the lateral root cap cell formation, a pectinaceous mucilage accumulates inside the sixth cell layer (c6) below the quiescent center to, *in fine* extrude root mucilage around the c7 (Maeda et al., 2019). As CuAOα1 seems to be involved in free RGI production in seed mucilage (Fabrissin et al., 2019) it would be interesting to investigate as well its putative role in root mucilage. Although a putative link between root mucilage and non-adherent mucilage remains elusive, there is an interesting parallel between their developmental and ecological roles. Both root cap cell and mucilage secretory cells accumulate pectinaceous mucilage after starch granule accumulation and undergo cell wall degradations (Francoz et al., 2019; Maeda et al., 2019). Additionally, in other species than *A. thaliana*, seed and root mucilages were shown to change soil physical properties, influencing the water retention and availability (Deng et al., 2014; Ahmed et al., 2016), and also induce microbial behavior modification (Sasse et al., 2018). If the seed mucilage is also implicated in biotic interaction that could be an explanation for the association of both seed mucilage layers with several candidates genes implicated in secondary metabolism highlighted above, knowing that plant secondary metabolites play crucial roles during plant-microorganism interactions (Kessler and Kalske, 2018).

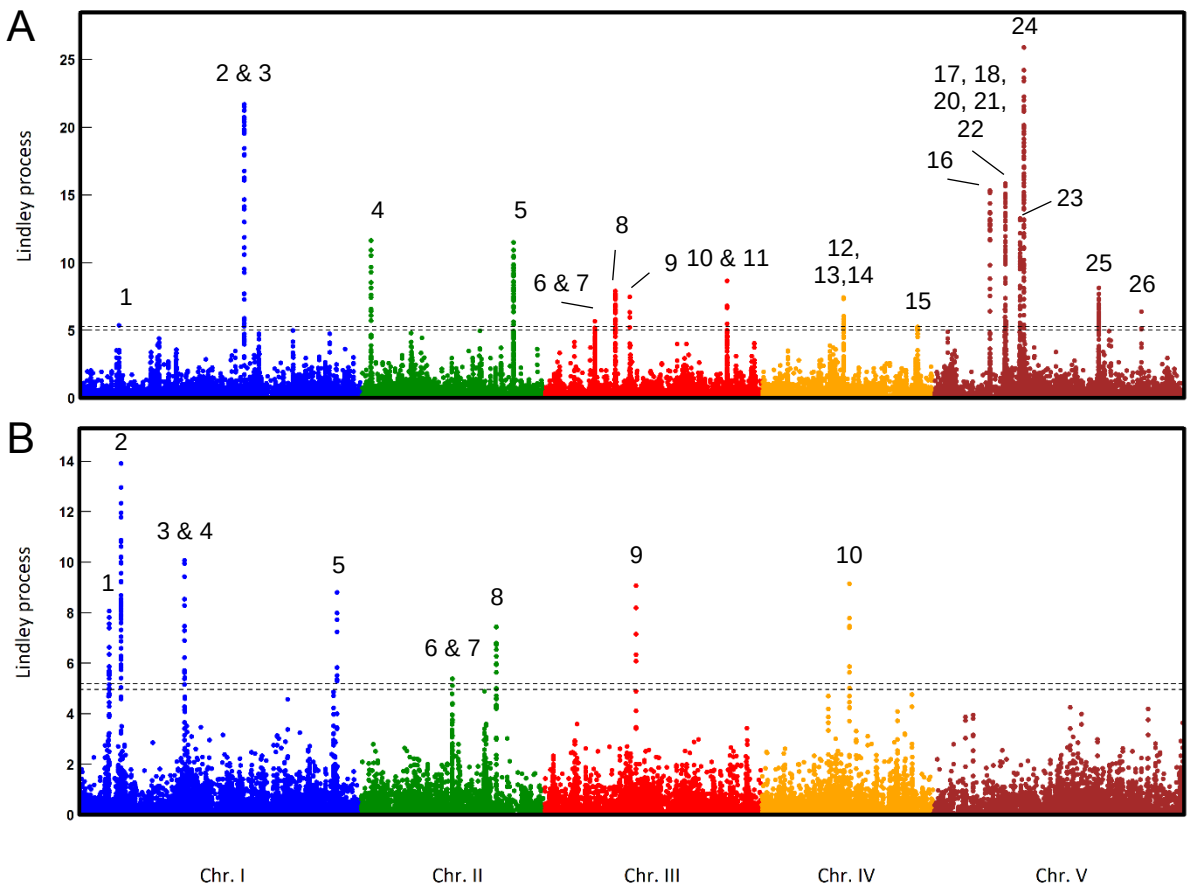


Figure 3: Highlight of genome QTLs explaining the morphological variability of adherent and non-adherent seed mucilage layers. Manhattan plots representing the Lindey process value for each SNPs among the genome of *A. thaliana* highlights 26 QTLs for adherent **(A)** and 10 QTLs for non-adherent **(B)** seed mucilage. The threshold used for QTLs selection is represented by the dotted line at Lindey value of 5. Everything below is considered as background noise.

Table 2: Distribution and number of significant SNPs in the QTLs. The table shows **the number** of SNPs identified with the GWAS performed on adherent mucilage (left) and non-adherent mucilage (right) layers and the gene ID of the gene overlapping with the SNPs, or found within a 2 kb upstream and downstream region.

Adherent mucilage

QTLs	SNPs	Overlap ID	Upstream ID	Downstream ID
QTL1	2	AT1G12280	AT1G12270	AT1G12290
	5	NA	AT1G12280	AT1G12290
QTL2	4	AT1G47990	AT1G47980	AT1G48000
QTL3	23	NA	AT1G47990	AT1G48000
QTL4	14	AT2G03150	AT2G03140	AT2G03160
QTL5	9	AT2G39110	AT2G39100	AT2G39120
	1	NA	AT2G39110	AT2G39120
	22	AT2G39120	AT2G39110	AT2G39130
	17	AT2G39130	AT2G39120	AT2G39140
QTL6	8	AT3G15940	AT3G15930	AT3G15950
QTL7	5	AT3G15950	AT3G15940	AT3G15960
QTL8	26	AT3G21620	AT3G21610	AT3G21630
QTL9	5	AT3G25170	AT3G25165	AT3G25180
	12	AT3G53090	AT3G53080	AT3G53100
QTL11	5	AT3G53090	AT3G53080	AT3G53110
	3	AT3G53100	AT3G53090	AT3G53110
QTL12	3	AT4G15380	AT4G15370	AT4G15390
QTL13	4	NA	AT4G15380	AT4G15390
QTL14	13	AT4G15390	AT4G15380	AT4G15393
	1	NA	AT4G15390	AT4G15393
QTL15	7	AT4G35250	AT4G35240	AT4G35260
QTL16	20	NA	AT5G18200	AT5G18210
QTL17	6	NA	AT5G22860	AT5G22870
QTL18	4	NA	AT5G22860	AT5G22870
QTL19	5	AT5G22880	AT5G22875	AT5G22890
QTL20	3	AT5G22890	AT5G22880	AT5G22900
	16	NA	AT5G22890	AT5G22900
QTL21	10	NA	AT5G23020	AT5G23027
QTL22	14	NA	AT5G26340	AT5G26360
QTL23	3	NA	AT5G26340	AT5G26360
QTL24	42	AT5G27360	AT5G27350	AT5G27370
QTL25	14	AT5G44070	AT5G44065	AT5G44080
	5	NA	AT5G44070	AT5G44080
QTL26	4	NA	AT5G55056	AT5G55060

Non-adherent mucilage

QTLs	SNPs	Overlap ID	Upstream ID	Downstream ID
QTL1	21	AT1G09840	AT1G09830	AT1G09850
QTL2	22	AT1G13120	AT1G13110	AT1G13130
QTL3	15	AT1G31690	AT1G31670	AT1G31710
QTL4	5	AT1G31770	AT1G31760	AT1G31772
QTL5	8	AT1G73880	AT1G73875	AT1G73885
QTL6	2	AT2G22980	AT2G22970	AT2G22990
	4	NA	AT2G22980	AT2G22990
QTL7	3	NA	AT2G22990	AT2G23000
QTL8	17	AT2G34490	AT2G34480	AT2G34500
QTL9	10	NA	AT3G26870	AT3G26880
QTL10	6	AT4G16790	AT4G16780	AT4G16800

A

AGI code	SNPs inside	Gene name	Preglobular	Globular	Heart	Linear cotyledon	Bending cotyledon	Mature green
At5g44080			1194	1158	898	1019	965	617
At5g22880	5	*HTB2*	736	807	217	184	94	112
At4g15390	13		11	6	5	473	727	109
At3g21610			295	487	276	340	423	370
At5g26360		*CCT3*	374	349	244	257	282	136
At1g12270		*HOP1*	344	205	228	234	107	162
At4g35260		*IDH1*	326	193	319	216	234	196
At5g27360	42	*SFP2*	213	301	276	94	42	27
At3g53110		*LOS4*	237	273	157	96	103	173
At3g21630		*CERK1*	195	87	241	83	116	107
At2g03150	14	*RSA1*	238	186	69	96	128	120
At5g44070	14	*PCS1*	124	229	121	121	193	192
At5g18200			142	124	223	93	75	83
At5g22875			149	204	184	157	148	181
At3g15940	8	*UGT*	67	75	97	158	196	143
At5g55060			64	85	105	80	89	171
At2g39130	17		65	88	110	142	84	65
At5g23020		*IMS2*	4	9	9	9	89	126
At4g35240			45	119	37	28	51	92
At3g53090	12	*UPL7*	26	48	29	32	31	96
At2g39110	9	*PBL38*	24	25	31	41	23	85
At1g47980			2	13	1	31	25	79
At2g03140			60	63	36	17	67	73
At3g25170	5	*RALFL26*	15	69	4	61	11	18
At2g39140		*SVR1*	17	48	16	22	26	15
At1g12290		*CC-NBS-LRR*	16	45	22	22	32	48
At2g39120	22	*WTF9*	45	9	35	27	20	25
At3g15950	5	*NAI2*	16	11	7	7	5	40
At1g47990	4	*GA2OX4*	2	2	1	2	15	38
At5g18210			6	18	19	14	34	12
At3g15930			15	6	2	7	15	32
At4g35250	7	*HCF244*	13	16	18	17	30	10
At1g12280	2	*SUMM2*	5	23	14	7	10	29
At2g39100			17	21	13	26	10	21
At5g27350		*SFP1*	10	8	10	10	15	25
At3g53100	3		22	25	19	10	9	11
At3g25180		*CYP82G1*	4	3	7	6	15	23
At4g15380	3	*CYP75A4*	0	6	7	23	15	17
At3g21620	26		5	11	2	19	14	21
At5g22870			0	3	8	4	8	19
At5g22900		*CHX3*	10	1	5	1	1	17
At5g22890	3	*STOP2*	2	11	2	3	4	15
At5g22860			9	4	2	4	5	14
At3g53080			5	1	2	10	9	12
At5g26340		*STP13*	4	11	5	1	1	9
At3g25165		*RALFL25*	3	11	2	1	6	10
At4g15370		*BARS1*	2	2	6	3	10	3
At1g48000		*MYB112*	1	1	5	3	3	9
At2g03160		*ASK19*	6	6	1	5	3	7
At5g44065								
At5g55056								
At3g15960								
At4g15393		*CYP72A5*						
At5g23027								
At5g27370		*DMP10*						

B

AGI code	SNPs inside	Gene name	Preglobular	Globular	Heart	Linear cotyledon	Bending cotyledon	Mature green
At2g34480		*RPL18AB*	947	761	654	832	924	532
At2g22970		*SCPL11*	675	361	258	92	56	40
At2g34490	17	*CYP71A2*	17	52	154	101	527	463
At1g09830		*PUR2*	283	362	361	231	238	172
At1g31670		*CuAOα1*	1	17	5	4	18	354
At2g22990		*SNG1*	245	299	81	55	22	43
At1g13130		*GH5*	35	21	172	66	6	18
At1g13120	22	*GLE1*	118	72	105	78	139	102
At1g09840	21	*SK41*	116	70	76	39	47	43
At4g16800			42	59	103	59	46	72
At2g23000		*SCPL10*	78	89	96	24	50	8
At1g31760		*SWIB5*	82	51	48	69	93	53
At1g09850		*XBCP3*	40	51	79	45	37	62
At1g31690	15	*CuAOα2*	18	17	15	14	22	39
At1g73880	8	*UGT89B1*	16	35	14	14	14	26
At1g13110		*CYP71B7*	22	29	18	34	28	32
At1g31710		*CuAOα3*	8	16	6	3	16	27
At4g16790	6	*HRGP*	4	12	18	23	24	21
At1g31770	5		10	11	4	7	6	23
At2g34500		*CYP71A1*	2	22	1	10	9	5
At4g16780		*HB-2*	3	1	8	9	17	6
At3g26880			2	3	0	1	3	15
At3g26870			8	2	2	9	13	10
At2g22980	2	*SCPL13*	7	11	2	1	5	11
At1g73880	8	*UGT89B1*						
At1g73875		*CCR4E*						
At1g31772		*DEFL*						
At1g73885								

C

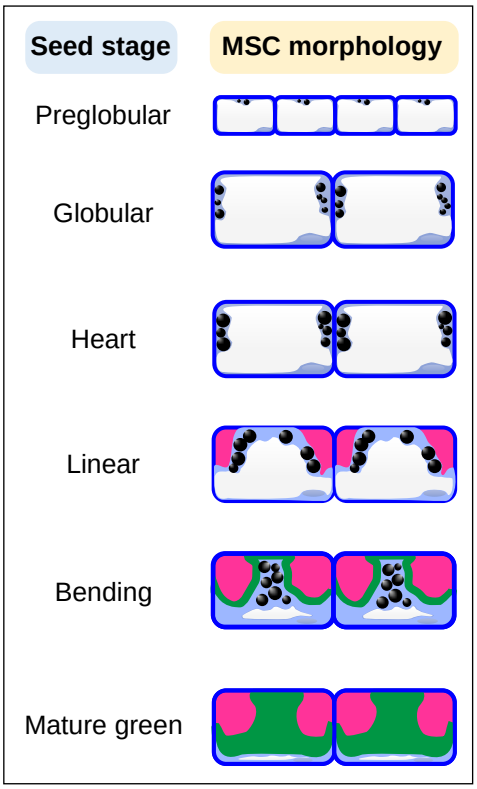


Figure 4: About 50% of the 83 candidate genes from the GWAS are expressed in the seed coat during seed development. Fifty five and 28 selected genes are present within 26 and 10 QTLs

highlighted by the GWAS for adherent **(A)** and non-adherent **(B)** mucilage layers, respectively. These are confronted to previously published transcriptomic data of the isolated seed coat during six seed developmental stages (Belmonte et al., 2013). No significant values below the threshold of "45" are in grey, low expressed genes are in green, highly expressed genes are in red. Note that 10 genes had no spot ID in the tissue arrays. Number of significant SNPs found within the genes region are indicated, genes with no value were located downstream or upstream of significant SNPs. **(C)** Corresponding mucilage secretory cell morphology for every seed developmental stage available in transcriptomic (adapted from Francoz et al., 2019).

Relationship between area of seeds and of seed mucilage layers and between area of seed mucilage layers

Seed area is a relatively stable feature between studied accessions compared to the mucilage (Fig. 1). But interestingly, a slight variation in seed size seems to influence the size of both mucilage layers. Seed area is positively correlated to the adherent and non-adherent mucilage with Pearson coefficient indicating a moderate correlation (respectively 0.365 and 0.494) (Fig. 5). If the cells keep a proportional size in seed, it could be logical to have larger mucilage secretory cells in larger seeds (not evaluated here) and, as a result, more mucilage. Relationships between both mucilage layers are more surprising. Adherent mucilage and non-adherent mucilage are positively correlated with a low but significant Pearson coefficient of 0.19 (Fig. 5). Adherent and non-adherent mucilage are different due to their polysaccharidic composition which provides structure to the former and not the latter (Macquet et al., 2007b). According to the hypothesis that for an equal amount of synthetized mucilage, the adherent and non-adherent area of mucilage are dependent on structural polysaccharide content, we should observe a negative correlation between the two layers of mucilage. The larger the adherent layer is, the lower the non-adherent mucilage should remain and reciprocally. However, the slight positive correlation observed here could rather indicate that the larger is one of the two mucilage layers, the larger is the other one. Consequently, the differences of mucilage between accessions are more likely due to differences in total amount of synthetized mucilage than a change in mucilage composition between all analysed accessions. Note that at 0.191 the correlation is weak which could indicate that both compositional and total amount of mucilage are implicated with a dominance of total amount.

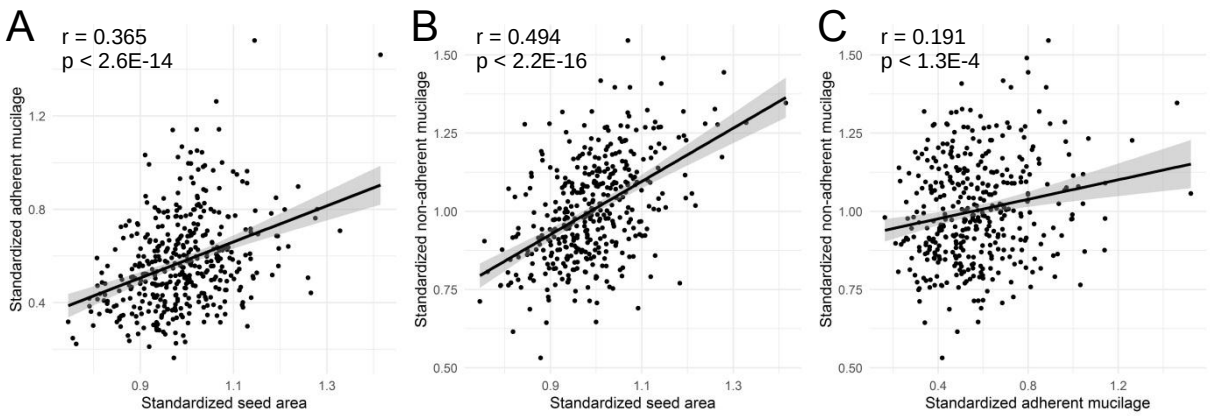


Figure 5: The two mucilage layers are both positively correlated to the area of the seed and weakly to each other. Scatter plots of the mean value of each accession for standardized seed area versus adherent mucilage area **(A)** and non-adherent mucilage area **(B)**, or adherent mucilage versus non-adherent mucilage areas **(C)**. Black line is the linear model with it confidence interval around in grey. r is the Pearson correlation coefficient and p is associated p-value.

Relationship between mucilage and environmental biotic and abiotic parameters

The variation of the adherent mucilage size between the populations shows interesting correlation with several abiotic parameters measured on their harvesting sites (Frachon et al., 2018; Frachon et al., 2019) (Fig. 6). For climatic parameters, the adherent seed mucilage is negatively correlated to the mean annual temperature (-0.28) and to the mean coldest month temperature (-0.35). The adherent mucilage is also highly and positively correlated with the precipitation in winter (0.44) and autumn (0.38), and scarcely in summer (0.16). Therefore, the adherent mucilage seems to be an adaptation to both cold temperature and abundant precipitation. Note that during winter and autumn, the summer annual *A. thaliana* is persisting in the environment as seeds. For these populations, the adherent mucilage could help them to prevent early germination in order to persist in the environment under the seed form, protecting them from the cold temperature of autumn and winter seasons. It was recently demonstrated that the thick adherent seed mucilage of *Lepidium perfoliatum*, another myxospermous Brassicaceae species, prevents the germination under low temperature (10°C and 15°C) and enhances germination under high temperature (25°C and 30°C) especially with a long light time exposure (Zhou et al., 2021). Concerning the correlation with the precipitation it is probably more complex. In *A. thaliana*, seed mucilage was often proposed to help germination by providing more water to the embryo (Penefield, 2001, Arsovsky et al., 2009). However, at 24 h after imbibition, the embryos of *A. thaliana* seeds contain less water for mucilaginous seeds than for non-mucilaginous seeds (Saez-Aguayo et al., 2014). In the previously mentioned study concerning *L. perfoliatum*, seed mucilage was shown to increase drastically germination on abundant or excess water condition while no impact was found under osmotic stress (Zhou et al., 2021). Consequently, the adherent mucilage of *A. thaliana* seed could help winter annual populations to germinate during winter even under water submersion explaining the positive correlation with precipitation in autumn and winter. Correlation between soil composition and adherent mucilage are rare and low but the positive correlations observed with manganese content (0.22) and the nitrogen content (0.16) and the negative correlation with soil PH (-0.15) can be highlighted without current explanation.

The variation of non-adherent mucilage area between populations is not correlated at all with any climatic or soil compositional parameter of their harvesting site. It is an expected result as the non-adherent mucilage will rapidly spread into the ground during the first rain and thus do not influence directly the seed anymore. However, as it spreads into the ground it is surprising that physical parameters such as soil water holding capacity are not correlated.

Populations are not present in their harvesting sites under vegetative form during the same season. Consequently the microbiota communities present on their roots and leaves were sampled in spring and autumn dependent of the population (Bartoli et al., 2018). For spring sampling, the adherent mucilage is negatively correlated with the root pathobiota richness (abundancy) (-0.18) and the Shannon index (diversity) (-0.19) (Fig. 7A). It is the only significant correlation with the microbial communities present on plant observed here. As the radicle passes through the adherent mucilage upon germination, the adherent mucilage might act as hosting a pool of beneficial bacteria that will pre-inseminate the root with a dominance of non-pathogenic bacteria to form the future root microbiota. This hypothesis could explain why the abundance and the diversity of pathogens are globally lower on root when the adherent mucilage is larger. This "selective media" effect can be indirectly triggered by enhancing the growth of a natural competitor within the mucilage as shown with *Streptomyces lividans* (soil bacteria) and *Verticilium dahlia* (plant pathogen fungi) (Meschke and Schrempf, 2010).

A much more unexpected and unexplainable result is the positive correlation between the diversity of pathobiotha on leaves and non-adherent mucilage abundancy (0.21). For autumn sampling, the above reported negative correlation between adherent mucilage and root pathobiotha richness is still occurring (-0.27) but not anymore for the pathobiotha diversity (Fig. 7B). As for spring sampling, it is the only observed significant correlation. Non-adherent mucilage is no more correlated with leaves pathobia diversity but rather negatively correlated with the second axis of the principal component coefficient plotting the pathogen genetic diversity (-0.22) (Fig. 7B). Species corresponding to the low values of axis 2 are mainly *Xanthomonas campestris* and *Pseudomonas viridiflava* (Bartoli et al., 2018). Thus, our data do not show any implication of the non-adherent mucilage in the root microbiota but rather on the leaf microbiota. This could occur by influencing the composition of pathogen species community, maintaining a wide diversity of pathogens (large number and/or equilibrated representation of species through enhancement of the rarest ones). Non-adherent mucilage may have a beneficial role for the plant by avoiding the domination of a single pathogen species to prevent its immune system from being overwhelmed.

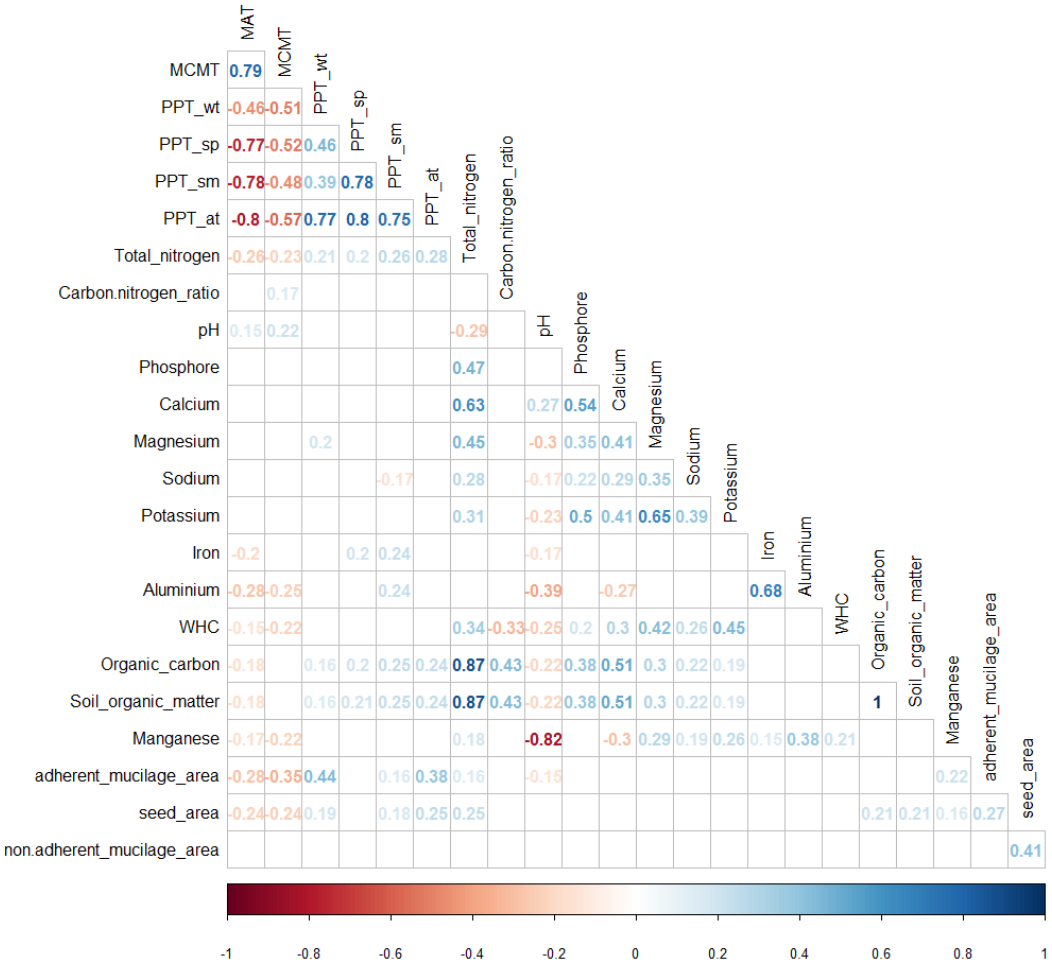
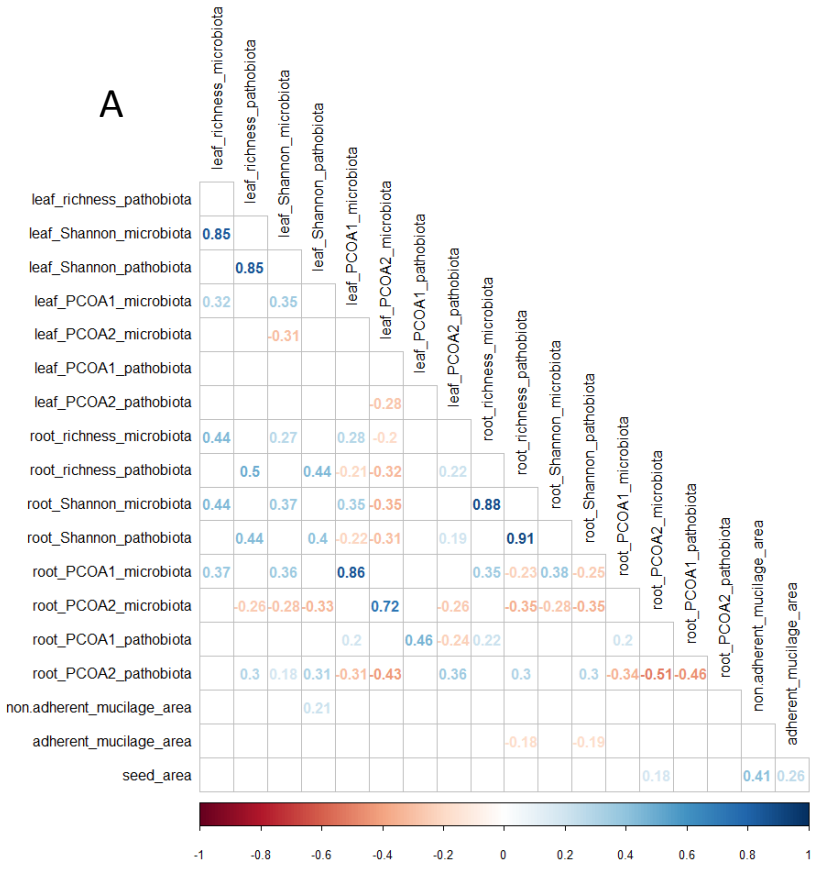


Figure 6: The adherent mucilage is correlated to several abiotic parameters while the non-adherent mucilage does not show such correlation. Heatmap showing statistically significant (p-value < 0.05) Spearman correlation coefficient values. The low and high coefficient values appear in transparency and bold, respectively. Climatic parameters are: mean annual temperature (MAT), mean coldest month temperature (MCMT), precipitation (PPT) in winter (wt), spring (sp), summer (sm), autumn (at). Soil composition parameters have lower coefficient except for water holding capacity (WHC). All details about these abiotic data are available in Frachon et al., 2018, Frachon et al., 2019.


A
leaf_richness_microbiota
leaf_richness_pathobiota
leaf_Shannon_microbiota
leaf_Shannon_pathobiota
leaf_PCOA1_microbiota
leaf_PCOA2_microbiota
leaf_PCOA1_pathobiota
leaf_PCOA2_pathobiota
root_richness_microbiota
root_richness_pathobiota
root_Shannon_microbiota
root_Shannon_pathobiota
root_PCOA1_microbiota
root_PCOA2_microbiota
root_PCOA1_pathobiota
root_PCOA2_pathobiota
non.adherent_mucilage_area
adherent_mucilage_area
seed_area
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

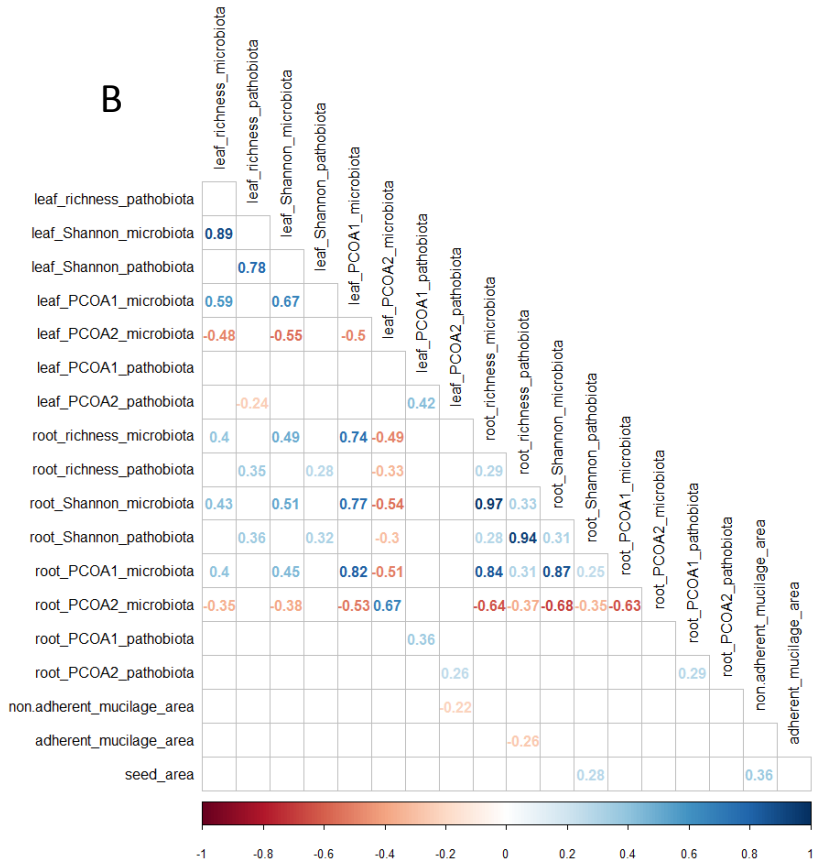

B
leaf_richness_microbiota
leaf_richness_pathobiota
leaf_Shannon_microbiota
leaf_Shannon_pathobiota
leaf_PCOA1_microbiota
leaf_PCOA2_microbiota
leaf_PCOA1_pathobiota
leaf_PCOA2_pathobiota
root_richness_microbiota
root_richness_pathobiota
root_Shannon_microbiota
root_Shannon_pathobiota
root_PCOA1_microbiota
root_PCOA2_microbiota
root_PCOA1_pathobiota
root_PCOA2_pathobiota
non.adherent_mucilage_area
adherent_mucilage_area
seed_area
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1

Figure 7: The correlation between both mucilage layers and biotic parameter measured on developed plant is sparse. Heatmap showing statistically significant (p-value < 0.05) Spearman correlation coefficient values. The low and high coefficient values appear in transparency and bold, respectively. The sampling was realized in two rounds, one during spring **(A)**, and the other one during autumn **(B)**. Both mucilage layers were correlated with the bacterial richness (abundancy), Shannon index (diversity), and PCo (principal component coordinate reflecting the composition) for the root and the leaf. Putative pathobiota are distinguished from the rest of the microbiota thanks to confrontation with a pathogen species list. All details about these abiotic data are available in Bartoli et al., 2018.

Conclusion

The seed mucilage is highly variable between the studied populations for both layers. Several QTLs are underlying this morphological variability containing one previously characterized gene and 82 putative new seed mucilage actors. Considering the one hundred genes already characterized in previous studies, this highlights the extreme polygenic character of this complex trait in *A. thaliana*. CuAOα1 (At1g31670) was previously shown to be implicated in non-adherent mucilage and also associated here with non-adherent mucilage, validating the significance of the approach. New candidate genes coming from our study await functional validation. The correlation of adherent seed mucilage with temperature and water amount seems to be important and will be interesting to validate through germination assays using seed with or without mucilage. The following results might be expected: a role of mucilage in germination inhibition under cold temperature until warmer condition and possibly an enhancement of germination or a longer viability under water excess. Interestingly, both GWAS and correlation with environmental data pointed toward a potential regulation of microbial behavior by seed mucilage. Both layer of seed mucilage were associated with genes implicated in secondary metabolites and correlated with microbial changes in the developed plants. The putative relationship between seed mucilage, secondary metabolism and seed-associated microorganisms is largely understudied so far and might constitute a promising field of research. Altogether this study shed light on putative new genetic actors and ecological role of seed mucilage.

Bibliography

Ahmed MA, Kroener E, Benard P, Zarebanadkouki M, Kaestner A, Carminati A (2016) Drying of mucilage causes water repellency in the rhizosphere of maize: measurements and modelling. Plant Soil **407**: 161–171

Alonso-Blanco C, Andrade J, Becker C, Bemm F, Bergelson J, Borgwardt KMM, Cao J, Chae E, Dezwaan TMM, Ding W, et al (2016) 1,135 Genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana*. Cell **166**: 481–491

Aoun N, Desaint H, Boyrie L, Bonhomme M, Deslandes L, Berthomé R, Roux F (2020) A complex network of additive and epistatic quantitative trait loci underlies natural variation of *Arabidopsis thaliana* quantitative disease resistance to *Ralstonia solanacearum* under heat stress. Mol Plant Pathol **21**: 1405–1420

Arsovski AA, Popma TM, Haughn GW, Carpita NC, McCann MC, Western TL (2009) AtBXL1

encodes a bifunctional -D-Xylosidase/ -L-Arabinofuranosidase required for pectic arabinan modification in Arabidopsis mucilage secretory cells. Plant Physiol **150**: 1219–1234

Bartoli C, Frachon L, Barret M, Rigal M, Huard-Chauveau C, Mayjonade B, Zanchetta C, Bouchez O, Roby D, Carrère S, et al (2018) In situ relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. ISME J **12**: 2024–2038

Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, Hashimoto M, Fei J, Harada CM, Munoz MD, et al (2013) Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. Proc Natl Acad Sci **110**: E435–E444

Bonhomme M, Fariello MI, Navier H, Hajri A, Badis Y, Miteul H, Samac DA, Dumas B, Baranger A, Jacquet C, et al (2019) A local score approach improves GWAS resolution and detects minor QTL: application to *Medicago truncatula* quantitative disease resistance to multiple *Aphanomyces euteiches* isolates. Heredity (Edinb) **123**: 517–531

Brennan AC, Méndez-Vigo B, Haddioui A, Martínez-Zapater JM, Picó FX, Alonso-Blanco C (2014) The genetic structure of *Arabidopsis thaliana* in the south-western Mediterranean range reveals a shared history between North Africa and southern Europe. BMC Plant Biol. **14**: 17

Bueso E, Munoz-Bertomeu J, Campos F, Brunaud V, Martinez L, Sayas E, Ballester P, Yenush L, Serrano R (2014) *Arabidopsis thaliana* HOMEOBOX25 uncovers a role for gibberellins in seed longevity. Plant Physiol **164**: 999–1010

Deng W, Hallett PD, Jeng DS, Squire GR, Toorop PE, Iannetta PPM (2014) The effect of natural seed coatings of *Capsella bursa-pastoris* L. Medik. (shepherd's purse) on soil-water retention, stability and hydraulic conductivity. Plant Soil **387**: 167–176

Fabrissin I, Cueff G, Berger A, Granier F, Sallé C, Poulain D, Ralet M-C, North HM (2019) Natural variation reveals a key role for rhamnogalacturonan I in seed outer mucilage and underlying genes. Plant Physiol **181**: 1498–1518

Frachon L, Bartoli C, Carrère S, Bouchez O, Chaubet A, Gautier M, Roby D, Roux F (2018) A genomic map of climate adaptation in *Arabidopsis thaliana* at a micro-geographic scale. Front Plant Sci **9**: 1–15

Frachon L, Mayjonade B, Bartoli C, Hautekèete NC, Roux F (2019) Adaptation to plant communities across the genome of *Arabidopsis thaliana*. Mol Biol Evol **36**: 1442–1456

François O, Blum MGB, Jakobsson M, Rosenberg NA (2008) Demographic history of European populations of *Arabidopsis thaliana*. PLoS Genet. **5**: e1000075

Francoz E, Ranocha P, Le Ru A, Martinez Y, Fourquaux I, Jauneau A, Dunand C, Burlat V (2019) Pectin demethylesterification generates platforms that anchor peroxidases to remodel plant cell wall domains. Dev Cell **48**: 261-276

Gifford ML, Dean A, Gutierrez RA, Coruzzi GM, Birnbaum KD (2008) Cell-specific nitrogen responses mediate developmental plasticity. Proc Natl Acad Sci U S A **105**: 803–808

Kessler A, Kalske A (2018) Plant secondary metabolite diversity and species interactions. Annu Rev Ecol Evol Syst **49**: 115–138

Kim YC, Nakajima M, Nakayama A, Yamaguchi I (2005) Contribution of gibberellins to the formation of Arabidopsis seed coat through starch degradation. Plant Cell Physiol **46**: 1317–1325

Klepikova A V., Kasianov AS, Gerasimov ES, Logacheva MD, Penin AA (2016) A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. Plant J **88**: 1058–1070

Korte A, Farlow A (2013) The advantages and limitations of trait analysis with GWAS: A review. Plant Methods **9**: 1–9

Krämer U (2015) Planting molecular functions in an ecological context with *Arabidopsis thaliana*. Elife **4**: 1–13

Lepiniec L, Debeaujon I, Routaboul J-M, Baudry A, Pourcel L, Nesi N, Caboche M (2006) Genetics and biochemistry of seed flavonoids. Annu Rev Plant Biol **57**: 405–430

Li C, Zhang B, Chen B, Ji L, Yu H (2018) Site-specific phosphorylation of TRANSPARENT TESTA GLABRA1 mediates carbon partitioning in Arabidopsis seeds. Nat Commun **9**: 571

Macquet A, Ralet M-C, Loudet O, Kronenberger J, Mouille G, Marion-Poll A, North HM (2007a) A naturally occurring mutation in an Arabidopsis accession affects a β-d-Galactosidase that increases the hydrophilic potential of rhamnogalacturonan I in seed mucilage. Plant Cell **19**: 3990–4006

Macquet A, Ralet MC, Kronenberger J, Marion-Poll A, North HM (2007b) In situ, chemical and macromolecular study of the composition of *Arabidopsis thaliana* seed coat mucilage. Plant Cell Physiol **48**: 984–999

Maeda K, Kunieda T, Tamura K, Hatano K, Hara-Nishimura I, Shimada T (2019) Identification of periplasmic root-cap mucilage in developing columella cells of *Arabidopsis thaliana*. Plant Cell Physiol **60**: 1296–1303

Meschke H, Schrempf H (2010) *Streptomyces lividans* inhibits the proliferation of the fungus *Verticillium dahliae* on seeds and roots of *Arabidopsis thaliana*. Microb Biotechnol **3**: 428–443

Meyerowitz EM (1987) *Arabidopsis thaliana*. Annu Rev Genet **21**: 93–111

Miart F, Fontaine JX, Pineau C, Demailly H, Thomasset B, Wuytswinkel O, Pageau K, Mesnard F (2018) MuSeeQ, a novel supervised image analysis tool for the simultaneous phenotyping of the soluble mucilage and seed morphometric parameters. Plant Methods **14**: 1–17

Penfield S (2001) MYB61 is required for mucilage deposition and extrusion in the Arabidopsis seed coat. Plant Cell Online **13**: 2777–2791

Petersen A, Hansen LG, Mirza N, Crocoll C, Mirza O, Halkier BA (2019) Changing substrate specificity and iteration of amino acid chain elongation in glucosinolate biosynthesis through targeted mutagenesis of Arabidopsis methylthioalkylmalate synthase 1. Biosci Rep **39**: 1–15

Phan JL, Burton RA (2018) New insights into the composition and structure of seed mucilage. Annu. Plant Rev. John Wiley & Sons, Ltd, Chichester, UK, pp 1–41

Poulain D, Botran L, North HM, Ralet M-C (2019) Composition and physicochemical properties of outer mucilage from seeds of Arabidopsis natural accessions. AoB Plants **11**: plz031

Quirino BF, Reiter WD, Amasino RD (2001) One of two tandem Arabidopsis genes homologous to monosaccharide transporters is senescence-associated. Plant Mol Biol **46**: 447–457

Renard J, Niñoles R, Martínez-Almonacid I, Gayubas B, Mateos-Fernández R, Bissoli G, Bueso E, Serrano R, Gadea J (2020) Identification of novel seed longevity genes related to oxidative stress and seed coat by genome-wide association studies and reverse genetics. Plant Cell Environ **43**: 2523–2539

Saez-Aguayo S, Rondeau-Mouro C, Macquet A, Kronholm I, Ralet MC, Berger A, Sallé C, Poulain D, Granier F, Botran L, et al (2014) Local evolution of seed flotation in Arabidopsis. PLoS Genet **10**: e1004221

Salem MA, Li Y, Wiszniewski A, Giavalisco P (2017) Regulatory-associated protein of TOR (RAPTOR) alters the hormonal and metabolic composition of Arabidopsis seeds, controlling seed morphology, viability and germination potential. Plant J **92**: 525–545

Sasse J, Martinoia E, Northen T (2018) Feed your friends: Do plant exudates shape the root

microbiome? Trends Plant Sci **23**: 25–41

Schuler MA, Werck-Reichhart D (2003) Functional genomics of P450s. Annu Rev Plant Biol **54**: 629–667

Shirley AM, McMichael CM, Chapple C (2001) The *sng2* mutant of Arabidopsis is defective in the gene encoding the serine carboxypeptidase-like protein sinapoylglucose:choline sinapoyltransferase. Plant J **28**: 83–94

Sullivan S, Ralet M-C, Berger A, Diatloff E, Bischoff V, Gonneau M, Marion-Poll A, North HM (2011) CESA5 is required for the synthesis of cellulose with a role in structuring the adherent mucilage of Arabidopsis seeds. Plant Physiol **156**: 1725–1739

Tsai AY-L, Kunieda T, Rogalski J, Foster LJ, Ellis BE, Haughn GW (2017) Identification and characterization of Arabidopsis seed coat mucilage proteins. Plant Physiol **173**: 1059–1074

Visscher PM, Brown MA, McCarthy MI, Yang J (2012) Five years of GWAS discovery. Am J Hum Genet **90**: 7–24

Viudes S, Burlat V, Dunand C (2020) Seed mucilage evolution: diverse molecular mechanisms generate versatile ecological functions for particular environments. Plant Cell Environ **43**: 2857–2870

Voiniciuc C, Zimmermann E, Schmidt MH-W, Günl M, Fu L, North HM, Usadel B (2016) Extensive natural variation in Arabidopsis seed mucilage structure. Front Plant Sci **7**: 803

Weigel D, Mott R (2009) The 1001 genomes project for *Arabidopsis thaliana*. Genome Biol **10**: 107

Western TL (2001) Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis. Plant Physiol **127**: 998–1011

Zhou Z, Xing J, Zhao J, Liu L, Gu L, Lan H (2021) The ecological roles of seed mucilage on germination of *Lepidium perfoliatum*, a desert herb with typical myxospermy in Xinjiang. Plant Growth Regul. 1-17

Supplemental Figures

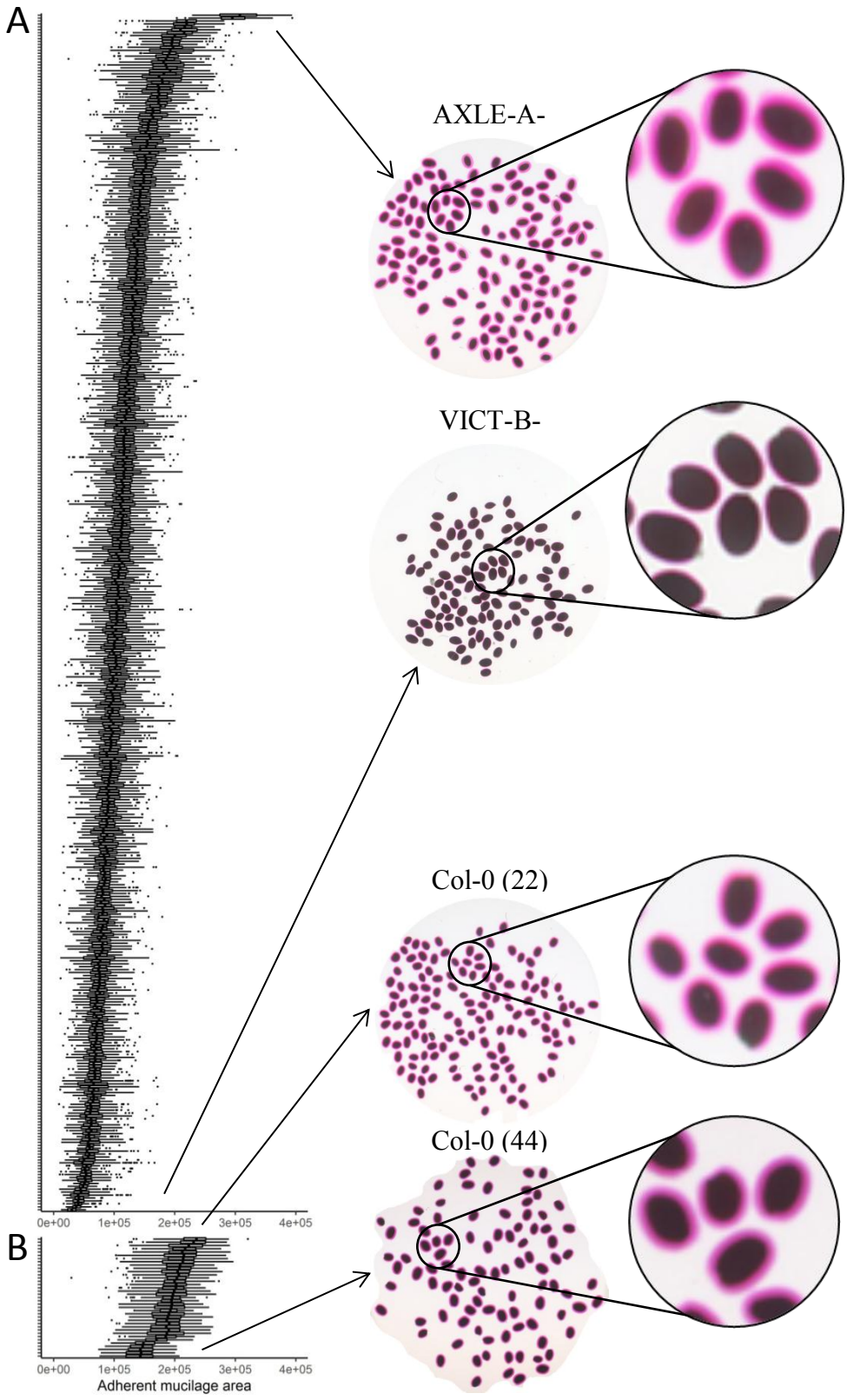


Figure S1: Measurement of adherent mucilage ordered by their mean value before the standardization relatively to Col-0 standard, along with source picture examples illustrating the extreme area values. (A) Plot showing the distribution of the mean area (μm^2) of the adherent mucilage from the 424 accessions. **(B)** Plot showing the distribution of the mean area (μm^2) of the adherent mucilage from the Col-0 measurement across the 37 experimental batches. For two selected accessions and Col-0 references, images corresponding to the extreme values are shown. Each

individual boxplot represents the measured area of the adherent mucilage of approximately 100 seeds. Images come from high resolution scan after vigorous shaking and ruthenium red staining.

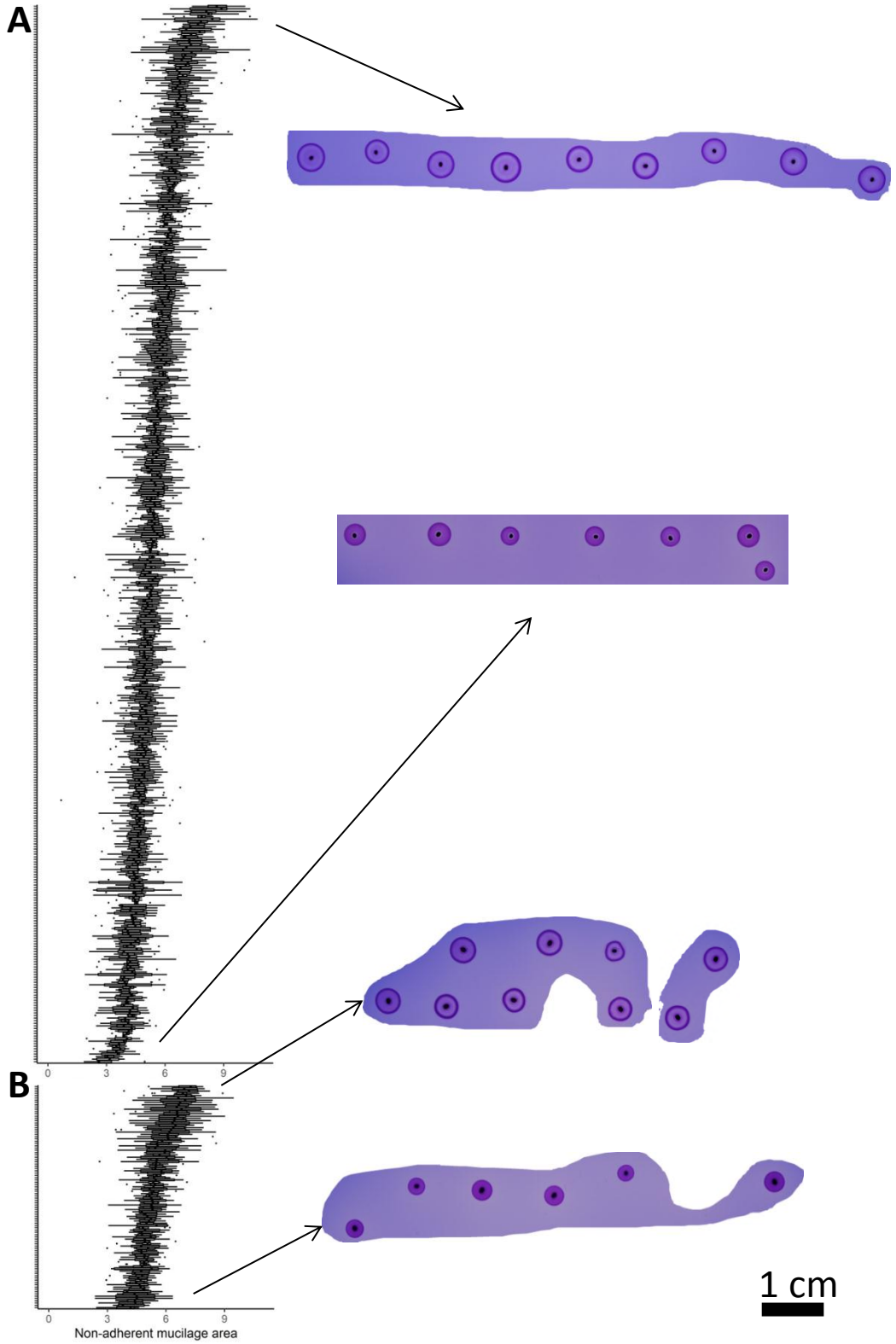


Figure S2: Measurement of the non-adherent mucilage area ordered by their mean value before the standardization relatively to the Col-0 standard, along with source picture examples

illustrating the extreme values. (A) Plot showing the distribution of the mean area (mm^2) of the non-adherent mucilage from the 424 accessions. **(B)** Plot showing the distribution of the mean area (mm^2) of the non-adherent mucilage from the different Col-0 measurement across the 98 experimental batches. For both accessions and Col-0 references, image corresponding to the extreme values are shown. Each boxplot represents the measured area of non-adherent mucilage of approximately 8 seeds. Images come from pictures taken 24 h after dry seed deposition on a Toluidine blue agarose media.

Supplementary Table 1: Detail of every SNPs under the QTLs highlight by GWAS on adherent seed mucilage. The table shows for each QTL every SNP position, statistical significance (pval), Lindley process (plotted in the Manhattan plot Figure 3) and the gene ID of the gene overlapping with the SNPs, or found within a 2 kb upstream and downstream region.

QTL	chrom	pos	pval	lindley	Overlap_ID	Upstream_ID	Downstream_ID
QTL1	1	4177485	0,002389683	0,621659784	AT1G12280	AT1G12270	AT1G12290
QTL1	1	4177523	0,001675666	1,397472329	AT1G12280	AT1G12270	AT1G12290
QTL1	1	4177655	0,002990273	1,921761542	#N/A	AT1G12280	AT1G12290
QTL1	1	4177741	0,004622015	2,256930195	#N/A	AT1G12280	AT1G12290
QTL1	1	4177749	0,000755034	3,378963484	#N/A	AT1G12280	AT1G12290
QTL1	1	4177764	0,006699029	3,552951617	#N/A	AT1G12280	AT1G12290
QTL1	1	4177799	0,000154963	5,362722787	#N/A	AT1G12280	AT1G12290
QTL2	1	17699260	3,8529E-05	2,414212731	AT1G47990	AT1G47980	AT1G48000
QTL2	1	17699272	0,00028192	3,964086736	AT1G47990	AT1G47980	AT1G48000
QTL2	1	17699286	0,000288593	5,503800686	AT1G47990	AT1G47980	AT1G48000
QTL2	1	17699404	6,59869E-05	7,684342746	AT1G47990	AT1G47980	AT1G48000
QTL3	1	17701095	0,000432626	1,363886998	#N/A	AT1G47990	AT1G48000
QTL3	1	17701097	0,000201187	3,060287048	#N/A	AT1G47990	AT1G48000
QTL3	1	17701109	0,00082357	4,144586706	#N/A	AT1G47990	AT1G48000
QTL3	1	17701117	0,000722214	5,285921038	#N/A	AT1G47990	AT1G48000
QTL3	1	17701130	0,002623952	5,866965123	#N/A	AT1G47990	AT1G48000
QTL3	1	17701185	0,504766076	4,163874963	#N/A	AT1G47990	AT1G48000
QTL3	1	17701375	0,00607745	4,380153544	#N/A	AT1G47990	AT1G48000
QTL3	1	17701410	0,012898006	4,269630955	#N/A	AT1G47990	AT1G48000
QTL3	1	17701455	9,65359E-06	7,284942085	#N/A	AT1G47990	AT1G48000
QTL3	1	17701519	0,000106353	9,258190836	#N/A	AT1G47990	AT1G48000
QTL3	1	17701582	0,00047633	10,58028298	#N/A	AT1G47990	AT1G48000
QTL3	1	17701598	0,000508501	11,87399093	#N/A	AT1G47990	AT1G48000
QTL3	1	17701633	8,26883E-05	13,95654688	#N/A	AT1G47990	AT1G48000
QTL3	1	17701634	0,006457262	14,14649849	#N/A	AT1G47990	AT1G48000
QTL3	1	17701669	0,000117969	16,07472878	#N/A	AT1G47990	AT1G48000
QTL3	1	17701670	0,00011685	18,00709988	#N/A	AT1G47990	AT1G48000
QTL3	1	17701734	0,537279907	16,27689928	#N/A	AT1G47990	AT1G48000
QTL3	1	17701821	0,00328603	16,76022776	#N/A	AT1G47990	AT1G48000
QTL3	1	17701841	1,71054E-05	19,52709409	#N/A	AT1G47990	AT1G48000
QTL3	1	17701878	0,006389389	19,72163476	#N/A	AT1G47990	AT1G48000
QTL3	1	17701891	0,1882444	18,44691269	#N/A	AT1G47990	AT1G48000
QTL3	1	17701905	0,000208847	20,12708459	#N/A	AT1G47990	AT1G48000
QTL3	1	17701911	0,005481431	20,38819061	#N/A	AT1G47990	AT1G48000
QTL3	1	17701966	0,001396037	21,24329381	#N/A	AT1G47990	AT1G48000
QTL3	1	17701978	0,047117385	20,57011263	#N/A	AT1G47990	AT1G48000
QTL3	1	17702006	0,000763891	21,68708099	#N/A	AT1G47990	AT1G48000
QTL4	2	958124	0,000736278	1,132958426	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958133	0,003012441	1,654039893	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958136	0,005570886	1,908115614	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958139	0,004700086	2,236009849	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958154	0,000677933	3,40482329	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958163	0,000812678	4,494904588	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958166	0,000116166	6,429825172	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958178	0,001000488	7,429613192	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958216	0,527461282	5,707422607	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958277	0,001391875	6,563822329	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958296	0,0002677	8,136174741	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958314	0,000705209	9,28785708	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958397	0,004125912	9,672337089	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958470	0,000547581	10,93388868	AT2G03150	AT2G03140	AT2G03160
QTL4	2	958473	0,002008672	11,63097969	AT2G03150	AT2G03140	AT2G03160
QTL5	2	16320782	0,005033556	0,298125098	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16320814	0,003121283	0,803791909	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16320823	0,000527933	2,081212731	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16320951	0,001100868	3,039477408	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321502	0,002690233	3,609687467	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321564	0,003446587	4,072298186	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321588	0,069004999	3,233417631	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321608	0,003364972	3,706436151	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321635	0,015095573	3,527586539	AT2G39110	AT2G39100	AT2G39120

QTL5	2	16321636	0,006096762	3,742487271	AT2G39110	AT2G39100	AT2G39120
QTL5	2	16321780	0,010848304	3,707125413	#N/A	AT2G39110	AT2G39120
QTL5	2	16321785	0,96992241	1,72038842	#N/A	AT2G39110	AT2G39120
QTL5	2	16321891	0,021289578	1,392221368	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16321958	0,00157553	2,194794766	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16321980	0,001972711	2,899731242	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322049	0,502428552	1,198656931	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322138	0,006245991	1,403055574	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322144	0,002897115	1,941089783	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322145	0,00321983	2,43325688	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322147	0,006298886	2,633993122	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322163	0,003029594	3,152608642	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322327	0,006738042	3,324074899	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322419	0,002687763	3,894683915	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322504	0,001372577	4,757147328	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322564	0,108169616	3,723042039	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322597	0,006161207	3,933376273	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322607	0,022764566	3,576116892	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322612	0,000575959	4,815725448	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322726	0,099292766	3,818807837	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322732	0,001023995	4,808509926	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322741	0,000121522	6,723856616	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322759	0,000151748	8,542733626	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322805	0,577304114	6,781328973	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322834	0,003904055	7,189813101	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322849	0,006223959	7,395746366	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322876	0,002333278	8,027779932	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322888	0,000266477	9,602119724	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16322919	0,001692114	10,37369003	AT2G39120	AT2G39110	AT2G39130
QTL5	2	16323234	0,600707516	8,59502696	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323302	0,018935774	8,317743904	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323319	0,02753983	7,877782644	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323343	0,016047467	7,672376149	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323376	0,017365804	7,43268126	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323419	0,001839347	8,168017703	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323479	0,001011831	9,162909855	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323556	0,01282998	9,054683859	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323565	0,004487697	9,402660357	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323616	0,015095573	9,223810745	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323649	0,003648615	9,661682762	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323681	0,006648021	9,838990398	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16323974	0,018268682	9,577283184	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16324133	0,027237152	9,142121487	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16324136	0,001959349	9,850009744	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16324154	0,00292815	10,38341643	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16324367	0,002901743	10,92075745	AT2G39130	AT2G39120	AT2G39140
QTL5	2	16324418	0,002691469	11,49076804	AT2G39130	AT2G39120	AT2G39140
QTL6	3	5394792	0,000548842	1,26055232	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5394850	0,000166945	3,037978637	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5394880	0,012384862	2,945087449	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5394976	0,125915121	1,845009563	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5394987	0,002318118	2,479874037	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5394994	0,001511137	3,300570261	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5395020	0,000180094	5,045071665	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5395056	0,145668292	3,881706635	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5395114	0,001212267	4,798108408	AT3G15940	AT3G15930	AT3G15950
QTL6	3	5395122	0,001306421	5,682025238	AT3G15940	AT3G15930	AT3G15950
QTL7	3	5402250	0,000382923	1,416888016	AT3G15950	AT3G15940	AT3G15960
QTL7	3	5402252	0,00014776	3,247330114	AT3G15950	AT3G15940	AT3G15960
QTL7	3	5402302	0,018808513	2,972975663	AT3G15950	AT3G15940	AT3G15960
QTL7	3	5402403	0,000272867	4,537024227	AT3G15950	AT3G15940	AT3G15960
QTL7	3	5402512	0,002339938	5,167819809	AT3G15950	AT3G15940	AT3G15960
QTL8	3	7611283	0,004798295	0,318913028	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611294	0,019484544	0,029222774	AT3G21620	AT3G21610	AT3G21630

QTL8	3	7611315	0,001017801	1,021560011	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611324	0,002523217	1,619605337	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611360	0,003313151	2,099364156	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611531	0,022063773	1,75568437	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611533	0,036972163	1,187809512	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611534	0,000353601	2,639296426	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611605	0,100419447	1,6374786	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611606	0,008744051	1,695765937	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611644	0,014465625	1,535428728	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611651	0,000530165	2,811017311	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611766	0,000187616	4,537746377	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7611817	0,002383447	5,160540856	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612235	0,002785957	5,715566476	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612478	0,624621533	3,919949525	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612488	0,00101424	4,913808689	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612521	0,008861102	4,966320955	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612529	0,015074052	4,788090943	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612530	0,00738884	4,919514657	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612557	0,006041867	5,138343497	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612560	0,013759159	4,999751615	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612570	0,006988039	5,155396275	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612892	0,000449277	6,502882558	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612981	0,001354598	7,371072129	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7612990	0,005445195	7,635058709	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7613043	0,125153601	6,537615359	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7613053	0,000717165	7,681996123	AT3G21620	AT3G21610	AT3G21630
QTL8	3	7613079	0,005970625	7,905976335	AT3G21620	AT3G21610	AT3G21630
QTL9	3	9165812	6,55315E-05	2,183550134	AT3G25170	AT3G25165	AT3G25180
QTL9	3	9165819	0,000269393	3,753163431	AT3G25170	AT3G25165	AT3G25180
QTL9	3	9165826	0,000333353	5,230258806	AT3G25170	AT3G25165	AT3G25180
QTL9	3	9165831	0,000793215	6,330868105	AT3G25170	AT3G25165	AT3G25180
QTL9	3	9165836	0,000712562	7,478045256	AT3G25170	AT3G25165	AT3G25180
QTL10	3	19683271	0,001158812	0,935986864	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683394	0,001462852	1,770786356	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683402	0,004414251	2,125929372	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683407	0,001615299	2,917676499	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683452	0,001629443	3,705637425	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683494	0,017290527	3,467829198	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683497	0,001862997	4,197617079	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683568	0,012648064	4,095593033	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683621	0,004167447	4,475722974	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683744	0,003873038	4,887671218	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683795	0,000123168	6,797172923	AT3G53090	AT3G53080	AT3G53100
QTL10	3	19683908	0,000137879	8,65767408	AT3G53090	AT3G53080	AT3G53100
QTL11	3	19684898	0,004299734	0,366558381	AT3G53090	AT3G53080	AT3G53110
QTL11	3	19684926	0,008150668	0,455365196	AT3G53090	AT3G53080	AT3G53110
QTL11	3	19684976	0,009425558	0,481058122	AT3G53090	AT3G53080	AT3G53110
QTL11	3	19685085	0,000918893	1,517793254	AT3G53090	AT3G53080	AT3G53110
QTL11	3	19685264	0,015464091	1,328468843	AT3G53090	AT3G53080	AT3G53110
QTL11	3	19685402	0,000442547	2,682509827	AT3G53100	AT3G53090	AT3G53110
QTL11	3	19685422	0,00011746	4,612619089	AT3G53100	AT3G53090	AT3G53110
QTL11	3	19685447	0,001371656	5,475373954	AT3G53100	AT3G53090	AT3G53110
QTL12	4	8789963	2,05518E-05	2,687150509	AT4G15380	AT4G15370	AT4G15390
QTL12	4	8790007	4,37559E-05	5,046114151	AT4G15380	AT4G15370	AT4G15390
QTL12	4	8790070	0,00097527	6,056989135	AT4G15380	AT4G15370	AT4G15390
QTL13	4	8792512	0,000518612	1,285157111	#N/A	AT4G15380	AT4G15390
QTL13	4	8792516	0,000308612	2,795744581	#N/A	AT4G15380	AT4G15390
QTL13	4	8792521	1,35618E-05	5,663426958	#N/A	AT4G15380	AT4G15390
QTL13	4	8792527	0,000219987	7,321029386	#N/A	AT4G15380	AT4G15390
QTL14	4	8793546	0,000930615	1,031229942	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793624	0,007029279	1,184319175	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793697	0,006964052	1,341457151	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793724	0,006460995	1,531157764	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793754	0,001107456	2,486831455	AT4G15390	AT4G15380	AT4G15393

QTL14	4	8793839	0,008545765	2,5550805	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793840	0,016985883	2,324992377	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793888	0,010124736	2,31960866	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8793951	0,002616261	2,901927543	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8794014	0,025132208	2,501696893	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8794180	0,000477375	3,822836778	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8794222	0,000246167	5,431606382	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8794284	0,017215534	5,19568589	AT4G15390	AT4G15380	AT4G15393
QTL14	4	8794486	6,10333E-05	7,410118814	#N/A	AT4G15390	AT4G15393
QTL15	4	16772271	0,001683821	0,773703978	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772283	0,000204803	2,462368174	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772350	0,001984047	3,164816298	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772359	0,000154763	4,975150342	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772407	0,017067993	4,742967891	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772559	0,00440364	5,099156044	AT4G35250	AT4G35240	AT4G35260
QTL15	4	16772561	0,006997509	5,254212582	AT4G35250	AT4G35240	AT4G35260
QTL16	5	6016887	0,001927386	0,715031364	#N/A	AT5G18200	AT5G18210
QTL16	5	6016894	0,0036688	1,150507316	#N/A	AT5G18200	AT5G18210
QTL16	5	6016904	0,001334711	2,025120028	#N/A	AT5G18200	AT5G18210
QTL16	5	6016917	0,002023413	2,719035485	#N/A	AT5G18200	AT5G18210
QTL16	5	6016920	0,001186216	3,644871704	#N/A	AT5G18200	AT5G18210
QTL16	5	6016986	0,127009203	2,541036512	#N/A	AT5G18200	AT5G18210
QTL16	5	6017011	0,001924763	3,256659328	#N/A	AT5G18200	AT5G18210
QTL16	5	6017032	0,480361566	1,575091076	#N/A	AT5G18200	AT5G18210
QTL16	5	6017059	0,007880947	1,678512676	#N/A	AT5G18200	AT5G18210
QTL16	5	6017100	0,003299008	2,160129315	#N/A	AT5G18200	AT5G18210
QTL16	5	6017144	0,072030478	1,302613016	#N/A	AT5G18200	AT5G18210
QTL16	5	6017280	0,05434641	0,567442155	#N/A	AT5G18200	AT5G18210
QTL16	5	6017282	0,001661356	1,346979445	#N/A	AT5G18200	AT5G18210
QTL16	5	6017410	0,000580692	2,583033896	#N/A	AT5G18200	AT5G18210
QTL16	5	6017442	0,216234706	1,248108495	#N/A	AT5G18200	AT5G18210
QTL16	5	6017468	0,010415523	1,23042741	#N/A	AT5G18200	AT5G18210
QTL16	5	6017520	0,000160741	3,024302091	#N/A	AT5G18200	AT5G18210
QTL16	5	6017573	0,0004459	4,375064794	#N/A	AT5G18200	AT5G18210
QTL16	5	6017604	0,120120196	3,295448762	#N/A	AT5G18200	AT5G18210
QTL16	5	6017619	0,000296717	4,8231067	#N/A	AT5G18200	AT5G18210
QTL16	5	6017620	1,94859E-05	7,533385652	#N/A	AT5G18200	AT5G18210
QTL16	5	6017665	0,000521358	8,816249428	#N/A	AT5G18200	AT5G18210
QTL16	5	6017713	1,46188E-05	11,65133751	#N/A	AT5G18200	AT5G18210
QTL16	5	6017743	0,000102381	13,64111709	#N/A	AT5G18200	AT5G18210
QTL16	5	6017762	0,001853834	14,37304622	#N/A	AT5G18200	AT5G18210
QTL16	5	6017765	0,001116364	15,32524031	#N/A	AT5G18200	AT5G18210
QTL17	5	7643645	0,001975394	0,704346338	#N/A	AT5G22860	AT5G22870
QTL17	5	7643670	0,000405284	2,096586628	#N/A	AT5G22860	AT5G22870
QTL17	5	7643672	0,000236594	3,722583057	#N/A	AT5G22860	AT5G22870
QTL17	5	7643680	0,001182046	4,649948646	#N/A	AT5G22860	AT5G22870
QTL17	5	7643699	0,000466331	5,981254162	#N/A	AT5G22860	AT5G22870
QTL17	5	7643729	0,002687609	6,551888114	#N/A	AT5G22860	AT5G22870
QTL18	5	7646591	7,08526E-05	2,149644277	#N/A	AT5G22860	AT5G22870
QTL18	5	7646594	0,000520533	3,433195957	#N/A	AT5G22860	AT5G22870
QTL18	5	7646733	0,00028939	4,971712204	#N/A	AT5G22860	AT5G22870
QTL18	5	7646898	0,002317039	5,606778951	#N/A	AT5G22860	AT5G22870
QTL19	5	7652289	0,001758549	0,754845411	AT5G22880	AT5G22875	AT5G22890
QTL19	5	7652304	0,000183878	2,490315325	AT5G22880	AT5G22875	AT5G22890
QTL19	5	7652352	0,00018767	4,216920932	AT5G22880	AT5G22875	AT5G22890
QTL19	5	7652442	0,062822191	3,418807854	AT5G22880	AT5G22875	AT5G22890
QTL19	5	7652443	0,000437312	4,778016631	AT5G22880	AT5G22875	AT5G22890
QTL19	5	7652457	0,000221803	6,432048469	AT5G22880	AT5G22875	AT5G22890
QTL20	5	7653811	8,17197E-05	2,087673306	AT5G22890	AT5G22880	AT5G22900
QTL20	5	7653884	0,000193707	3,800528355	AT5G22890	AT5G22880	AT5G22900
QTL20	5	7654200	0,949192613	1,823174005	AT5G22890	AT5G22880	AT5G22900
QTL20	5	7654621	0,000506253	3,118806226	AT5G22890	AT5G22880	AT5G22900
QTL20	5	7654654	0,063652961	2,314987618	AT5G22890	AT5G22880	AT5G22900
QTL20	5	7654884	8,58901E-05	4,38104428	#N/A	AT5G22890	AT5G22900

QTL20	5	7655055	3,84261E-05	6,796418204	#N/A	AT5G22890	AT5G22900
QTL20	5	7655059	0,118219257	5,723729979	#N/A	AT5G22890	AT5G22900
QTL20	5	7655069	0,000686974	6,886789707	#N/A	AT5G22890	AT5G22900
QTL20	5	7655089	0,323298101	5,377186553	#N/A	AT5G22890	AT5G22900
QTL20	5	7655308	8,23813E-06	8,461357971	#N/A	AT5G22890	AT5G22900
QTL20	5	7655316	0,00016586	10,24161613	#N/A	AT5G22890	AT5G22900
QTL20	5	7656024	0,455602609	8,583029924	#N/A	AT5G22890	AT5G22900
QTL20	5	7656147	0,00045741	9,922724076	#N/A	AT5G22890	AT5G22900
QTL20	5	7656229	0,0108682	9,886566443	#N/A	AT5G22890	AT5G22900
QTL20	5	7656266	0,000480908	11,20450443	#N/A	AT5G22890	AT5G22900
QTL20	5	7656286	0,004966073	11,50849135	#N/A	AT5G22890	AT5G22900
QTL20	5	7656288	0,000663823	12,68643888	#N/A	AT5G22890	AT5G22900
QTL20	5	7656307	0,003698546	13,11840781	#N/A	AT5G22890	AT5G22900
QTL20	5	7656359	0,000530865	14,39342394	#N/A	AT5G22890	AT5G22900
QTL20	5	7656400	0,004265201	14,76348439	#N/A	AT5G22890	AT5G22900
QTL20	5	7656411	0,009284587	14,79572181	#N/A	AT5G22890	AT5G22900
QTL20	5	7656423	0,00505996	15,09157474	#N/A	AT5G22890	AT5G22900
QTL20	5	7656440	0,00178274	15,84048673	#N/A	AT5G22890	AT5G22900
QTL21	5	7723484	0,000119624	1,922183001	#N/A	AT5G23020	AT5G23027
QTL21	5	7723651	0,003609953	2,36468143	#N/A	AT5G23020	AT5G23027
QTL21	5	7723657	0,021835334	2,025521595	#N/A	AT5G23020	AT5G23027
QTL21	5	7723807	0,03499287	1,481542033	#N/A	AT5G23020	AT5G23027
QTL21	5	7723840	0,002513179	2,081318528	#N/A	AT5G23020	AT5G23027
QTL21	5	7723954	0,000289749	3,619295887	#N/A	AT5G23020	AT5G23027
QTL21	5	7724007	0,000780794	4,726759416	#N/A	AT5G23020	AT5G23027
QTL21	5	7724064	0,002864416	5,269723284	#N/A	AT5G23020	AT5G23027
QTL21	5	7724193	0,01727034	5,032422407	#N/A	AT5G23020	AT5G23027
QTL21	5	7724210	0,002478833	5,638175142	#N/A	AT5G23020	AT5G23027
QTL22	5	9249415	0,001974694	0,70450028	#N/A	AT5G26340	AT5G26360
QTL22	5	9249487	0,000181433	2,445783513	#N/A	AT5G26340	AT5G26360
QTL22	5	9249542	0,014724203	2,277751716	#N/A	AT5G26340	AT5G26360
QTL22	5	9249548	0,0007857	3,382494712	#N/A	AT5G26340	AT5G26360
QTL22	5	9249581	0,015801544	3,183795185	#N/A	AT5G26340	AT5G26360
QTL22	5	9249632	0,028701858	2,72588517	#N/A	AT5G26340	AT5G26360
QTL22	5	9249661	0,007797898	2,833907624	#N/A	AT5G26340	AT5G26360
QTL22	5	9249711	0,000111236	4,787660983	#N/A	AT5G26340	AT5G26360
QTL22	5	9249712	0,033306322	4,265134308	#N/A	AT5G26340	AT5G26360
QTL22	5	9249758	0,092969025	3,296796033	#N/A	AT5G26340	AT5G26360
QTL22	5	9249827	0,000407169	4,687021099	#N/A	AT5G26340	AT5G26360
QTL22	5	9250101	3,05641E-05	7,201808928	#N/A	AT5G26340	AT5G26360
QTL22	5	9250130	0,965643362	5,216992168	#N/A	AT5G26340	AT5G26360
QTL22	5	9250281	5,61663E-05	7,467516414	#N/A	AT5G26340	AT5G26360
QTL22	5	9250299	1,05895E-05	10,44264235	#N/A	AT5G26340	AT5G26360
QTL22	5	9250311	1,51981E-05	13,26085332	#N/A	AT5G26340	AT5G26360
QTL23	5	9252433	0,000125521	1,90128304	#N/A	AT5G26340	AT5G26360
QTL23	5	9252474	2,2512E-06	5,548868009	#N/A	AT5G26340	AT5G26360
QTL23	5	9252539	0,002890351	6,087917439	#N/A	AT5G26340	AT5G26360
QTL24	5	9658739	0,002019123	0,694837174	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658783	5,99733E-05	2,916879142	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658784	0,001917481	3,634148119	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658857	0,00261822	4,21614202	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658885	0,000651209	5,402421779	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658895	0,001891967	6,125508266	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658904	0,000666062	7,301993322	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658965	0,002657517	7,877517352	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658979	0,000955253	8,897398944	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9658981	0,010204387	8,888612017	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659069	0,001319446	9,768220245	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659112	0,007107602	9,916497117	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659122	0,633598934	8,114682679	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659141	0,205986084	6,800844798	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659144	0,000975148	7,811774322	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659168	0,000101176	9,806698017	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659171	7,05472E-05	11,95821796	AT5G27360	AT5G27350	AT5G27370

QTL24	5	9659186	1,07419E-05	14,92713692	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659262	0,574450366	13,16788441	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659288	0,001614328	13,95989277	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659527	0,002096756	14,63834481	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659554	0,297794071	13,16442876	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659579	0,011107824	13,11879975	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659592	0,000509511	14,41164629	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659661	0,005374516	14,68130697	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659677	0,004895162	14,99153993	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659681	0,002211631	15,64682727	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659723	0,390143264	14,05560316	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659768	0,001662553	14,83482778	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659797	0,000513432	16,12434478	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659817	3,41495E-05	18,59096019	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659892	0,385408145	17,0050393	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659912	0,008079996	17,09762816	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659921	0,000174915	18,85480148	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659934	0,005282782	19,13193875	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9659949	0,095948876	18,14989886	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660122	0,000110575	20,10624191	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660172	0,000131163	21,98843014	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660194	0,000367842	23,42276875	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660220	0,744017264	21,55118574	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660226	0,430892553	19,91681675	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660464	0,023875671	19,53886117	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660466	0,003601099	19,98242606	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660553	0,630230523	18,18292662	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660554	0,03246839	17,67146587	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660555	0,169745208	16,44166835	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660557	0,007477283	16,5679245	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660759	0,000520464	17,85153351	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660776	0,274569015	16,41288199	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660782	0,000167772	18,18816192	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660783	0,000481164	19,50586921	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660820	0,000214184	21,17508257	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660826	3,30364E-05	23,6560892	AT5G27360	AT5G27350	AT5G27370
QTL24	5	9660902	5,79982E-05	25,89267439	AT5G27360	AT5G27350	AT5G27370
QTL25	5	17736682	0,005258445	0,279142649	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17736700	2,28006E-05	2,92119551	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17736716	0,001740889	3,680424368	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17736720	0,000867775	4,742017465	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17736767	0,006599434	4,922510789	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17736921	0,004012386	5,319108082	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737004	0,096703844	4,333664343	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737043	0,000335818	5,807560023	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737046	0,000711553	6,955352571	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737081	0,004303744	7,321506136	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737114	0,814137079	5,410808601	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737443	0,001706665	6,178660203	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737562	0,006844498	6,343318592	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737580	0,156461077	5,148912276	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737690	0,000975822	6,159541841	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737728	0,014313487	6,003796399	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737751	0,000207027	7,687768482	AT5G44070	AT5G44065	AT5G44080
QTL25	5	17737869	0,067119802	6,860917815	#N/A	AT5G44070	AT5G44080
QTL25	5	17737889	0,042814587	6,229326057	#N/A	AT5G44070	AT5G44080
QTL25	5	17737920	0,003746593	6,655689586	#N/A	AT5G44070	AT5G44080
QTL25	5	17737948	0,822507149	4,740549905	#N/A	AT5G44070	AT5G44080
QTL25	5	17737988	0,001207939	5,658505023	#N/A	AT5G44070	AT5G44080
QTL25	5	17738026	0,000257385	7,24792178	#N/A	AT5G44070	AT5G44080
QTL25	5	17738058	0,001319204	8,127609763	#N/A	AT5G44070	AT5G44080
QTL26	5	22341704	0,000134067	1,872677457	#N/A	AT5G55056	AT5G55060
QTL26	5	22341726	0,000140565	3,72480092	#N/A	AT5G55056	AT5G55060
QTL26	5	22341822	0,000453826	5,067911885	#N/A	AT5G55056	AT5G55060
QTL26	5	22341851	0,000480238	6,386455496	#N/A	AT5G55056	AT5G55060

Supplementary Table 2: Detail of every SNPs under the QTLs highlight by GWAS on non-adherent seed mucilage. The table shows for each QTL every SNP position, statistical significance (pval), Lindley process (plotted in the Manhattan plot Figure 3) and the gene ID of the gene overlapping with the SNPs, or found within a 2 kb upstream and downstream region.

QTL	chrom	pos	pval	lindley	Overlap_ID	Upstream_ID	Downstream_ID
QTL1	1	3197144	0,003016211	0,520538238	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197193	0,016879251	0,293185058	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197218	0,005001381	0,594095134	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197224	0,016597287	0,37405804	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197347	0,003085435	0,884741591	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197375	0,008429983	0,958914869	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197403	0,005249752	1,238776062	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197437	0,001635215	2,025201251	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197449	0,008590414	2,091187156	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197482	0,000727184	3,229543117	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197503	0,000287667	4,770653102	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197676	0,021391054	4,440420918	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197791	0,002652325	5,016794102	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197848	0,00319449	5,512392607	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3197932	0,008719813	5,571885417	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199078	0,001649914	6,354424134	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199215	0,031071219	5,862065842	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199288	0,001688252	6,634628581	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199315	0,001746918	7,392356128	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199348	0,003801336	7,812419885	AT1G09840	AT1G09830	AT1G09850
QTL1	1	3199402	0,00557571	8,066119725	AT1G09840	AT1G09830	AT1G09850
QTL2	1	4470666	0,004805368	0,318273346	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470767	0,002646537	0,8955954	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470772	0,003228935	1,386536163	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470776	0,000115436	3,324193194	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470836	0,000190413	5,044496781	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470855	9,773E-06	8,054469014	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470862	6,93016E-05	10,21372564	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4470978	0,002172371	10,87679164	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471009	0,076254113	9,994528367	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471019	0,010325372	9,980622647	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471029	0,197561726	8,684919835	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471074	0,001333307	9,559989686	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471126	0,106588947	8,532277514	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471156	0,000219869	10,19011438	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471207	0,67597217	8,360185563	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471216	0,030382285	7,87756513	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471223	0,002905718	8,414311738	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471343	0,232907513	7,047128239	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471349	0,002802581	7,599569997	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471524	0,00474733	7,923120596	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471525	0,005835105	8,1570719	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471568	0,000159295	9,954870418	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471616	0,002181038	10,61620711	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471636	0,005664728	10,86302806	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471730	0,001231068	11,77274602	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471736	0,002774818	12,32951157	AT1G13120	AT1G13110	AT1G13130
QTL2	1	4471973	0,000265322	13,90573798	AT1G13120	AT1G13110	AT1G13130
QTL3	1	11344459	0,000489943	1,309854077	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344502	0,002868998	1,852123765	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344604	0,000271894	3,417724761	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344776	0,714265478	1,5638651	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344778	0,000402769	2,958808915	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344782	6,28265E-05	5,160665862	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344798	0,003021526	5,680439591	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344817	0,253518362	4,276430171	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344826	0,000362802	5,716760778	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344842	0,367346573	4,151684785	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11344924	0,000594118	5,377811919	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345370	0,001427551	6,223220137	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345374	0,002107118	6,899531234	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345459	0,004025058	7,29475914	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345495	0,000571657	8,537623516	AT1G31690	AT1G31670	AT1G31710

QTL3	1	11345497	0,001300573	9,42348889	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345561	0,003043903	9,940058026	AT1G31690	AT1G31670	AT1G31710
QTL3	1	11345585	0,007452698	10,06774449	AT1G31690	AT1G31670	AT1G31710
QTL4	1	11377227	0,000319817	1,495098587	AT1G31770	AT1G31760	AT1G31772
QTL4	1	11377318	1,85267E-05	4,227301335	AT1G31770	AT1G31760	AT1G31772
QTL4	1	11377324	0,003568485	4,674817407	AT1G31770	AT1G31760	AT1G31772
QTL4	1	11377340	0,012142177	4,590520831	AT1G31770	AT1G31760	AT1G31772
QTL4	1	11377355	0,001435383	5,433553104	AT1G31770	AT1G31760	AT1G31772
QTL5	1	27785883	0,000121594	1,915086746	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27785896	0,000321881	3,407391857	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27785910	0,002540464	4,002478799	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27785912	0,000435604	5,363387268	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27786015	0,00712695	5,510483543	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27786042	0,004874323	5,822569282	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27786130	0,000126304	7,721151342	AT1G73880	AT1G73875	AT1G73885
QTL5	1	27786168	0,000827809	8,803221388	AT1G73880	AT1G73875	AT1G73885
QTL6	2	9782979	0,00222199	0,653257971	AT2G22980	AT2G22970	AT2G22990
QTL6	2	9782983	0,000262575	2,234005036	AT2G22980	AT2G22970	AT2G22990
QTL6	2	9783384	0,000295767	3,76305602	#N/A	AT2G22980	AT2G22990
QTL6	2	9783406	0,002313257	4,398832161	#N/A	AT2G22980	AT2G22990
QTL6	2	9783443	0,483968953	2,714014659	#N/A	AT2G22980	AT2G22990
QTL6	2	9783530	0,000582349	3,948830958	#N/A	AT2G22980	AT2G22990
QTL6	2	9783532	0,000676689	5,118441845	#N/A	AT2G22980	AT2G22990
QTL7	2	9791037	0,006253653	0,203866205	#N/A	AT2G22990	AT2G23000
QTL7	2	9791075	2,89628E-06	3,742025122	#N/A	AT2G22990	AT2G23000
QTL7	2	9791166	0,000229744	5,380780615	#N/A	AT2G22990	AT2G23000
QTL8	2	14536528	0,000434701	1,361809231	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536540	0,001316041	2,242539875	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536563	0,010638118	2,215675052	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536648	0,001715149	2,981373195	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536653	9,67085E-05	4,995908751	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536673	0,010578062	4,97150263	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536687	0,018123946	4,713249872	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536795	0,013672361	4,577406351	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536837	0,000203768	6,268270549	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14536934	0,97014925	4,281431996	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537050	0,011942515	4,204336184	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537077	0,00407578	4,594125496	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537122	0,017989719	4,339101109	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537131	0,000230276	5,976852343	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537146	0,021586098	5,642678201	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537161	0,005054556	5,938995214	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537170	0,000321595	7,43168658	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537189	0,050349326	6,729692915	AT2G34490	AT2G34480	AT2G34500
QTL8	2	14537347	0,001925776	7,445087046	AT2G34490	AT2G34480	AT2G34500
QTL9	3	9903939	0,000633754	1,1980793	#N/A	AT3G26870	AT3G26880
QTL9	3	9903942	0,002403447	1,81724481	#N/A	AT3G26870	AT3G26880
QTL9	3	9904097	0,183117708	0,554514465	#N/A	AT3G26870	AT3G26880
QTL9	3	9904124	0,027494939	0,115261698	#N/A	AT3G26870	AT3G26880
QTL9	3	9904136	0,002572764	0,704861755	#N/A	AT3G26870	AT3G26880
QTL9	3	9904141	0,000496514	2,008930438	#N/A	AT3G26870	AT3G26880
QTL9	3	9904142	0,000351814	3,462617928	#N/A	AT3G26870	AT3G26880
QTL9	3	9904151	0,000375629	4,887858938	#N/A	AT3G26870	AT3G26880
QTL9	3	9904158	0,000354911	6,337739739	#N/A	AT3G26870	AT3G26880
QTL9	3	9904267	0,000138624	8,195902553	#N/A	AT3G26870	AT3G26880
QTL9	3	9904268	0,001343125	9,067786013	#N/A	AT3G26870	AT3G26880
QTL10	4	9451870	0,000108031	1,96645083	AT4G16790	AT4G16780	AT4G16800
QTL10	4	9451876	5,539E-05	4,223019408	AT4G16790	AT4G16780	AT4G16800
QTL10	4	9451877	0,000388429	5,633707587	AT4G16790	AT4G16780	AT4G16800
QTL10	4	9451884	0,000171404	7,399686769	AT4G16790	AT4G16780	AT4G16800
QTL10	4	9451890	0,008467034	7,471955465	AT4G16790	AT4G16780	AT4G16800
QTL10	4	9452080	0,000215596	9,138315557	AT4G16790	AT4G16780	AT4G16800

Supplementary Table 3: Expression of every candidate genes in the different seed tissues during six seed developmental stages (Belmonte et al., 2013). No significant values below the threshold of “45” are in grey, low expressed genes are in green, highly expressed genes are in red. Note that 10 genes had no spot ID in the tissue arrays.

			SEED COAT						CHALAZAL SEED COAT						WHOLE SEED							EMBRYO						MICROPYLAR ENDOSPERM					PERIPHERAL ENDOSPERM						CHALAZAL ENDOSPERM					
GWAS trait	ORF/AgI	Gene Symbol	pg SC	g SC	h SC	lc SC	bc SC	mg SC	pg CZSC	g CZSC	h CZSC	lc CZSC	bc CZSC	mg CZSC	pg WS	g WS	h WS	lc WS	bc WS	mg WS	g SUS	pg EP	g EP	h EP	lc EP	bc EP	mg EP	pg MCE	g MCEN	h MCEN	lc MCEN	mg MCEN	pg PEN	g PEN	h PEN	lc PEN	bc PEN	mg PEN	pg CZEN	g CZEN	h CZEN	lc CZEN	bc CZEN	mg CZEN
Adherent mucilage	At5g44080		1194	1158	898	1019	965	617	1588	1421	1399	1429	877	941	1121	662	943	334	481	596	700	459	756	393	702	495	544	427	1773	509	351	536	101	425	513	521	331	527	1550	2993	3372	1434	473	694
Non-adherent mucilage	At2g34480		947	761	654	832	924	532	596	730	603	621	662	570	827	1422	1015	1733	1786	1410	1091	3191	2629	1966	2392	1955	1381	1593	1402	1155	1088	1001	1980	1679	1677	901	1627	1241	900	284	297	609	1363	523
Adherent mucilage	At5g22880	HTB2	736	807	217	184	94	112	597	833	262	185	110	110	498	364	309	382	551	230	194	1421	1752	905	814	1108	191	365	219	396	280	151	358	486	342	640	322	177	294	124	71	92	134	243
Adherent mucilage	At1g47980		2	13	1	31	25	79	2	1	3	4	1	6	1	1	2	19	8	62	1	8	2	5	2	1	5	4	6	75	271	1581	3	2	10	3	17	662	3	8	3	1	3	267
Non-adherent mucilage	At4g16790	HRGP	4	12	18	23	24	21	418	1022	919	902	1045	1346	74	46	22	55	19	10	11	9	6	0	10	8	7	3	1	10	41	58	6	17	8	13	29	29	22	7	20	52	13	57
Adherent mucilage	At1g12270	TPR	344	205	228	234	107	162	380	305	214	248	158	196	391	81	267	158	134	108	395	1230	782	1166	393	426	90	486	360	1151	85	47	582	485	747	198	169	54	539	305	583	376	413	36
Non-adherent mucilage	At1g09830		283	362	361	231	238	172	225	277	295	320	257	229	258	374	293	291	312	190	429	977	564	737	503	286	87	391	447	974	133	141	775	1168	763	399	164	139	392	262	479	753	561	173
Adherent mucilage	At5g26360		374	349	244	257	282	136	281	294	226	200	313	168	284	424	265	433	438	399	445	959	865	772	563	686	510	614	480	544	416	399	718	841	743	530	560	561	234	254	295	380	577	509
Adherent mucilage	At4g15390		11	6	5	473	727	109	1	7	3	31	23	30	6	10	18	253	345	305	5	7	17	27	109	243	275	12	3	12	100	16	4	2	9	21	314	108	4	21	6	13	11	24
Non-adherent mucilage	At2g22970	SCPL11	675	361	258	92	56	40	62	24	30	30	9	11	334	230	108	17	22	39	9	2	8	6	4	13	14	13	7	1	11	71	4	11	4	2	5	39	3	12	2	8	1	10
Adherent mucilage	At2g03150	RSA1	238	186	69	96	128	120	381	446	295	263	127	92	249	59	126	83	89	66	289	450	657	397	312	260	186	247	152	359	115	208	245	293	268	128	104	131	249	447	439	151	113	188
Adherent mucilage	At3g25170	RALFL26	15	69	4	61	11	18	14	11	43	14	93	11	14	8	13	2	2	3	12	32	9	1	189	7	11	29	9	13	12	14	2	9	10	181	8	13	14	11	25	593	11	8
Adherent mucilage	At3g21610		295	487	276	340	423	370	212	362	248	572	243	426	225	388	229	198	113	99	142	85	91	63	63	40	63	558	380	105	146	160	130	100	69	110	95	85	183	394	190	303	184	253
Non-adherent mucilage	At2g34490	CYP71A2	17	52	154	101	527	463	60	136	130	154	145	194	21	72	60	68	109	148	1	7	4	2	2	3	28	3	7	10	2	185	10	6	4	6	8	268	18	14	19	43	5	103
Adherent mucilage	At5g44070	CAD1	124	229	121	121	193	192	275	191	238	124	132	522	139	76	144	157	51	43	34	34	34	37	72	38	71	23	22	26	52	18	19	45	34	48	54	21	19	18	44	27	21	58
Adherent mucilage	At4g35260	IDH1	326	193	319	216	234	196	496	208	336	173	328	309	410	432	266	312	220	160	365	319	206	288	149	199	48	288	415	191	268	104	344	209	320	444	291	118	165	162	312	178	375	137
Non-adherent mucilage	At4g16780	ATHB-2	3	1	8	9	17	6	23	3	10	13	64	3	11	6	22	25	43	25	12	43	12	9	36	37	21	31	41	63	29	48	2	10	133	52	92	90	5	3	4	146	487	43
Non-adherent mucilage	At1g31760		82	51	48	69	93	53	152	133	62	99	73	72	84	133	115	173	280	301	300	338	446	467	290	324	310	102	107	204	155	283	215	204	303	89	174	179	93	33	28	56	184	200
Adherent mucilage	At5g22875		149	204	184	157	148	181	248	290	291	232	186	173	150	128	172	99	144	81	121	169	144	109	127	65	173	236	147	133	121	198	177	112	127	190	215	152	179	372	445	302	142	159
Non-adherent mucilage	At1g13120	emb1745	118	72	105	78	139	102	113	71	129	66	182	112	53	57	93	88	103	48	142	247	188	260	154	126	125	77	66	120	79	99	81	85	130	109	106	113	111	24	82	387	308	136
Adherent mucilage	At3g53110	LOS4	237	273	157	96	103	173	227	196	283	118	105	123	150	183	142	114	150	148	128	220	230	202	124	191	77	192	231	243	104	138	187	256	324	173	69	138	368	214	242	275	296	136
Non-adherent mucilage	At1g31670		1	17	5	4	18	354	4	7	5	13	6	16	7	16	13	8	17	32	3	9	12	3	12	1	7	5	1	5	12	10	18	4	3	6	11	2	14	2	29	2	2	2
Adherent mucilage	At5g55060		64	85	105	80	89	171	45	55	77	75	51	96	63	71	59	62	74	46	142	50	50	56	66	58	88	108	318	36	58	84	65	64	79	58	80	85	69	258	140	55	51	67
Adherent mucilage	At5g27360	SFP2	213	301	276	94	42	27	49	45	51	127	60	41	134	197	106	63	92	223	8	3	6	7	56	79	227	36	1	6	70	39	25	13	11	52	38	79	13	13	10	8	3	15
Non-adherent mucilage	At2g22990	SNG1	245	299	81	55	22	43	19	7	21	16	15	21	182	43	43	7	14	12	5	3	1	1	7	1	115	3	28	0	8	10	17	6	8	9	14	10	41	17	9	11	4	7
Adherent mucilage	At5g23020	IMS2	4	9	9	9	89	126	1	11	16	12	98	290	8	6	31	29	5	33	7	6	3	1	3	5	18	8	10	3	7	6	13	13	1	9	9	10	19	11	13	4	3	5
Non-adherent mucilage	At1g09840	ATSK41	116	70	76	39	47	43	276	150	182	93	143	86	126	84	62	67	58	68	83	96	195	131	90	113	66	65	59	54	156	138	58	51	16	68	118	101	64	75	46	41	41	74
Adherent mucilage	At3g21630	CERK1	195	87	241	83	116	107	158	99	135	77	139	139	159	83	128	72	63	83	48	47	43	75	63	71	93	91	95	63	38	82	24	2	32	22	58	100	40	16	74	3	28	71
Adherent mucilage	At2g39130		65	88	110	142	84	65	16	12	10	19	28	11	49	92	67	53	46	60	72	39	43	66	45	20	46	149	126	31	31	48	34	38	34	35	48	31	233	191	178	23	18	18
Adherent mucilage	At5g18200		142	124	223	93	75	83	159	88	93	76	101	134	96	107	106	125	93	107	112	51	33	69	68	56	82	120	50	43	81	54	41	33	67	84	76	35	76	58	66	59	104	53
Adherent mucilage	At4g35240		45	119	37	28	51	92	197	167	196	90	93	71	36	58	53	47	29	46	29	47	54	41	25	18	60	5	7	10	19	67	29	44	13	31	25	65	20	13	4	10	24	118
Adherent mucilage	At3g15940	UGT	67	75	97	158	196	143	97	89	52	48	57	37	59	106	53	79	104	118	52	53	40	45	67	73	79	36	23	18	74	80	32	8	44	82	32	108	23	8	22	24	27	84
Non-adherent mucilage	At4g16800		42	59	103	59	46	72	58	100	112	83	40	67	51	73	79	69	96	112	184	45	74	133	45	99	91	67	89	58	77	84	77	110	101	97	143	118	43	25	24	74	68	110
Non-adherent mucilage	At1g13130	GH5	35	21	172	66	6	18	11	13	12	4	1	15	68	56	117	20	11	20	3	9	4	6	6	7	5	21	7	3	9	12	8	5	15	9	2	31	51	11	9	15	10	10
Adherent mucilage	At3g21620		5	11	2	19	14	21	5	6	5	7	19	22	4	4	37	26	9	6	27	13	13	8	9	1	2	19	94	105	63	21	4	13	87	17	37	26	10	9	5	28	139	26
Adherent mucilage	At2g39120		45	9	35	27	20	25	30	25	50	25	49	32	39	23	42	34	34	25	52	96	56	96	78	39	31	83	72	74	29	28	59	24	73	26	34	24	33	23	78	114	66	8
Adherent mucilage	At2g03140		60	63	36	17	67	73	35	59	66	46	27	28	48	53	53	27	19	30	23	24	29	28	17	26	23	6	20	5	39	29	20	20	34	21	30	16	40	29	112	36	9	9
Non-adherent mucilage	At1g31710		8	16	6	3	16	27	6	7	10	4	16	2	8	17	23	10	7	5	7	7	2	1	7	1	8	14	17	1	1	7	2	1	8	4	8	1	24	112	34	21	8	5
Non-adherent mucilage	At1g73880	UGT89B1	16	35	14	14	14	26	18	11	24	37	26	40	10	20	50	57	69	82	47	67	101	50	89	57	26	52	29	28	45	4	52	99	75	48	62	24	20	30	23	6	32	33
Adherent mucilage	At3g53090	UPL7	26	48	29	32	31	96	62	51	29	50	68	59	31	55	39	35	27	49	29	21	29	22	32	28	48	29	44	5	32	45	10	3	13	16	34	44	37	68	55	68	42	45
Non-adherent mucilage	At2g23000	SCPL1	78	89	96	24	50	8	12	13	9	1	7	7	53	46	36	6	23	9	3	2	10	0	2	12	4	11	4	8	21	5	10	3	12	59	16	6	11	4	7	2	5	4
Adherent mucilage	At2g39110		24	25	31	41	23	85	62	28	35	30	37	47	25	22	23	34	23	32	51	28	24	23	23	22	32	3	13	22	23	34	22	5	12	17	20	48	65	10	19	9	27	29
Adherent mucilage	At4g35250	HCF244	13	16	18	17	30	10	7	11	5	17	4	14	36	28	24	34	33	20	38	4	15	22	17	35	16	23	27	34	81	37	53	73	55	16	40	38	10	6	32	7	6	29
Non-adherent mucilage	At1g09850	XBCP3	40	51	79	45	37	62	38	41	65	54	39	16	63	72	45	28	27	16	9	30	25	9	13	10	17	27	30	11	22	17	59	38	20	13	8	20	33	35	65	33	20	14
Adherent mucilage	At3g53100	GDSL	22	25	19	10	9	11	65	61	18	42	26	13	25	25	20	23	20	24	12	15	4	15	14	11	30	13	24	10	17	18	8	8	7	14	18	11	25	6	28	19	18	22
Non-adherent mucilage	At1g31770		10	11	4	7	6	23	44	33	34	52	39	28	21	7	8	7	4	12	4	5	6	4	17	9	11	3	10	10	2	7	7	2	4	8	9	13	9	3	4	8	5	13
Adherent mucilage	At5g22890		2	11	2	3	4	15	1	4	8	4	12	22	5	11	3	7	2	1	1	3	2	1	1	1	9	2	1	3	12	1	3	2	8	2	8	14	19	2	6	2	9	51
Adherent mucilage	At2g39140	SVR1	17	48	16	22	26	15	17	27	24	13	11	18	45	25	12	32	25	34	10	26	32	44	42	28	25	48	19	30	34	23	43	49	35	22	7	25	18	11	11	9	32	18
Adherent mucilage	At1g12290	CC-NBS-LRR	16	45	22	22	32	48	18	29	30	48	33	26	6	23	15	25	22	31	24	20	16	19	14	15	27	30	15	6	20	43	18	18	6	6	15	25	3	20	21	13	25	9
Adherent mucilage	At5g27350	SFP1	10	8	10	10	15	25	10	10	7	7	8	13	14	11	7	7	16	26	4	3	7	4	2	1	9	21	15	1	45	19	7	6	4	13	12	41	3	11	2	6	11	36
Non-adherent mucilage	At1g31690		18	17	15	14	22	39	22	9	16	20	16	16	26	16	17	19	20	2	11	17	8	4	13	10	11	32	24	13	18	10	12	19	11	17	18	14	29	21	25	20	42	9
Adherent mucilage	At3g15950	NAI2	16	11	7	7	5	40	3	4	16	7	5	12	16	11	8	5	21	11	2	2	3	6	5	1	16	18	23	1	7	7	8	11	7	10	19	2	24	13	20	9	26	3
Adherent mucilage	At5g18210		6	18	19	14	34	12	17	20	39	17	20	30	6	16	22	33	10	25	17	1	7	7	3	10	31	17	12	4	5	2	6	5	14	6	18	5	7	6	5	12	13	16
Adherent mucilage	At1g47990	ATGA2OX4	2	2	1	2	15	38	2	7	7	3	7	17	2	10	10	21	10	13	4	10	4	17	29	12	11	15	39	8	22	14	2	5	6	7	5	21	1	5	22	9	9	11
Non-adherent mucilage	At1g13110	CYP71B7	22	29	18	34	28	32	21	24	31	18	25	35	31	5	15	32	13	18	11	22	16	2	14	15	14	38	15	13	15	15	29	7	7	14	10	10	29	30	22	11	2	29
Adherent mucilage	At3g15930		15	6	2	7	15	32	2	7	5	15	5	1	11	2	3	13	12	9	5	7	18	19	5	5	5	3	9	10	4	2	8	6	2	2	7	1	11	13	1	11	6	8
Adherent mucilage	At2g39100		17	21	13	26	10	21	16	10	19	22	7	20	29	16	24	15	16	22	18	22	17	8	13	8	32	15	26	8	9	20	19	14	12	10	4	20	18	9	23	18	23	26
Adherent mucilage	At4g15380	CYP75A4	0	6	7	23	15	17	3	8	6	7	21	6	10	1	1	3	14	8	3	11	5	6	5	5	17	32	2	8	2	8	1	1	11	13	11	11	1	3	11	13	8	2
Adherent mucilage	At3g25165	RALFL25	3	11	2	1	6	10	1	1	14	7	13	1	2	6	1	12	6	2	2	8	14	1	6	1	6	10	12	1	1	2	13	4	4	4	1	5	1	5	4	31	8	12
Adherent mucilage	At1g48000	MYB112	1	1	5	3	3	9	4	2	4	3	27	19	4	2	4	14	4	5	29	11	8	2	10	4	13	11	8	4	20	17	15	5	4	3	8	9	15	10	9	12	8	23
Adherent mucilage	At1g12280	SUMM2	5	23	14	7	10	29	26	27	22	15	21	5	8	4	12	8	14	11	6	14	7	5	5	8	13	6	14	12	9	15	2	10	6	3	12	15	5	9	9	15	17	7
Non-adherent mucilage	At2g22980	SCPL13	7	11	2	1	5	11	1	2	7	11	3	5	2	3	2	4	1	8	6	11	12	2	5	6	9	29	8	4	1	1	2	2	5	2	0	10	19	11	6	18	6	1
Adherent mucilage	At3g25180	CYP82G1	4	3	7	6	15	23	1	10	13	8	20	3	8	10	10	6	14	21	8	1	1	1	2	5	22	4	1	10	9	10	7	7	10	12	10	10	3	15	18	15	14	2
Non-adherent mucilage	At2g34500	CYP71A1	2	22	1	10	9	5	4	3	18	7	1	23	1	4	2	9	13	2	5	15	1	4	1	1	1	18	1	8	2	1	3	3	8	7	13	1	3	10	12	5	2	15
Adherent mucilage	At5g22870		0	3	8	4	8	19	3	9	6	1	3	15	10	14	6	5	17	2	6	7	5	0	9	7	5	14	16	5	7	12	6	3	7	8	6	14	5	10	20	11	3	11
Non-adherent mucilage	At3g26880		2	3	0	1	3	15	6	0	3	1	1	0	9	5	1	3	1	8	1	4	7	0	1	0	11	19	7	1	0	1	3	1	1	6	0	2	4	2	2	9	1	5
Adherent mucilage	At4g15370	BARS1	2	2	6	3	10	3	2	7	2	7	2	1	3	4	2	3	5	13	1	3	1	3	9	2	13	17	2	9	3	5	1	2	1	2	1	5	3	11	7	7	2	2
Adherent mucilage	At5g26340	MSS1	4	11	5	1	1	9	1	3	8	2	17	6	4	4	10	7	6	1	0	1	3	4	1	1	3	16	2	0	11	7	1	6	2	1	1	1	1	5	5	1	6	1
Adherent mucilage	At5g22900	CHX3	10	1	5	1	1	17	3	1	9	12	3	1	1	1	3	2	8	1	5	7	1	1	3	1	1	3	1	3	1	1	3	1	1	4	5	1	5	7	1	2	3	7
Adherent mucilage	At2g03160	ASK19	6	6	1	5	3	7	2	8	7	2	1	2	3	9	2	2	1	8	1	3	4	0	6	4	4	5	15	2	2	1	2	6	1	8	6	1	12	5	16	2	9	4
Adherent mucilage	At3g53080	LECTIN	5	1	2	10	9	12	11	8	4	1	0	1	5	1	5	4	0	1	1	5	1	1	1	3	2	15	1	5	3	9	1	3	1	3	3	3	9	8	4	1	2	6
Adherent mucilage	At5g22860		9	4	2	4	5	14	7	9	10	4	6	10	6	1	1	11	12	9	1	8	9	3	3	2	1	3	11	1	2	4	10	1	5	9	6	2	8	10	11	2	11	1
Non-adherent mucilage	At3g26870		8	2	2	9	13	10	1	5	6	6	4	5	6	14	9	1	6	4	3	5	2	0	1	5	3	7	10	1	7	7	10	4	1	1	3	1	1	3	12	7	4	5
Adherent mucilage	AT5G44065																																											
Adherent mucilage	AT5G55056																																											
Adherent mucilage	AT3G15960																																											
Adherent mucilage	AT4G15393	CYP72A5																																										
Adherent mucilage	AT5G23027																																											
Adherent mucilage	AT5G27370	DMP1																																										
Non-adherent mucilage	AT1G73880	UGT89B1																																										
Non-adherent mucilage	AT1G73875	CCR4E																																										
Non-adherent mucilage	AT1G31772	DEFL																																										
Non-adherent mucilage	AT1G73885																																											

Conclusion générale et perspectives

L'introduction et notamment l'article de revue ont bien posé les bases de la complexité des traits liés à la myxospermie. L'approche de comparaison inter-espèces décrite dans le chapitre 1 avait pour objectif de mieux comprendre l'évolution de la myxospermie grâce au cadre phylogénétique relativement restreint que représente la famille des Brassicaceae. L'approche de comparaison intra espèces du chapitre 2 visait à élargir les connaissances disponibles chez *Arabidopsis thaliana* au niveau du son contrôle génétique et de ses rôles écologiques.

Diversité inter-espèce

La myxospermie est incroyablement variable entre les espèces de Brassicaceae qui possèdent ce trait complexe. La morphologie cellulaire de la couche épidermique des graines qui est responsable de l'extrusion du mucilage est relativement changeante en taille et en forme mais ce n'est rien comparé à la diversité de structure mucilagineuse que ces cellules libèrent lors de l'imbibition de la graine. La nature moléculaire de ce mucilage a partiellement été abordée par immunofluorescence avec une récurrence de la présence d'épitopes de RGI. La disponibilité réduite des graines ne nous a malheureusement pas permis d'établir un profil biochimique des polysaccharides du mucilage chez les espèces myxospermiques étudiées. Etonnamment, les orthologues des gènes caractérisés pour leur rôle dans la myxospermie chez *A. thaliana* sont globalement tous présents dans la sélection d'espèces de Brassicaceae myxospermiques mais également chez les espèces non-myxospermiques. Chez des espèces ne développant pas une fonction particulière impliquée dans la myxospermie, il n'y a donc pas l'absence espérée des gènes impliqués dans cette fonction. Les niveaux d'expression de ces gènes dans des graines non-myxospermiques ne montrent pas non plus de changement significatif comparé aux graines non myxospermiques. Il est important de noter que la sélection des espèces a été guidé par la disponibilité des données génomique chez les espèces de la famille des Brassicaceae ce qui représente un biais d'échantillonnage. L'enrichissement de deux sous familles (Camelineae et Brassiceae) a nui à la représentativité d'autres sous-groupes et les espèces non-myxospermiques analysées étaient globalement moins nombreuses avec des données génomiques de moins bonne qualité.

Le manque de corrélation entre la génomique et la morphologie peut provenir du manque d'adéquation exclusive entre les gènes impliqués dans la myxospermie d'*A. thaliana* et ceux impliqués dans la myxospermie d'autres espèces de la famille. Peut-être ont-ils été majoritairement recrutés tardivement pour la myxospermie chez *A. thaliana* ne rendant pas pertinente l'hypothèse de l'implication dans la myxospermie des gènes orthologues chez d'autres espèces parfois relativement éloignées phylogénétiquement. Considérant la particularité morphologique des trois espèces d'Arabidopsis étudiées comparée au reste de la famille, il est bien possible qu'une partie des gènes que l'on connait actuellement leur soient spécifiques aussi. Une manière de tester cela serait le recoupement chez les gènes orthologues obtenus dans le chapitre 1, de la présence de motifs régulateurs présents dans les

régions promotrices qui provoquent l'expression dans le testa. Par cette méthode indirecte les gènes orthologues qui ne sont potentiellement pas exprimés dans le testa aurait pu être éliminés pour réduire le nombre de faux positifs. Les motifs utilisés auraient pu être celui déjà caractérisé dans la littérature (Dean et al., 2017) ou identifiés *de novo* depuis les promoteur de gènes spécifiques du seed coat (Esfandiari et al., 2013, Jeong et al., 2014).

Pour contourner le problème de la spécificité de la myxospermie chez *A. thaliana* il faudrait développer de nouveaux modèles pour étudier l'évolution du mucilage des graines au niveau moléculaire. La cameline (*Camelina sativa*) est d'ores et déjà en bonne voie pour le devenir (espèce transformable (Liu et al., 2012), génome de bonne qualité (Kagale et al., 2014), transcriptomes de la graine en développement (Abdullah et al., 2016), grosse cellules sécrétrice de mucilage), mais il faudrait aussi une ou plusieurs espèces qui ont divergé plus tôt. *Aethionema arabicum* est intéressante car elle fait partie du groupe qui a divergé le plus tôt de l'ancêtre commun des Brassicaceae. cette espèce est séquencée (Nguyen et al., 2019) et potentiellement transformable (Schlager, 2016). De plus, cette espèce présente un dimorphisme des diaspores qui lui permet d'avoir à la fois des graines myxospermiques et non-myxospermiques, une particularité très pratique qui commence à être exploitée pour comprendre la régulation de l'établissement du mucilage (Arshad et al., 2021). *Arabis alpina* serait un bon deuxième choix d'espèce qui à divergé tôt car elle est séquencée (Willing et al., 2015) et possède de grandes cellules sécrétrices de mucilage. La sous famille des Brassicaea serait aussi un bon cadre d'étude pour comprendre comment les cellules épidermiques se différentient ou non en cellules sécrétrices de mucilage car elle contient des espèces phylogenetiquement proches, qui peuvent être myxospermiques (comme la moutarde blanche ou noire) ou non-myxospermiques (comme *Brassica rapa* ou *Brassica oleracea*). C'est de plus dans cette famille que la plupart des espèces de Brassicaceae d'intérêt agronomique sont regroupées renforçant l'intérêt d'une telle étude par anticipation du jour ou le rôle écologique du mucilage des moutardes sera connu.

L'ouverture des cellules sécrétrices de mucilages au sommet des parois radiales connue chez *A. thaliana* s'est révélé être spécifique au genre Arabidopsis. Grace aux outils et données disponibles chez la Cameline les acteurs moléculaires impliqués dans le processus d'ouverture chez *A. thaliana* pourraient être localisés par anticorps ou gène rapporteur. Si leur localisation coïncide avec l'ouverture centrale des cellules sécrétrices du mucilage de la Cameline ces gènes pourraient être transcomplementé chez le mutant correspondant d'*A. thaliana* pour tenter de délocaliser sa zone d'ouverture. Le mutant ainsi créé pourrait faire l'objet d'analyse physiologique (gestion des flux d'eau) ou développementale (germination) pour comprendre pourquoi la zone d'ouverture a changé au cours de l'évolution.

Un deuxième argument pour expliquer ce manque de corrélation entre la génomique et la morphologie est la forte polygénicité de la myxospermie. En effet, si la centaine de gènes actuellement caractérisés

chez *A. thaliana* ne représentent qu'une faible proportion de l'ensemble des gènes qui participent à l'établissement de ce trait, ceux qui sont pseudogenisés ou non-exprimés chez les espèces non-myxospermiques ont pu passer à travers le crible car ils ne sont peut-être pas encore connus. Une autre approche utilisant l'orthologie des gènes mais cette fois-ci sans a priori a été envisagée au cours de la thèse. Elle aurait permis l'identification de nouveaux gènes impliqués dans la myxospermie par la construction automatique d'orthogroupes puis la mise en avant des orthogroupes qui ne contiennent que des espèces myxospermiques, mais la faible qualité de nombreux génome et la complexité du trait étudié rendait cette approche peu propice. L'établissement d'une liste de gènes plus universels pourrait aussi être réalisée par le recoupement entre les transcriptomes des cellules sécrétrices de mucilage (ou du testa si la dissection laser ne permet pas une précision suffisante) au cours de leur développement chez plusieurs espèces. Quatre espèces myxospermiques éloignées phylogénétiquement et morphologiquement comme *A. arabicum, A. alpina, C. sativa,* et *L. stativum* pourraient apporter cette information mais aussi mettre en avant les gènes spécifiques du morphotype de mucilage relargué partagé par les deux premières espèces en comparaison de ceux liés au morphotype de mucilage relargué formant une matrice beaucoup plus cohésive partagé par les deux dernières espèces.

Diversité intra-espèce

Le caractère extrêmement polygénique du mucilage des graines est appuyé par les nombreux polymorphismes génétiques associés à la variabilité morphologique du mucilage entre les différentes populations naturelles d'*A. thaliana.* La variation de taille des mucilages adhérent et non-adhérent est en effet associée respectivement à 26 et 10 QTLs, ce qui porte le nombre de gènes candidats présents sur, en aval, ou en amont des SNPs significatifs qu'ils contiennent à 83. Même si l'un d'entre eux a déjà été caractérisé dans une étude précédente (Fabrissin et al., 2019), ces gènes candidats ne seront probablement pas tous impliqués dans la myxospermie, mais cela montre néanmoins le caractère hautement polygénique de la myxospermie. Parmi les gènes candidats le transporteur de sucres SFP2 (At5g27360) et une cellulase (At1g13130) sont à caractériser en priorité par génétique inverse car étant les plus prometteur au vu des données provenant de la GWAS, des données d'expression et de leur annotation fonctionnelle en accord avec un possible rôle dans la myxospermie. Le mutant *sfp2* pourrait avoir une composition en monosaccharide différente de celle du sauvage et le mutant de At1g13130 pourrait posséder potentiellement plus de cellulose cristalline et par conséquent un mucilage non-adhérent moins abondant. Outre ces gènes candidats, la caractérisation morphologique du mucilage de population naturelle a permis de montrer une forte corrélation entre le mucilage adhérent et des paramètres abiotiques tels que la température et les précipitations. Ces corrélations viennent renforcer des hypothèses sur le rôle du mucilage d'*A. thaliana* dans le contrôle de la germination relatif à la température et face à la disponibilité en eau qu'il serait intéressant de tester. Des graines de Col-0 pourraient être démucilaginées par sonication ou par traitements enzymatiques pour ensuite tester leur capacité à germer au froid, au chaud, ou en excès d'eau comparativement aux

graines intactes. Quelques corrélations plus faibles avec des paramètres biotiques indiquent aussi qu'à l'instar du mucilage racinaire, le mucilage des graines pourrait être impliqué dans les étapes précoces de l'établissement du microbiote qui existera sur la plante une fois développée. Ces corrélations sont à mettre en parallèle avec certains gènes impliqués dans le métabolisme secondaire qui ont été mis en avant par la GWAS. Il serait intéressant d'étudier le rôle du mucilage dans le stockage et la diffusion de molécules actives en testant l'impact de la myxospermie sur des communautés microbiennes ou d'autres formes d'organismes du sol. Le choix des espèces pourrait être guidé par les résultats obtenus sur la nature des organismes pouvant être influencés ou interagir avec le mucilage des graines (Meschke et Schrempf, 2010; Hu et al., 2019; Hu et al., 2021; Yi-Lun Tsai et al., 2019). Une fois une communauté ou une espèce microbienne sensible trouvée, les mutants de gènes du métabolisme secondaire précédemment évoqués pourraient être testés pour leur capacité antimicrobienne qui pourrait être réduite.

Bilan

L'étude de la diversité inter-espèces a été la plus longue et difficile car elle a nécessité beaucoup de remises en question méthodologiques face à la complexité de la myxospermie chez les Brassicaceae. Cette famille était un bon cadre phylogénétique mais probablement trop grand pour en tirer des conclusions claires. Ce type d'approche gagnera en pertinence avec l'obtention de davantage de données génomiques, surtout si elles sont de bonne qualité. Néanmoins les difficultés rencontrées ont été stimulantes grâce au sentiment de participer à des recherches pionnières basées sur des méthodologies que je désirais pratiquer. À l'inverse l'étude de la diversité intra espèces a porté plus facilement et rapidement ses fruits grâce à la qualité des données associées aux populations utilisées et à la maitrise de la méthodologie de la GWAS par l'équipe de Fabrice Roux. D'un côté je suis donc très satisfait des résultats obtenus mais de l'autre je regrette de ne pas avoir pu me former à cette méthodologie.

Je pense (et j'espère) que la revue bibliographique et le premier chapitre apporteront un socle utile pour de future études de l'évolution de la myxospermie et que mon équipe d'accueil et celle de Fabrice Roux réussiront à valoriser les pistes prometteuses mise en avant dans le chapitre 2.

Bibliographie générale

Abdullah HM, Akbari P, Paulose B, Schnell D, Qi W, Park Y, Pareek A, Dhankher OP (2016) Transcriptome profiling of *Camelina sativa* to identify genes involved in triacylglycerol biosynthesis and accumulation in the developing seeds. Biotechnol Biofuels **9**: 136

Adlassnig W, Lendl T, Peroutka M, Lang I (2010) Deadly glue — Adhesive traps of carnivorous plants. In Biological adhesive systems (pp. 15–28). Springer, Vienna.

Albenne C, Canut H, Hoffmann L, Jamet E (2014) Plant cell wall proteins: A large body of data, but what about runaways? Proteomes **2**: 224–242

Alvarado-Cárdenas LO, Martínez-Meyer E, Feria TP, Eguiarte LE, Hernández HM, Midgley G, Olson ME (2013) To converge or not to converge in environmental space: Testing for similar environments between analogous succulent plants of North America and Africa. Ann Bot **111**: 1125–1138

Anderson CT, Kieber JJ (2020) Dynamic construction, perception, and remodeling of plant cell walls. Annu Rev Plant Biol **71**: 39–69

Arber A (1910) On the structure of the Palaeozoic seed *Mitrospermum compressum* (Will .). Ann Bot **24**: 491–509

Archibald JM (2015) Endosymbiosis and eukaryotic cell evolution. Curr Biol **25**: R911–R921

Arshad W, Lenser T, Wilhelmsson PKI, Chandler JO, Steinbrecher T, Marone F, Pérez M, Collinson ME, Stuppy W, Rensing SA, et al (2021) A tale of two morphs: developmental patterns and mechanisms of seed coat differentiation in the dimorphic diaspore model *Aethionema arabicum* (Brassicaceae). Plant J **107**: 166–181

Baldwin BG, Sanderson MJ, Porter JM, Martin F, Campbell CS, Donoghue MJ, Its THE, Of R, Baldwin BG, Source AV, et al (1995) The its region of nuclear ribosomal DNA : A valuable source of evidence on Angiosperm phylogeny source. Annals of the Missouri Botanical Garden **82**: 247-277

Bartoli C, Frachon L, Barret M, Rigal M, Huard-Chauveau C, Mayjonade B, Zanchetta C, Bouchez O, Roby D, Carrère S, et al (2018) In situ relationships between microbiota and potential pathobiota in *Arabidopsis thaliana*. ISME J **12**: 2024–2038

Basu D, Tian L, DeBrosse T, Poirier E, Emch K, Herock H, Travers A, Showalter AM (2016) Glycosylation of a fasciclin-like arabinogalactan-protein (SOS5) mediates root growth and seed mucilage adherence via a cell wall receptor-like kinase (FEI1/ FEI2) pathway in Arabidopsis. PLoS One **11**: 1–27

Basu D, Wang W, Ma S, DeBrosse T, Poirier E, Emch K, Soukup E, Tian L, Showalter AM (2015) Two Hydroxyproline galactosyltransferases, GALT5 and GALT2, function in arabinogalactan-protein glycosylation, growth and development in Arabidopsis. PLoS One **10**: e0125624

Bertrand MC-E (1907) Les caractéristiques du genre Diplotesta de Brongniart. Bull la Soc Bot Fr **54**: 389–402

Brown AR, Gordon RA, Hyland SN, Siegrist MS, Grimes CL (2020) Chemical biology tools for examining the bacterial cell wall. Cell Chem Biol **27**: 1052–1062

Burki F, Roger AJ, Brown MW, Simpson AGB (2020) The new tree of Eukaryotes. Trends Ecol Evol **35**: 43–55

Busch A, Hess S (2021) Sunscreen mucilage: a photoprotective adaptation found in terrestrial green

algae (Zygnematophyceae). Eur J Phycol **00**: 1–18

Carafa A, Duckett JG, Ligrone R (2003) Subterranean gametophytic axes in the primitive liverwort Haplomitrium harbour a unique type of endophytic association with aseptate fungi. New Phytol **160**: 185–197

CNRT (2012) Centre National de Ressources Textuelles et Lexicales https://www.cnrtl.fr/definition/mucilage

Coen O, Magnani E (2018) Seed coat thickness in the evolution of angiosperms. Cell Mol Life Sci **75**: 2509–2518

Darwin C (1859) L'origine des espèces.

Dawkins R (1976) Le gène égoïste

Debeaujon I, Lepiniec L, Pourcel L, Routaboul JM (2007) Seed Coat Development and Dormancy. In Annual plant reviews, seed development, dormancy and germination (pp. 25–49)

Domozych DS (1999) Disruption of the Golgi apparatus and secretory mechanism in the desmid *Closterium acerosum*, J Exp Bot **50**: 1323–1330

Doyle JA (2013) Phylogenetic analyses and morphological innovations in land plants. Annu. Plant Rev. **45**: 1–50

Driouich A, Follet-Gueye ML, Bernard S, Kousar S, Chevalier L, Vicré-Gibouin M, Lerouxel O (2012) Golgi-mediated synthesis and secretion of matrix polysaccharides of the primary cell wall of higher plants. Front Plant Sci **3**: 79

Duckett JG, Carafa A, Ligrone R (2006) A highly differentiated glomeromycotean association with the mucilage-secreting, primitive antipodean liverwort Treubia (Treubiaceae): Clues to the origins of mycorrhizas. Am J Bot **93**: 797–813

Eder M, Lütz-Meindl U (2008) Pectin-like carbohydrates in the green alga Micrasterias characterized by cytochemical analysis and energy filtering TEM. J Microsc **231**: 201–214

Eder M, Lütz-Meindl U (2010) Analyses and localization of pectin-like carbohydrates in cell wall and mucilage of the green alga *Netrium digitus*. Protoplasma **243**: 25–38

Ehling-Schulz M, Bilger W, Scherer S (1997) UV-B-induced synthesis of photoprotective pigments and extracellular polysaccharides in the terrestrial cyanobacterium *Nostoc commune*. J Bacteriol **179**: 1940–1945

Evenari M (1984) Seed physiology: From ovule to maturing seed. Bot Rev **50**: 143–170

Fabrissin I, Cueff G, Berger A, Granier F, Sallé C, Poulain D, Ralet M-C, North HM (2019) Natural variation reveals a key role for rhamnogalacturonan I in seed outer mucilage and underlying genes. Plant Physiol **181**: 1498–1518

Frachon L, Bartoli C, Carrère S, Bouchez O, Chaubet A, Gautier M, Roby D, Roux F (2018) A genomic map of climate adaptation in *Arabidopsis thaliana* at a micro-geographic scale. Front Plant Sci **9**: 1–15

Frachon L, Mayjonade B, Bartoli C, Hautekèete NC, Roux F (2019) Adaptation to plant communities across the genome of *Arabidopsis thaliana*. Mol Biol Evol **36**: 1442–1456

Francoz E, Ranocha P, Nguyen-Kim H, Jamet E, Burlat V, Dunand C (2015a) Roles of cell wall peroxidases in plant development. Phytochemistry **112**: 15–21

Francoz E, Ranocha P, Burlat V, Dunand C (2015b) Arabidopsis seed mucilage secretory cells: regulation and dynamics. Trends Plant Sci **20**: 515–524

Francoz E, Ranocha P, Le Ru A, Martinez Y, Fourquaux I, Jauneau A, Dunand C, Burlat V (2019) Pectin demethylesterification generates platforms that anchor peroxidases to remodel plant cell wall domains. Dev Cell **48**: 1-16

Frangedakis E, Shimamura M, Villarreal JC, Li FW, Tomaselli M, Waller M, Sakakibara K, Renzaglia KS, Szövényi P (2021) The hornworts: morphology, evolution and development. New Phytol **229**: 735–754

Galatis B, Apostolakos P (1977) On the fine structure of differentiating mucilage papillae of Marchantia. Can J Bot 772–795

Galloway AF, Knox P, Krause K (2020) Sticky mucilages and exudates of plants: putative microenvironmental design elements with biotechnological value. New Phytol **225**: 1461–1469

Gow NAR, Latge JP, Munro CA (2017) The fungal cell wall: Structure, biosynthesis, and function. The Fungal Kingdom 267–292

Griffiths JS, Crepeau M-J, Ralet M-C, Seifert GJ, North HM (2016) Dissecting seed mucilage adherence mediated by FEI2 and SOS5. Front Plant Sci **7**: 1–13

Griffiths JS, Tsai AY-L, Xue H, Voiniciuc C, ola K, Seifert GJ, Mansfield SD, Haughn GW (2014) SALT-OVERLY SENSITIVE5 Mediates Arabidopsis seed coat mucilage adherence and organization through pectins. Plant Physiol **165**: 991–1004

Grove GG, Rothwell GW (1980) *Mitrospermum Vinculum* Sp. Nov., a Cardiocarpalean ovule from the upper pennsylvanian of Ohio. Am J Bot **67**: 1051–1058

Haughn GW, Western TL (2012) Arabidopsis seed coat mucilage is a specialized cell wall that can be used as a model for genetic analysis of plant cell wall structure and function. Front Plant Sci **3**: 1–5

Holden HS (1954) The morphology of a new species of Pteridosperm seed from the Yorkshire coal measures. Ann Bot **18**: 407–415

Hu D, Baskin JM, Baskin CC, Liu R, Yang X, Huang Z (2021) A seed mucilage-degrading fungus from the rhizosphere strengthens the plant-soil-microbe continuum and potentially regulates root nutrients of a cold desert shrub. Mol Plant-Microbe Interact **34**: 538–546

Hu D, Zhang S, Baskin JM, Baskin CC, Wang Z, Liu R, Du J, Yang X, Huang Z (2019) Seed mucilage interacts with soil microbial community and physiochemical processes to affect seedling emergence on desert sand dunes. Plant Cell Environ **42**: 591–605

Hug LA, Baker BJ, Anantharaman K, Brown CT, Probst AJ, Castelle CJ, Butterfield CN, Hernsdorf AW, Amano Y, Ise K, et al (2016) A new view of the tree of life. Nat Microbiol **1**: 1–6

Kagale S, Koh C, Nixon J, Bollina V, Clarke WE, Tuteja R, Spillane C, Robinson SJ, Links MG, Clarke C, et al (2014) The emerging biofuel crop *Camelina sativa* retains a highly undifferentiated hexaploid genome structure. Nat Commun **5**: 1–11

Kreitschitz A, Gorb SN (2018) The micro- and nanoscale spatial architecture of the seed mucilage—comparative study of selected plant species. PLoS One **13**: e0200522

Kreitschitz A, Haase E, Gorb SN (2021) The role of mucilage envelope in the endozoochory of selected plant taxa. Sci Nat. **108**: 2

Laehnemann D, Borkhardt A, McHardy AC (2016) Denoising DNA deep sequencing data-high-throughput sequencing errors and their correction. Brief Bioinform **17**: 154–179

Lampugnani ER, Khan GA, Somssich M, Persson S (2018) Building a plant cell wall at a glance. J

Cell Sci **131**: jcs207373

Leebens-Mack JH, Barker MS, Carpenter EJ, Deyholos MK, Gitzendanner MA, Graham SW, Grosse I, Li Z, Melkonian M, Mirarab S, et al (2019) One thousand plant transcriptomes and the phylogenomics of green plants. Nature **574**: 679–685

Li C, Zhang B, Chen B, Ji L, Yu H (2018) Site-specific phosphorylation of TRANSPARENT TESTA GLABRA1 mediates carbon partitioning in Arabidopsis seeds. Nat Commun **9**: 571

Liu X, Brost J, Hutcheon C, Guilfoil R, Wilson AK, Leung S, Shewmaker CK, Rooke S, Nguyen T, Kiser J, et al (2012) Transformation of the oilseed crop *Camelina sativa* by Agrobacterium-mediated floral dip and simple large-scale screening of transformants. Vitr Cell Dev Biol - Plant **48**: 462–468

Linkies A, Graeber K, Knight C, Leubner-Metzger G (2010) The evolution of seeds. New Phytol **186**: 817–831

LoPresti EF, Pan V, Goidell J, Weber MG, Karban R (2019) Mucilage-bound sand reduces seed predation by ants but not by reducing apparency: a field test of 53 plant species. Ecology **100**: 1–11

Maeda K, Kunieda T, Tamura K, Hatano K, Hara-Nishimura I, Shimada T (2019) Identification of periplasmic root-cap mucilage in developing columella cells of *Arabidopsis thaliana*. Plant Cell Physiol **60**: 1296–1303

Melzer B, Steinbrecher T, Seidel R, Kraft O, Schwaiger R, Speck T (2010) The attachment strategy of English ivy: A complex mechanism acting on several hierarchical levels. J R Soc Interface **7**: 1383–1389

Mendu V, Griffiths JS, Persson S, Stork J, Downie AB, Voiniciuc C, Haughn GW, DeBolt S (2011) Subfunctionalization of cellulose synthases in seed coat epidermal cells mediates secondary radial wall synthesis and mucilage attachment. Plant Physiol **157**: 441–453

Meschke H, Schrempf H (2010) *Streptomyces lividans* inhibits the proliferation of the fungus *Verticillium dahliae* on seeds and roots of *Arabidopsis thaliana*. Microb Biotechnol **3**: 428–443

Minjares-Fuentes R, Medina-Torres L, González-Laredo RF, Rodríguez-González VM, Eim V, Femenia A (2017) Influence of water deficit on the main polysaccharides and the rheological properties of Aloe vera (*Aloe barbadensis* Miller) mucilage. Ind Crops Prod **109**: 644–653

Nazari M (2021) Plant mucilage components and their functions in the rhizosphere. Rhizosphere **18**: 100344

Nguyen T-P, Mühlich C, Mohammadin S, van den Bergh E, Platts AE, Haas FB, Rensing SA, Schranz ME (2019) Genome improvement and genetic map construction for *Aethionema arabicum* , the first divergent branch in the Brassicaceae family. G3 Genes|Genomes|Genetics **9**: 3521–3530

Niklas KJ (2004) The cell walls that bind the tree of Life. Bioscience **54**: 831

Nikolov LA, Shushkov P, Nevado B, Gan X, Al-Shehbaz IA, Filatov D, Bailey CD, Tsiantis M (2019) Resolving the backbone of the Brassicaceae phylogeny for investigating trait diversity. New Phytol **222**: 1638–1651

North HM, Berger A, Saez-Aguayo S, Ralet MC (2014) Understanding polysaccharide production and properties using seed coat mutants: future perspectives for the exploitation of natural variants. Ann Bot **114**: 1251–1263

Nurk S, Koren S, Rhie A, Rautiainen M, Bzikadze A V, Mikheenko A, Vollger MR, Altemose N, Uralsky L, Gershman A, et al (2021) The complete sequence of a human genome. bioRxiv

2021.05.26.445798

Oliveira CS, Salino A, Paiva EAS (2017) Colleters in Thelypteridaceae: Unveiling mucilage secretion and its probable role in ferns. Flora Morphol Distrib Funct Ecol Plants **228**: 65–70

Olivier F. W., Salisbury E. J. (1911) On the structure and affinities of the Palaeozoic seeds of the Conostoma group. Ann Bot **25**: 1–50

Pan VS, McMunn M, Karban R, Goidell J, Weber MG, LoPresti EF (2021) Mucilage binding to ground protects seeds of many plants from harvester ants: A functional investigation. Funct Ecol **00:** 1–13

Phan JL, Burton RA (2018) New insights into the composition and structure of seed mucilage. Annu. Plant Rev. **1**: 1–41

Porter JR (1976) Antony Van Leeuwenhoek. Tercentenary of his discovery of bacteria. Bacteriol Rev **40**: 260–269

Potts M (1994) Desiccation tolerance of prokaryotes. Microbiol Rev **58**: 755–805

Radhakrishnan G V., Keller J, Rich MK, Vernié T, Mbadinga Mbadinga DL, Vigneron N, Cottret L, Clemente HS, Libourel C, Cheema J, et al (2020) An ancestral signalling pathway is conserved in intracellular symbioses-forming plant lineages. Nat Plants **6**: 280–289

Rajjou L, Debeaujon I (2008) Seed longevity: Survival and maintenance of high germination ability of dry seeds. Comptes Rendus - Biol **331**: 796–805

Ren Y, Song Y, Zhang L, Guo D, He J, Wang L, Song S, Xu W, Zhang C, Lers A, et al (2021) Coding of non-coding RNA: Insights into the regulatory functions of pri-microRNA-encoded peptides in plants. Front Plant Sci **12**: 641351

Robins RJ, Hall DO, Shi D -J, Turner RJ, Rhodes MJC (1986) Mucilage acts to adhere cyanobacteria and cultured plant cells to biological and inert surfaces. FEMS Microbiol Lett **34**: 155–160

Ropitaux M, Bernard S, Schapman D, Follet-Gueye M-L, Vicré M, Boulogne I, Driouich A (2020) Root border cells and mucilage secretions of soybean, *Glycine Max* (Merr) L.: Characterization and role in interactions with the oomycete *Phytophthora parasitica*. Cells **9**: 2215

Sasse J, Martinoia E, Northen T (2018) Feed your friends: Do plant exudates shape the root microbiome? Trends Plant Sci **23**: 25–41

Satpathy S, Sen SK, Pattanaik S, Raut S (2016) Review on bacterial biofilm: An universal cause of contamination. Biocatal Agric Biotechnol **7**: 56–66

Schaffner G (1979) Extraflorale Nektarien bei Cuscuta. Ber Dtsch Bot Ges **92**: 721–729

Schlager D (2016) Transgene expression and plant regeneration in the Brassicaceae *Aethionema arabicum*. Master's thesis at the university of Vienna

Soukoulis C, Gaiani C, Hoffmann L (2018) Plant seed mucilage as emerging biopolymer in food industry applications. Curr Opin Food Sci **22**: 28–42

Tosif MM, Najda A, Bains A, Kaushik R, Dhull SB, Chawla P, Walasek-Janusz M (2021) A comprehensive review on plant-derived mucilage: Characterization, functional properties, applications, and its utilization for nanocarrier fabrication. Polymers **13**: 1066

Tsai AYL, Higaki T, Nguyen CN, Perfus-Barbeoch L, Favery B, Sawa S (2019) Regulation of root-knot nematode behavior by seed-coat mucilage-derived attractants. Mol Plant **12**: 99–112

Tsai AYL, Kunieda T, Rogalski J, Foster LJ, Ellis BE, Haughn GW (2017) Identification and

characterization of Arabidopsis seed coat mucilage proteins. Plant Physiol **173**: 1059–1074

Vaughan JG, Whitehouse JM (1971) Seed structure and the taxonomy of the Cruciferae. Bot J Linn Soc **64**: 383–409

Vaughn KC (2002) Attachment of the parasitic weed dodder to the host. Protoplasma **219**: 227–237

Western TL (2012) The sticky tale of seed coat mucilages: production, genetics, and role in seed germination and dispersal. Seed Sci Res **22**: 1–25

Westhoff M, Zimmermann D, Zimmermann G, Gessner P, Wegner LH, Bentrup FW, Zimmermann U (2009) Distribution and function of epistomatal mucilage plugs. Protoplasma **235**: 101–105

Willing EM, Rawat V, Mandáková T, Maumus F, James GV, Nordström KJV, Becker C, Warthmann N, Chica C, Szarzynska B, et al (2015) Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation. Nat Plants **1**: 1–7

Zhang RJ, Peng WJ, Liu WZ, Li XD, He YK (2016) From spore germination to gametophyte development: The culture, propagation and anatomical protonemal structure of *Takakia lepidozioides* (Bryophyta) in Tibet plateau. Cryptogam Bryol **37**: 383–397

Zhang T, Huang L, Wang Y, Wang W, Zhao X, Zhang S, Zhang J, Hu F, Fu B, Li Z (2017) Differential transcriptome profiling of chilling stress response between shoots and rhizomes of *Oryza longistaminata* using RNA sequencing. PLoS One **12**: 1–21