



HAL
open science

MuSeeQ, a novel supervised image analysis tool for the simultaneous phenotyping of the soluble mucilage and seed morphometric parameters

Fabien Miart, Jean-Xavier J.-X. Fontaine, Christophe Pineau, Hervé Demailly, Brigitte Thomasset, Olivier Van Wuytswinkel, Karine Pageau, François Mesnard

► To cite this version:

Fabien Miart, Jean-Xavier J.-X. Fontaine, Christophe Pineau, Hervé Demailly, Brigitte Thomasset, et al.. MuSeeQ, a novel supervised image analysis tool for the simultaneous phenotyping of the soluble mucilage and seed morphometric parameters. *Plant Methods*, 2018, 14 (1), pp.1-17. 10.1186/s13007-018-0377-5 . hal-01992925

HAL Id: hal-01992925

<https://utc.hal.science/hal-01992925v1>

Submitted on 26 May 2020

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

METHODOLOGY

Open Access



MuSeeQ, a novel supervised image analysis tool for the simultaneous phenotyping of the soluble mucilage and seed morphometric parameters

Fabien Miart^{1,4}, Jean-Xavier Fontaine¹, Christophe Pineau¹, Hervé Demailly², Brigitte Thomasset³, Olivier Van Wuytswinkel¹, Karine Pageau^{1*} and François Mesnard^{1*}

Abstract

Background: The mucilage is a model to study the polysaccharide biosynthesis since it is produced in large amounts and composed of complex polymers. In addition, it is of great economic interest for its technical and nutritional value. A fast method for phenotyping the released mucilage and the seed morphometric parameters will be useful for fundamental, food, pharmaceutical and breeding researches. Current strategies to phenotype soluble mucilage are restricted to visual evaluations or are highly time-consuming.

Results: Here, we developed a high-throughput phenotyping method for the simultaneous measurement of the soluble mucilage content released on a gel and the seed morphometric parameters. Within this context, we combined a biochemical assay and an open-source computer-aided image analysis tool, MuSeeQ. The biochemical assay consists in sowing seeds on an agarose medium containing the dye toluidine blue O, which specifically stains the mucilage once it is released on the gel. The second part of MuSeeQ is a macro developed in ImageJ allowing to quickly extract and analyse 11 morphometric data of seeds and their respective released mucilages. As an example, MuSeeQ was applied on a flax recombinant inbred lines population (previously screened for fatty acids content.) and revealed significant correlations between the soluble mucilage shape and the concentration of some fatty acids, e.g. C16:0 and C18:2. Other fatty acids were also found to correlate with the seed shape parameters, e.g. C18:0 and C18:2. MuSeeQ was then showed to be used for the analysis of other myxospermous species, including *Arabidopsis thaliana* and *Camelina sativa*.

Conclusions: MuSeeQ is a low-cost and user-friendly method which may be used by breeders and researchers for phenotyping simultaneously seeds of specific cultivars, natural variants or mutants and their respective soluble mucilage area released on a gel. The script of MuSeeQ and video tutorials are freely available at <http://MuSeeQ.free.fr>.

Keywords: Soluble mucilage, Seed shape, High-throughput phenotyping, Image analysis, Computer-aided, Low-cost, User-friendly, *Linum usitatissimum*, *Arabidopsis thaliana*, Breeding

*Correspondence: karine.pageau@u-picardie.fr;
francois.mesnard@u-picardie.fr

Karine Pageau and François Mesnard are co-senior authors

¹ Laboratoire de Biologie des Plantes et Innovation, EA-3900, UJV, UFR des Sciences, 33 rue St Leu, 80039 Amiens, France

Full list of author information is available at the end of the article



Background

Seeds of many angiosperms, including *Arabidopsis thaliana* and flax (*Linum usitatissimum* L.), become surrounded by a hydrophilic capsule called seed coat mucilage on imbibition [1–4]. This sticky secretion is usually composed of two distinctive parts with a water-soluble outer layer, poorly adherent to the seed surface, and a strongly adherent inner layer [5–8]. Screens for the mucilage content have been widely used in *Arabidopsis* to find out mutants affected in the mucilage biosynthesis and release [9–11]. Current methods were mainly based on visual appreciation of the adherent part of the mucilage, the latter being easily observable because it is tightly attached to the seed coat of *Arabidopsis* seeds [6]. The soluble part of mucilage is released in all directions and diluted when the seeds are soaked in water. This renders the visual appreciation impossible and the quantification much more complex [9, 12–14]. Access to these mucilage parameters are however crucial to better-understand the ecophysiological role of soluble mucilage layers [6, 16]. Indeed, the majority of the mucilage polysaccharides are contained in this layer [5, 7, 17, 18] and some species release almost only soluble mucilage [9, 15]. It has also been proposed that the soluble mucilage may play a role in the seedling germination capacity [6] and in increasing seed adhesion to soil substrates [1].

Seed coat mucilages are mainly composed of pectic polysaccharides [6, 7, 19, 20]. The classical method to detect seed coat mucilage phenotypes consists in incubating seeds in water containing a specific marker of these pectic polysaccharides such as the Ruthenium Red [5, 7, 21–24]. The dye toluidine blue O is also commonly used to stain the mucilage because of its metachromatic properties [25–27]. Once the seeds soaked, the soluble mucilage disperses in the surrounding medium in all directions and is stained by absorbing the dye. Only major differences can be detected with such a method. A second range of methods consists in the extraction of the soluble mucilage in water. The efficiency of that extraction process is highly variable over a broad range of plant species and varies according to the type of process used [9, 15, 28, 29]. For example, for the *Arabidopsis* reference accession Col-0, the soluble mucilage can be easily recovered by gentle extraction in water. The process of extraction on other natural accessions or mutants often requires the use of dilute chelators such as EDTA or HCL-NaOH [6, 12, 30–33]. Also, a sequential water extraction improves the mucilage extraction efficiency [18, 34, 35], considering a combination of physical parameters such as the time of extraction, the temperature, an ethanol precipitation (or not) and the stirring intensity [18, 34, 36, 37]. Number of mucilage extraction procedures on *Arabidopsis* are summarized, compared and discussed in

[9]. Another method consists in measuring a mucilage indicator value (MIV) corresponding to the viscosity of a hot-water extraction of mucilage [38, 39]. Finally, all these water extraction methods are time consuming and particularly hard to apply for the high-throughput quantitative phenotyping of the soluble mucilage content. A mucilage extrusion on agarose plates may be a good strategy for soluble mucilage phenotyping [16]. Since there is any dye in the gel, this is particularly dedicated to the visual evaluation of the released mucilage for the research of mutants showing major differences, but this is not a precise quantitative method. So far none of these methods are suitable for the high-throughput measurement of the soluble mucilage content.

Recent advances in the field of computer-assisted image processing, called bioimage informatics [40], have facilitated the development of a number of image-analysis algorithms, software packages and image-analysis platforms dedicated to the plant morphological traits analysis [41, 42]. Among the wide range of plant species studied with these methods, many of them deal with the seed shape and the seed development [43–47]. Recently, a method for the quantitative analysis of the mucilage area based on image analysis was developed by [48], but it can only be applied for the phenotyping of the adherent part of the mucilage.

Here, we developed a high-throughput phenotyping method called MuSeeQ for the simultaneous extraction of morphometric data from the seeds and their soluble mucilage areas released on the gel. Flaxseed was chosen as a model species for evaluating and validating the method due its high content in mucilage, phenylpropanoids and fatty acids (α -linolenic acid) [28, 38, 49–53]. The first MuSeeQ step allows the uniform release of mucilage around the seed and its staining via the absorption of the dye toluidine blue O contained in the gel. The second step is a digital image analysis tool performing the detection of several seed shape- and soluble mucilage-related traits. Using a combination of image-processing operators, MuSeeQ automatically identifies and segments seeds and soluble mucilages from the agarose stained gel background. It was found good correlations between automated and manual measurements for the area of the soluble mucilage released and for the area of the seeds. The automated procedure is ready to use to extract 11 relevant morphometric parameters related to the seeds and the soluble mucilages. MuSeeQ was next used for studying the carbon partitioning between the seedcoat and the embryo. Significant correlations were found between the soluble mucilages morphometric parameters and the content of some fatty acids, i.e. C16:0, C18:2 and the ratio C18:3/C18:2. (Interestingly, some fatty acids were also found to correlate with the seed

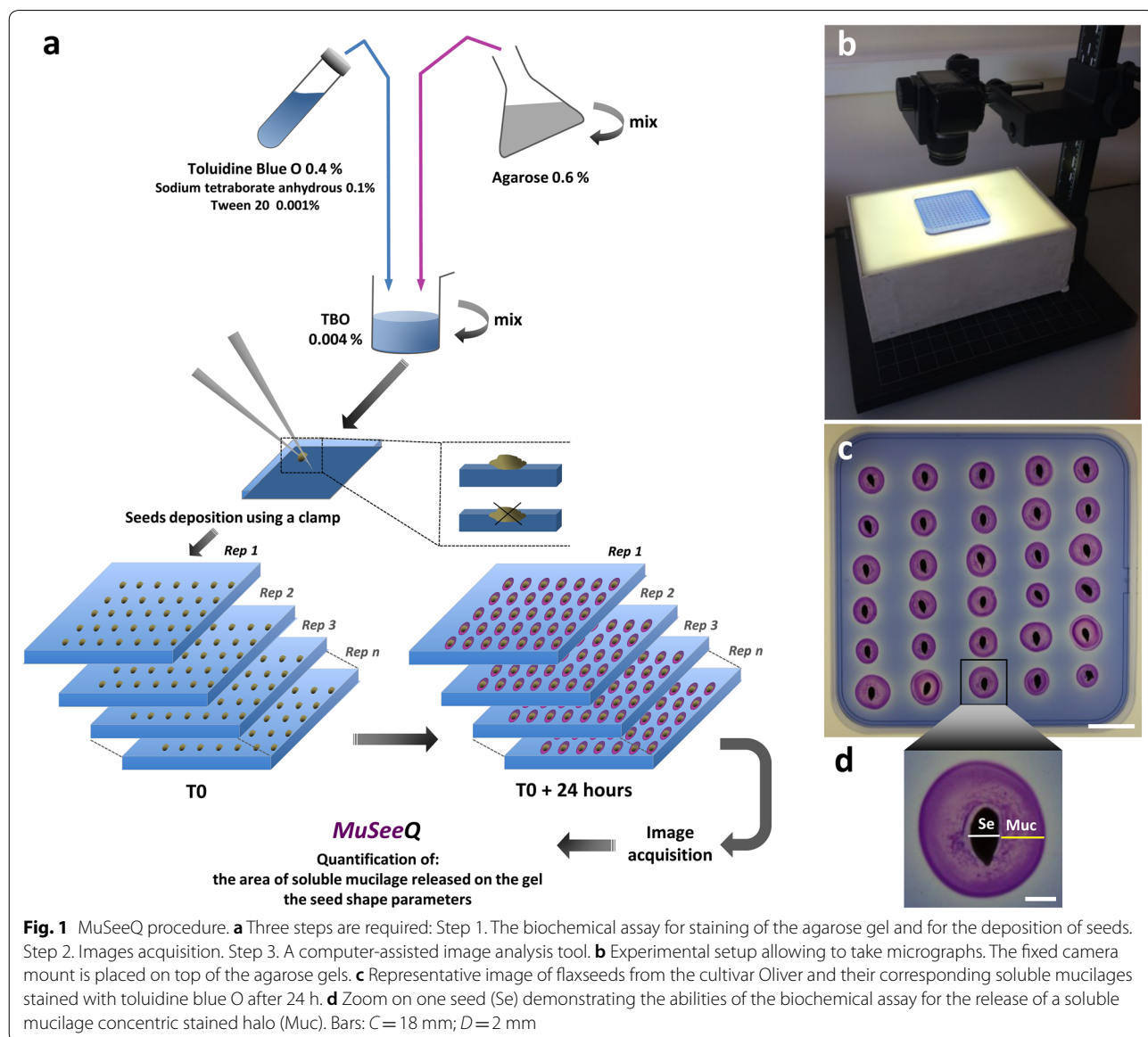
shape parameters, i.e. C18:0, C18:2 and the ratio C18:3/C18:2.) Finally, MuSeeQ was shown to be usable for other mycospermous species, including *Arabidopsis thaliana* and *Camelina sativa*. To conclude, MuSeeQ represents a novel promising low-cost method for the high-throughput phenotyping of the soluble mucilage area and the simultaneous measurement of the seed morphometric parameters.

Results and discussion

MuSeeQ’s biochemical assay: toluidine blue O staining of the mucilage on a gel

The starting point of our phenotyping method consists of obtaining regular mucilage stained halos (example on

flaxseeds: Fig. 1c, d). An agarose gel was chosen as technical support, a method commonly used for imaging root architecture [54, 55]. Almost directly after being deposited on the gel, seeds absorbed the water from the gel and started releasing the mucilage. The dye toluidine blue O was used for staining the translucent mucilage because it turns into a purple pink colour in presence of the pectic mucilage [26]. This stain was more efficient than the dye Ruthenium Red to segment the mucilage from the image background. Seeds have to be carefully deposited on the gel and the released mucilage progressively absorbed the dye toluidine blue O from the gel (Fig. 1a). We experimentally estimated that 24 h were required for a maximum soluble mucilage release and staining efficiency..



MuSeeQ's bioimage informatics: the computer-aided image analysis tool

Introduction to MuSeeQ's bioimage informatics

The image-processing program ImageJ [56] was chosen as a platform to host MuSeeQ because it is in the public domain and it is commonly used by the biologist community; ImageJ is presented as the world's fastest image processing program written in pure Java (<https://imagej.nih.gov/ij/features.html>). It supports and deals with a wide range of file formats and works on many operating systems. Moreover, the open source platform Fiji (Fiji Is Just ImageJ) [57] provides many powerful image analysis solutions and a curated selection of plugins. MuSeeQ was thus written in the ImageJ's Java-like macro language to be easily implemented in Fiji as a module. An additional movie file shows how to install MuSeeQ on Fiji [see Additional file 1].

Image acquisition and pre-processing

Image capture is a key factor of the accuracy of the high-throughput phenotyping method. To optimize the resolution of the input image, MuSeeQ was applied on digital images of agarose stained gels acquired using a high-resolution flatbed photo scanner [44, 47] or using standard consumer cameras. The image resolution was found to be higher when using standard consumer cameras, including low-budget imaging setups such as built-in phone cameras. However, high-resolution digital cameras will be preferred in the case of small seeds as it is the case for *Arabidopsis*. The better contrast between the soluble mucilage, the seed and the background, crucial for the segmentation process, was obtained using a fixed camera above the agarose stained gel placed on a light box (Fig. 1b).

The user needs to provide information only at the pre-processing step. Once the digital image is opened and MuSeeQ activated, an exact conversion of the image pixels in the calibrated millimeter unit is required. MuSeeQ requests the user to select both extremities of the agarose stained gel (Fig. 2; Additional file 2). If the dimensions of the square Petri dish are known, the user can directly enter the real dimensions and accept to continue the process. If unknown, the user has to place a graduated ruler next to the Petri dish. The segmentation steps applied by MuSeeQ are based on the identification of objects having a specific pixel colour intensity. To avoid any problems with the edges of the plates, i.e. to detect edges and assimilate them as objects of interest, the user can select a rectangular window within the plate and crop it.

Segmentation settings and image processing

It is possible in an 8-bit grayscale image to discriminate two pixels by measuring their difference in term of

brightness values ranging from 0 to 255. In colour images, two pixels of distinct colours can have the same brightness. The solution to distinguish pixels with the same brightness consists of expressing images in the RGB 3-D cubic space as a composite of the three primary colours, i.e. red, blue and green. Each colour channel can be independently selected and segmented. However, the HSV colour space (for hue, saturation and brightness value), allows to obtain a better and more natural depiction of colours [58–61]. Although segmentation is not a trivial task for a low signal-to-noise ratio [40], the HSV colour space also demonstrated its higher ability to discriminate plant organs from their background [62–64]. Due to the close colours between flaxseeds (brown to dark brown), soluble mucilages (purple-pink) and the agarose stained gels background (light blue), our algorithm processes input images in the HSV colour space.

Depending on the goal, a number of commonly used methods are available for the segmentation of plant objects in digital images [65], such as noise reduction, skeletonization or thresholding. Threshold-based methods generally require to automatically or manually set fixed threshold levels [66–68]. These methods make irrevocable decisions on the fact that a pixel is part or not of the object of interest and are more likely prone to make errors. To overcome recurrent problems such as an uneven illumination of the image or the background and non-homogenous feature intensities, adaptive threshold solutions have been developed and are freely available on Fiji [57]. Here, the threshold is computed locally pixel by pixel based on the image characteristics on a round window with a defined radius. Dynamic approaches also allow to automatically detect image objects whatever the lighting conditions and the monitoring system by fitting a topological or geometrical model of the targeted object to the input data [54, 69, 70]. These top-down methods generally consist of the peak detection into the histogram of all colour components [71, 72], to individualize and segment plant objects showing important differences in terms of contrast, such as soil, leaves, hypocotyls and roots. When working in the HSV colour space, an approximation of the Gaussian curves describing the green (hue) component is sufficient to segment the green parts of plants such as rosettes [64, 73]. In our case, it is impossible to assimilate the soluble mucilage or the seed only to peaks near the hue value component. Consequently, the histograms of hue, brightness and saturation colour components have to be approximated to determine peaks corresponding to each object of interest.

Identification and segmentation tasks may also have been performed using the more and more popular neural networks. This type of machine learning methods provides state-of-the-art performances [74, 75]. Basically,

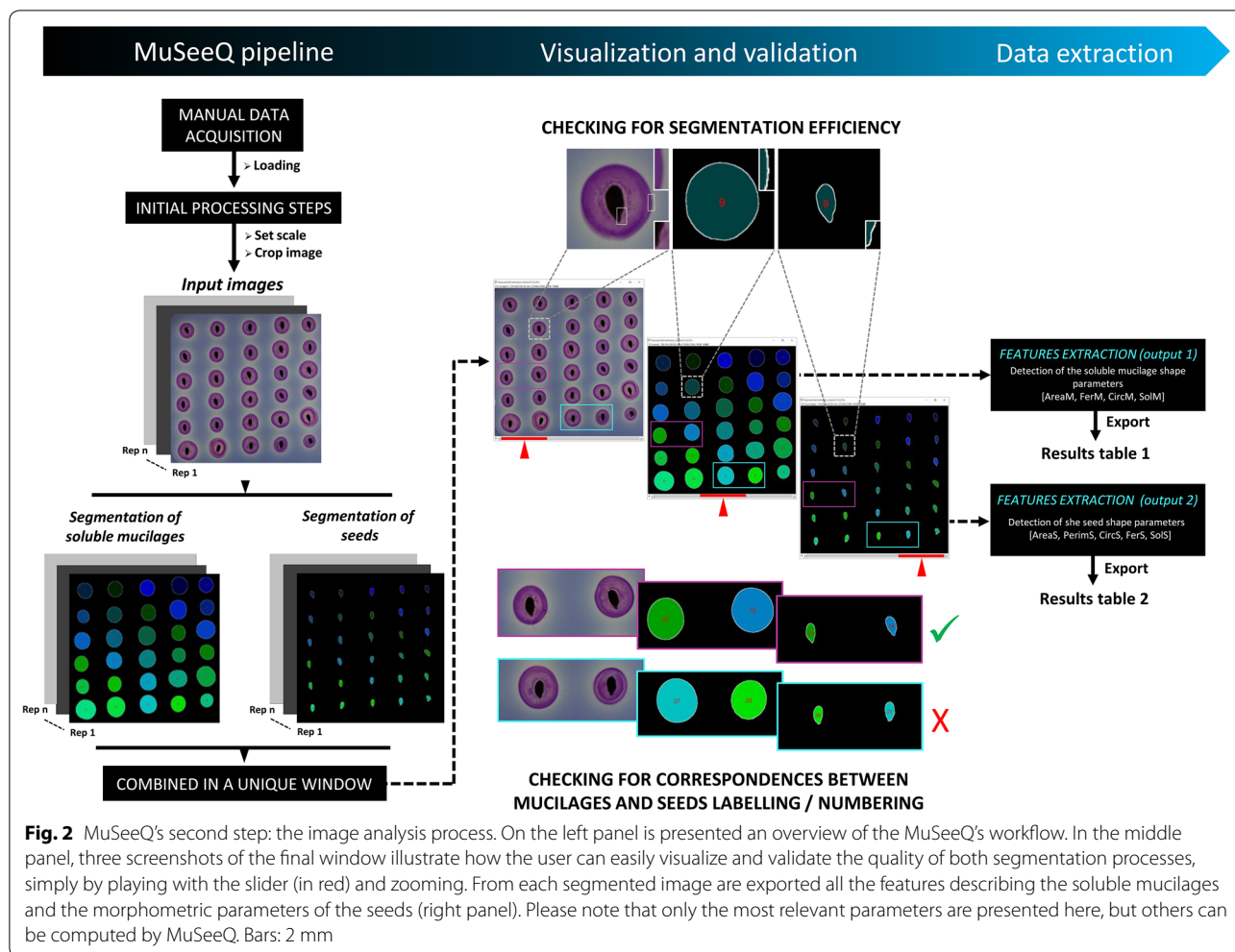


Fig. 2 MuSeeQ’s second step: the image analysis process. On the left panel is presented an overview of the MuSeeQ’s workflow. In the middle panel, three screenshots of the final window illustrate how the user can easily visualize and validate the quality of both segmentation processes, simply by playing with the slider (in red) and zooming. From each segmented image are exported all the features describing the soluble mucilages and the morphometric parameters of the seeds (right panel). Please note that only the most relevant parameters are presented here, but others can be computed by MuSeeQ. Bars: 2 mm

features are computed from raw images and the output of the features is used to extract and separate classes, such as seed morphometric parameters, using a classifier. Among the machine learning methods, Convolutional Neural Networks (CNNs) have become popular for unsupervised classification [76, 77]. Among CNNs, we find Region-based CNNs (R-CNNs) [78] which classifies regions using deep CNN. Nevertheless, this method is very slow to apply. Fast R-CNN was developed to overcome this limitation but the latter is based on a selective search algorithm. Faster R-CNN [79] and R-FCN [80] were also developed to improve the speed and accuracy of the process. While the first method is time-consuming when a large number of proposals is used in its first-stage, R-FCN might be too local to be discriminative enough. Deep machine learning methods have proved nevertheless its performance in image-based plant phenotyping for fully automated quantitative trait identification and localization [75] as well as plant disease detection and diagnosis [80]. Despite the unrivaled accuracy of these

techniques, they often require hundreds, sometimes thousands of images and diverse datasets to be trained [81]. Moreover, differences between the image format, size and depth of the dataset used for the training of the network and that used for the analysis can also lead to a misclassification and segmentation troubles for non experts, which led us to avoid CNNs.

In our study, there is no necessity for the power of a top-down thresholding approach or machine learning methods. The algorithm used by MuSeeQ was thought as the combination of the global threshold-based and dynamic segmentation principles. We first modified the G. Landini Colour Threshold function based on an algorithm derived from the iterative intermeans method (IsoData algorithm; [82]). Then, we manually trained our threshold algorithm to a set of digital images. This allowed us to detect peaks in the hue, saturation and brightness value histograms describing as much as possible pixels corresponding to the soluble mucilages and seeds (Additional file 3). Finally, the range of pixel

intensity values describing peaks in each of the colour components were computed in our algorithm as default threshold values.

Once digital input images are opened (just by dragging and dropping them in the Fiji toolbar) and scale setting done, the algorithm duplicates the row image (Fig. 2, 3) and processes independently both daughter images. On the first one, MuSeeQ splits the original image in three grayscale channels corresponding to the H, S and V colour components with pixel values ranging from 0 to 255 (Additional file 3). To segment the soluble mucilages, the default threshold values are set to 172, 24 and 85 for H, S and V, respectively (Additional files 3a, 9). If the quality of the threshold is unsatisfactory, the user is asked to refine the range of the default threshold values previously fixed in the tool box. Because the saturation colour component did not influence in a significant way pixels describing the mucilage, we chose to keep only the H and V colour components as manually adjustable parameters (Additional file 2). Segmentation of the soluble mucilage has to consider the entire area of the soluble mucilage (Fig. 3a–c), which comprises the pixels present in the soluble mucilage, between the outline of the external circle and the border of the seed, and the pixels detected in the seed area. This segmented area, AreaM, corresponds to all of the pixels contained inside the perimeter defined by the soluble mucilage outlines (Fig. 3d). Finally, we computed a predetermined range of size for the halo of soluble mucilage, from 18 to 315 mm², which allows to automatically delete two or more soluble mucilages overlapping on the gel. For seed segmentation, MuSeeQ exactly applies the same process on the second image duplicata (Figs. 2, 3e–g). Only the default threshold values, based on the peaks defining each colour component, are modified and set to 255, 255 and 70 for H, S and V, respectively (Additional file 3). At the end of the process, only the threshold values corresponding to the V component can be manually modified by the user to improve the quality of the seed segmentation. Due to the macro format of MuSeeQ, the user can easily modify the default threshold values for the seeds and soluble mucilages (Additional file 4). But to simplify as much as possible the use of our tool for those that are not familiar with Fiji, different versions of the macro are proposed, each one well-suited to five important myxospermous species (*Arabidopsis thaliana*, *Linum usitatissimum* L., *Camelina sativa*, *Plantago major* and *Capsella bursa-pastoris*), freely available at [83].

Segmentation display and validation

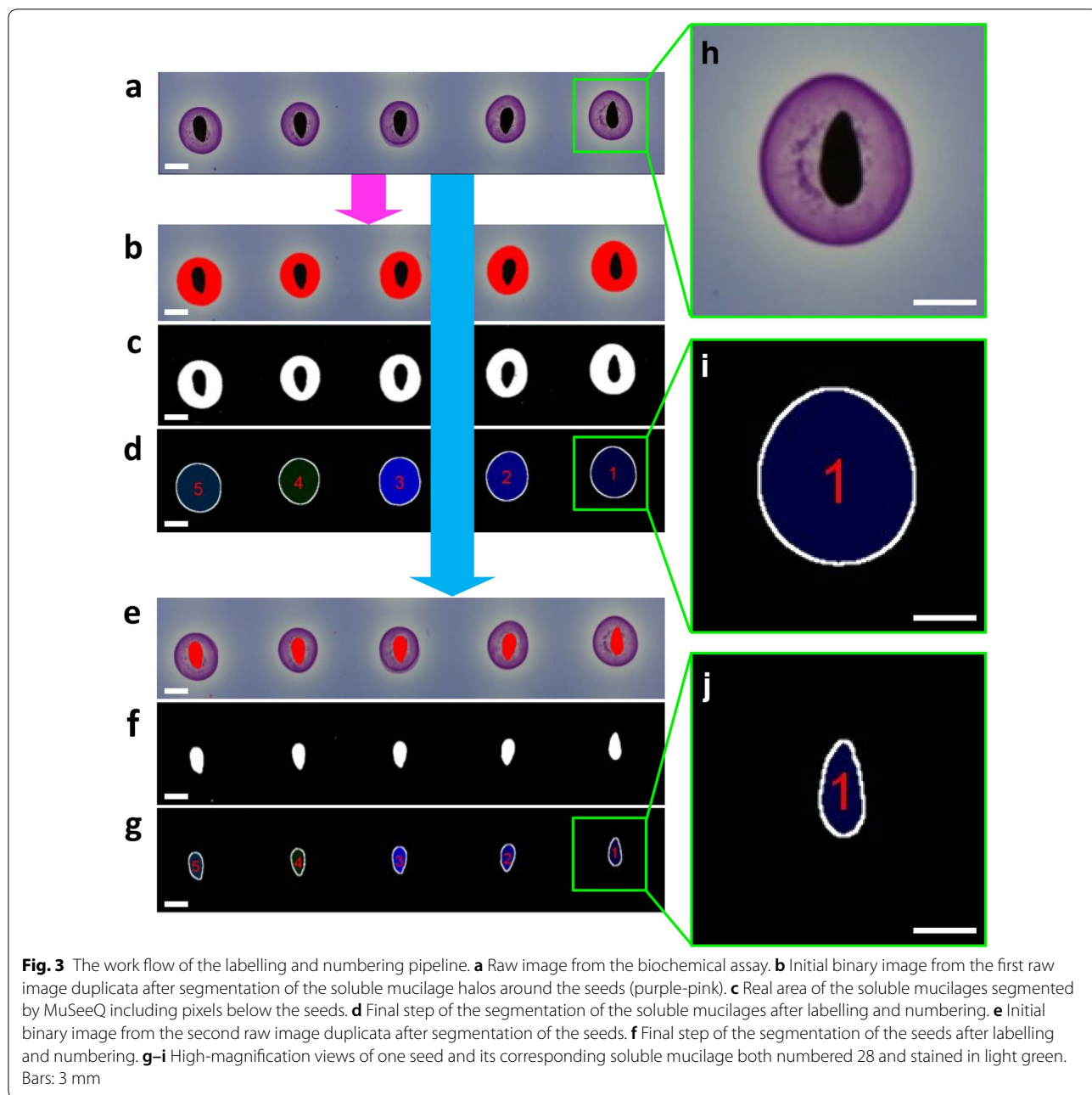
MuSeeQ proposes an intuitive way to validate segmentation processes. It assigns a specific number to the soluble mucilage of every analysed seed, allowing the user to

easily search the measured parameters in the tables of results (Figs. 2, 3h–j, Additional file 2). The same process is applied in parallel for the seeds. So far, despite the fact that no correlation was found between the area of the seeds and the quantity of soluble mucilage extracted [84], we assumed that the soluble mucilage area may potentially be linked to the seed size and/or the seed shape. We decided to normalize the area of each soluble mucilage by the area of their respective seed (AreaS). The quantification of the soluble mucilage area was assimilated to the ratio (AreaM/AreaS), i.e. the «adjusted mucilage content». Combining both kinds of data in order to compute this ratio can not be made automatically by MuSeeQ. To do so, the easiest way consists in using the labelling and numbering of the segmented seeds and soluble mucilages. MuSeeQ attributes colours and numbers to seeds and mucilages line by line from the top to the bottom of the image. For a perfect matching between seeds and mucilages, seeds have to be sown on the agarose gel according to a diagonal. Thus, the uppermost seed and its corresponding mucilage will be numbered “one” and the lowest seed and its mucilage will be associated to the last number. Therefore, if seeds of the same size do not release the same quantity of soluble mucilage, the soluble mucilages with the biggest areas may be detected first by MuSeeQ, which would lead to a mismatch between seeds and mucilages numbers and colours. For example, seed no. 28 corresponds to mucilage no. 27 (Fig. 2). To solve this problem and because it is not necessarily straightforward for a beginner to deposit the seeds properly on the medium, we propose several templates for the seed deposition of three major species (Additional file 5).

Since threshold pixels corresponding to the seeds and their mucilages overlap (Fig. 3a–i), another contribution of this work for the validation of the segmentation results was to display them separately. Once the segmentation steps finished, the three screens corresponding to the input image, the segmented mucilages and the segmented seeds are combined together in a unique window (Fig. 2). For checking the quality of both segmentations as well as the correct matching between labels, the user is finally asked to zoom on the objects of interest and play with the slider to move among the three images (Fig. 2, Additional file 2).

Extraction of the parameters of interest

The measurements performed by MuSeeQ are extracted from input images and exported in two tables of results, either for the soluble mucilage- or seed-related parameters (Fig. 2). MuSeeQ computes eleven mathematical shape descriptors. Five of the most relevant, both for the seed and the soluble mucilage, are summarized in Table 1 with their respective unit, description and abbreviation.



Regarding the mucilage, the most important one corresponds to the area of the surface of released soluble mucilage in 2D (AreaM). MuSeeQ also computes the area of the seeds surface projected in 2D (AreaS), a classical way to access the seed size [44, 47].

Interest in computing our algorithm under the Java-like macro language resides in its flexibility, which implies that the user can easily increase the number of parameters to analyse by simply adding them into the script (Additional file 6). MuSeeQ can also be complemented

by a high number of plugins and macros in Fiji such as editing, colour processing, measures or image enhancement [85, 86].

Statistical validation

First, we checked the technical reproducibility of the biochemical assay. Seeds from the cultivar Oliver were sown on three different biochemical assays. No statistical difference was found between the technical replicates for AreaM, AreaS and the ratio (AreaM/

Table 1 Overview of the main soluble mucilage- and seed shape-related parameters computed by MuSeeQ

Traits	MuSeeQ designation	Description	Code
Seed	Area (mm ²)	The area of the seed surface projected in 2D. Computed as: $\text{AreaS} = \pi \times \text{radius}^2$	AreaS
	Circularity	Corresponds to a more or less elongated seed shape. Circularity is comprised between 0.0 and 1.0. A value of 1.0 indicates a perfectly round seed. Computed as: $\text{CircS} = 4\pi(\text{AreaS}/\text{PeriS}^2)$	CircS
	Feret (mm)	The longest distance between two points of the segmented seed. For seeds, it can be interpreted as the longest distance between the peduncle and the awn	FerS
	Perimeter (mm)	The length of the limit of the segmented seed. Computed as: $\text{PeriS} = 2 \times \pi \times \text{radius}$	PeriS
	Solidity	Corresponds to the ratio between the exact and the approximated seed area, using a sort of rubber band all around the segmented seed. Computed as: $\text{SolS} = \text{AreaS}/\text{Convex AreaS}$	SolS
Soluble Mucilage	Area (mm ²)	The area of the surface of released soluble mucilage projected in 2D. Computed as: $\text{AreaM} = \pi \times \text{radius}^2$	AreaM
	Circularity	Corresponds to a more or less elongated mucilage shape. Circularity is comprised between 0.0 and 1.0. A value of 1.0 indicates a perfectly round mucilage. Computed as: $\text{CircM} = 4\pi(\text{AreaM}/\text{PeriM}^2)$	CircM
	Perimeter (mm)	The length of the limit of the segmented mucilage. Computed as: $\text{PeriM} = 2 \times \pi \times \text{radius}$	PeriM
	Solidity	Corresponds to the ratio between the exact and approximated mucilage areas using a sort of rubber band all around the segmented mucilage. Computed as: $\text{SolM} = \text{AreaM}/\text{Convex AreaM}$	SolM
Ajusted mucilage content		The 2D surface of released mucilage corrected by its corresponding seed surface projected in 2D	Ratio (AreaM/ AreaS)

Please note that the number of parameters which could be analysed can be increased using other plugins and functions in ImageJ

AreaS) (Fig. 4a–c). In order to statistically demonstrate the efficiency of MuSeeQ for the precise automated segmentation of the mucilages and seeds, the most relevant parameters, i.e. AreaM, AreaS, and the ratio (AreaM/AreaS), were measured on seeds of 21 randomly selected recombinant inbred lines (RILs) from a cross between the cultivars Oliver and Viking using MuSeeQ, and compared to manual measurements using Fiji. The Pearson's product-moment correlation analyses between automated (MuSeeQ) and manual (Fiji) measurements have demonstrated strong correlations for AreaM ($r_2 = 0.9970$; $P < 0.0001$), AreaS ($r_2 = 0.9743$; $P < 0.0001$) (Additional file 7), and the ratio (AreaM/AreaS) ($r_2 = 0.9962$; $P < 0.0001$) (Fig. 4d).

It is currently not possible to statistically compare our results obtained on the area of the soluble mucilage with other image-based analysis methods since they do not exist yet. However, SmartGrain is an image analysis software dedicated to the quantification of the seed shape parameters [45]. To compare the results obtained with MuSeeQ versus SmartGrain, we applied a Bland–Altman comparison method on the most important seed shape parameter, i.e. AreaS, which confirms the precision of the measurements using MuSeeQ (Fig. 4f). The measurements performed with MuSeeQ slightly exceed those obtained with SmartGrain. This bias is explained by the scale setting step which can slightly vary between the two

softwares. Nevertheless, a Pearson's product-moment correlation analysis

$$r = \frac{N \sum xy - (\sum x)(\sum y)}{\sqrt{[N \sum x^2 - (\sum x)^2][N \sum y^2 - (\sum y)^2]}}$$

on the same dataset reveals a highly acceptable r_2 ($r_2 = 0.8961$; $P < 0.001$) (Fig. 4g).

Evaluation of the technical efficiency

When flaxseeds are placed on a biochemical assay both the soluble and adherent parts of the mucilage seem to be released on the gel (Fig. 5a). To check it, flaxseeds were immersed in water for 24 h in order to extract the soluble part of the mucilage [7, 84] and seeds were then placed on a biochemical assay. Since a thin halo of mucilage is still detected around the seed (Fig. 5b), this means that our biochemical assay does not allow the full release of the adherent part of the mucilage. However, we assume that it can be used as a high-throughput phenotyping of the soluble mucilage of flaxseeds.

To validate the efficiency of MuSeeQ to perform precise phenotyping of the soluble mucilage, we chose to apply our tool on three randomly selected flax RILs, i.e. the RILs 84, 44 and 62. The soluble mucilage content was visually accessed using the toluidine blue O water staining method [26]. Here, seeds are directly soaked in water

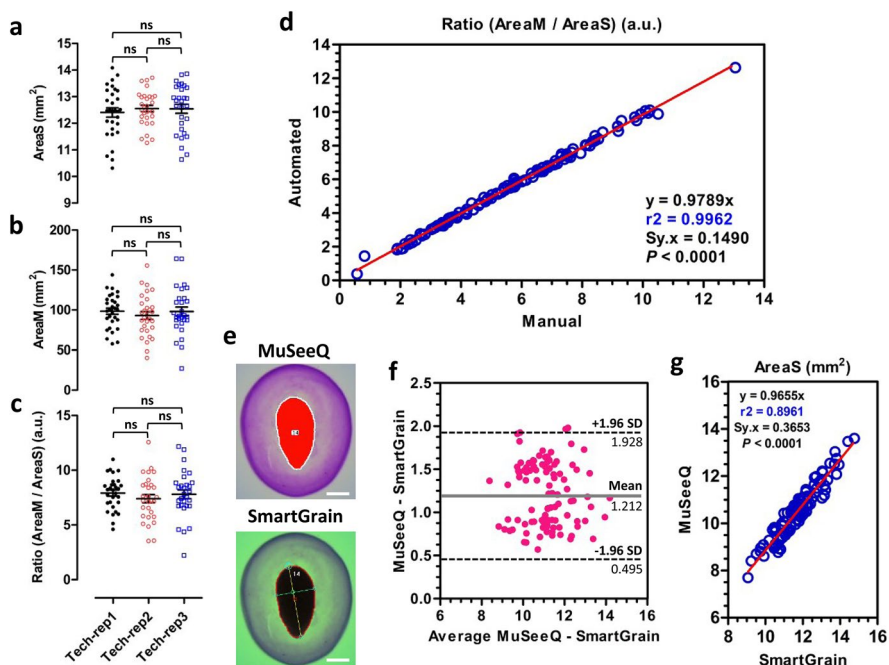


Fig. 4 Statistical validation of MuSeeQ. **a–c** Technical reproducibility assays for the measurement of the area of the seeds (**a**), the area of released soluble mucilage (**b**) and the ratio (AreaM/AreaS) (**c**) on seeds from the cultivar Oliver sown on three different agarose stained gels. ns indicate no significant differences ($p \leq 0.05$) as determined by one-way ANOVA followed by a Tukey’s HSD (honest significant difference) post hoc test ($n = 30$ seeds for each technical replicat). Bars represent mean \pm SEM **d** Correlation analysis of automated (MuSeeQ) versus manual (Fiji) measurements of the ratio (AreaM/AreaS) ($n = 104$ seeds from 21 randomly selected flax RILs, 5 seeds each). **e** A representative result of segmentation using MuSeeQ and SmartGrain on the same seed. Note that both image analysis tools numbered the seed 14. Opposite to SmartGrain, MuSeeQ does not draw the computed parameters on the seed, except for the seed outlines. **f** Bland–Altman plot showing no major statistical differences between the results of AreaS obtained using both methods. **g** Correlation analysis of automated (MuSeeQ) versus manual (SmartGrain) measurements of the AreaS. Pearson’s correlation coefficients were determined by linear regression analyses ($n = 104$ seeds from 21 randomly selected flax RILs, 5 seeds each). Red lines are least squares regression curves. Bars: 1.5 mm

containing the dye toluidine blue O. The soluble mucilage quickly diffuses in all directions and is stained by absorbing the dye. Differences in the phenotypes between the three RILs were not clearly obvious, with «soluble mucilage strips» more or less thick depending on the RILs, and were impossible to quantify (Fig. 5c). Using the MuSeeQ’s biochemical assay, differences between RILs became clearer (Fig. 5d). Applying our image analysis tool, we were able to measure and statistically discriminate RILs from each other (Fig. 5e).

Another goal of our phenotyping tool consists of considerably reducing the time of analysis. We estimate that

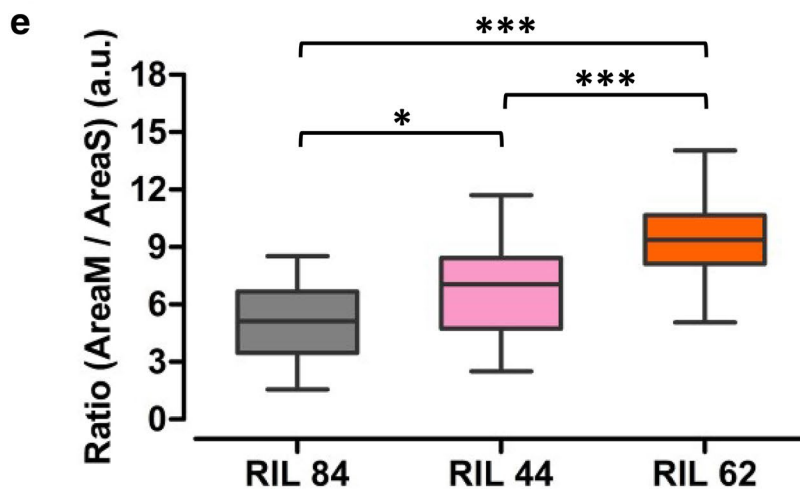
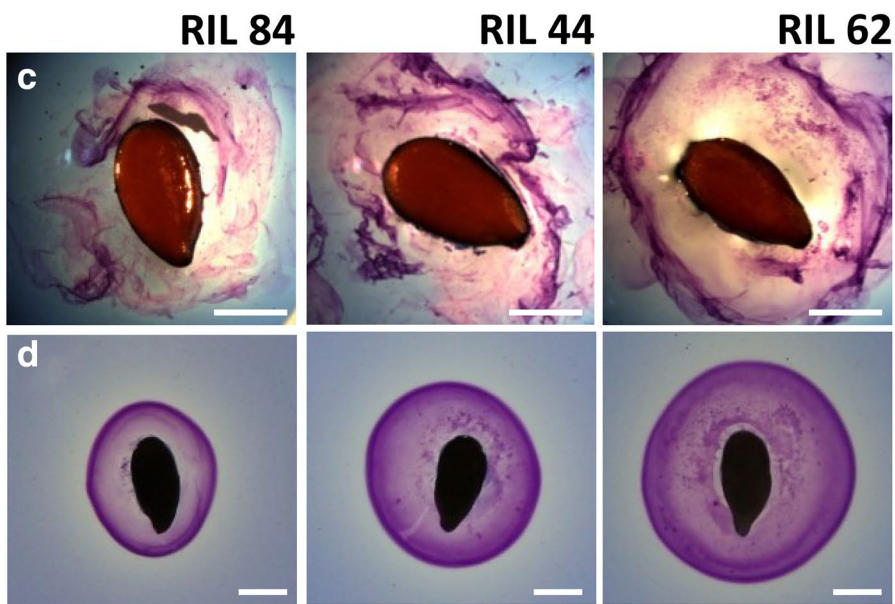
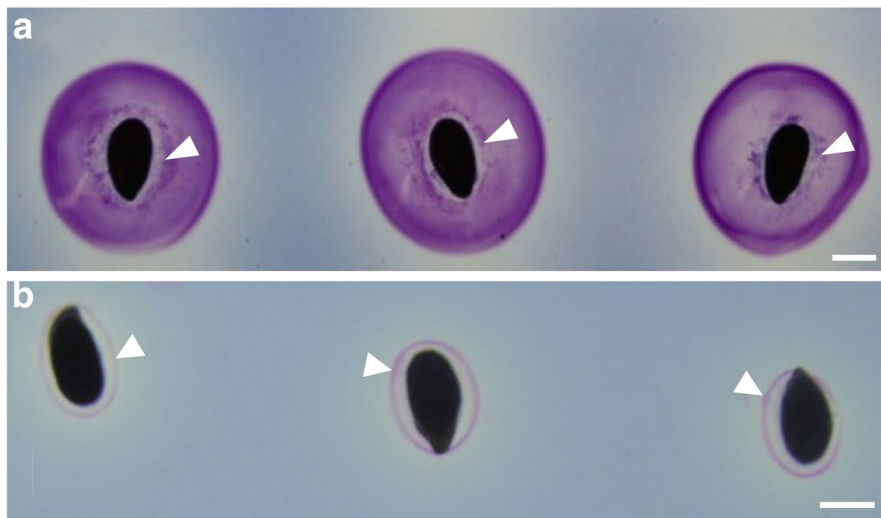
MuSeeQ allows to analyse the soluble mucilage area at least 40 times faster compared with a mucilage extraction in water. Compared with SmartGrain [45], the seed shape parameters analysis takes the same time.

MuSeeQ, a tool to better-understand the seed physiology

It has been proposed in *Arabidopsis* that during the seed development the main carbon source from photosynthesis could be shared between the seed coat, for the mucilage polysaccharides biosynthesis, and the embryo, for the oil biosynthesis [87]. However, functional studies of *tt2* mutants have shown that a significant reduction of

(See figure on next page.)

Fig. 5 MuSeeQ allows to quantitatively discriminate between close soluble mucilage phenotypes. Pictures showing flaxseeds sown on **(a)** a biochemical assay directly or **(b)** after 24 h of soluble mucilage extraction in water. In both cases, pictures have been taken after 24 h of mucilage release on the gels. White arrows indicate that the seeds are capable of releasing the adherent part of the mucilage on the gel **(b)**. **c** Soluble mucilage phenotyping in water containing toluidine blue O after 24 h of release. **d** Soluble mucilage phenotyping using the MuSeeQ’s biochemical assay after 24 h of release. Seeds in **(c)** and **(d)** were chosen as the most representative of the corresponding RILs. **e** Statistical analysis of the ratio (AreaM/AreaS) showing different phenotypes among the three RILs. Significant differences were determined using a one-way ANOVA followed by a Tukey’s HSD post hoc test, * $p < 0.05$; *** $p < 0.001$ ($n > 21$ seeds for each RIL). Bars: 2 mm



the seed oil content in the embryo and a modification of the fatty acids (FA) composition do not affect the seed coat mucilage production [88]. Other studies on flax have also revealed that the mucilage would not be correlated with the oil and protein contents [39, 89]. This suggests disconnected and/or parallel regulatory pathways for the mucilage and oil biosynthesis, and emphasizes the need for a better understanding of these complex mechanisms. In this context, MuSeeQ was applied on a RILs population in F8 derived from the cross between the cultivars Oliver and Viking in order to detect putative correlations between seed shape- and soluble mucilage-related traits. At least, eleven shape descriptors were computed on the seeds and the mucilages (Additional file 2), but the most important are shown in Fig. 6. The RILs population was screened in parallel for its FA content. The oil content

and sixteen FAs ratios were also added to the analysis since the ratio omega-3 to omega-6, i.e. C18:3/C18:2, has been found to play an important role in human health by preventing many diseases [90, 91]. In general, the Pearson's product-moment correlation coefficients observed are small (Fig. 6; Additional files 8, 9).

It has been shown that the soluble mucilage content could be linked to the seed shape in *Arabidopsis* [92–96]. Looking at the correlations matrix between seed shape parameters and other traits, the solidity of the seed (SolS), which indicates how curved the borders of the seeds are (Table 1), was found to positively correlate with AreaM and the ratio (AreaM/AreaS) (Additional files 8, 9). It suggests that the more flaxseed release soluble mucilage, the less flaxseed present curvarures on both sides. Among FAs, C18:0, C18:2,

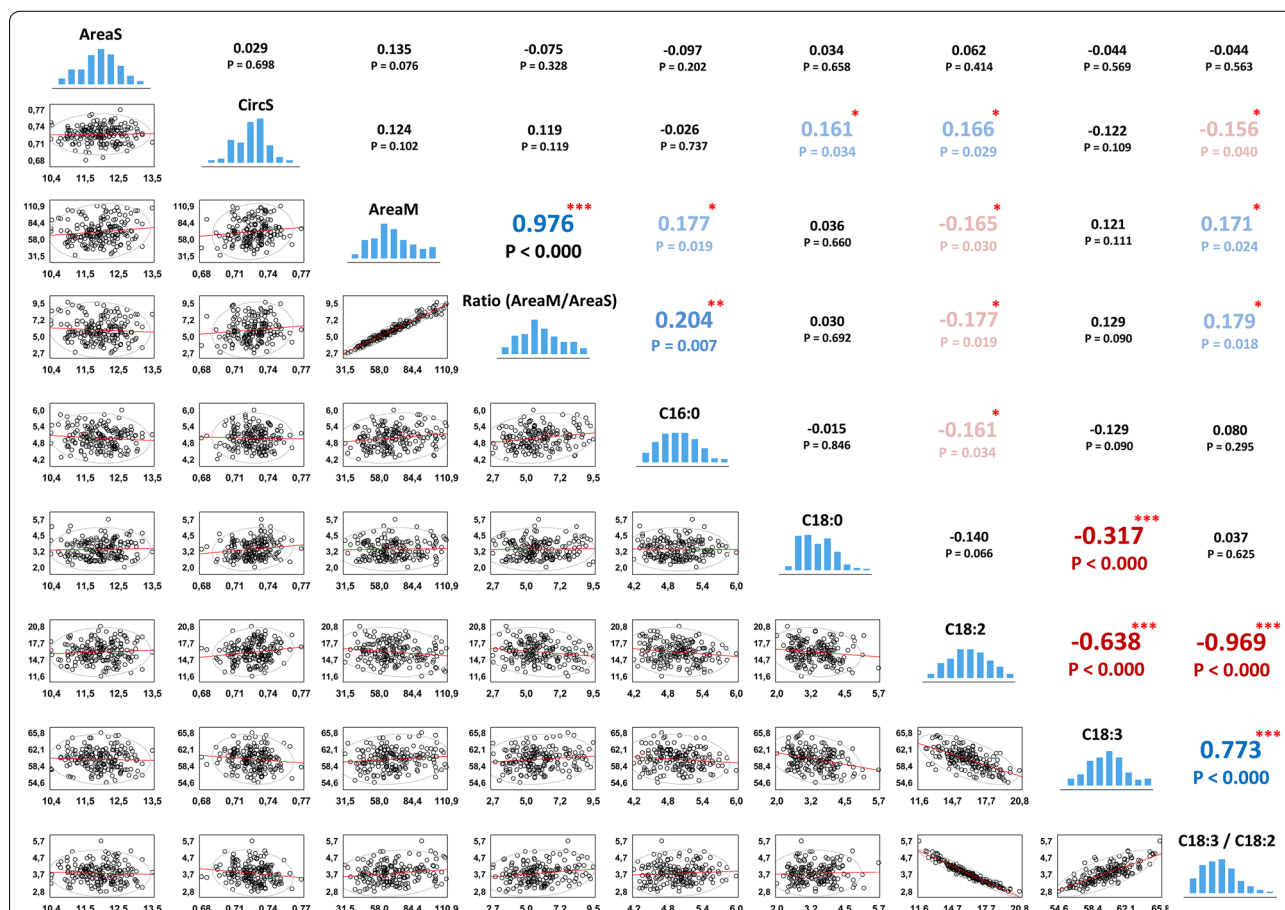


Fig. 6 Correlation matrix between traits linked to the soluble mucilage, seed shape and fatty acids. Only pairwise correlations between the most relevant parameters were displayed in the matrix. Correlations between other traits can be found in Additional file 3. The datasets used to build the matrix were obtained using MuSeeQ for the seed shape- and the soluble mucilage-related traits (10 biological replicates) and from a GC-FID analysis for the FA content (2 biological replicates), both performed on the same recombinant inbred lines population in generation F8. Below the diagonal of histograms are displayed the scatterplots of each pairwise correlation. Values above the diagonal represent the Pearson's product-moment correlation coefficients (r). The negative correlations are depicted in red while positive correlations appear in blue. The colour intensity follows the strength of the correlation. *, ** and *** asterisks indicate significant difference with $p < 0.05$, $p < 0.01$ and $p < 0.001$, respectively

C18:3 and several FAs ratios linked to C18:0 and C18:2, e.g. C18:3/C18:2, correlate with both CircS and SolS (Fig. 6; Additional files 8, 9). These results point out physiological connections between the seed morphometric parameters and the FAs content. On the other hand, the major part of the correlations found between each FA species were significantly negatives (Fig. 6; Additional files 8, 9), which confirms the results obtained on three years and two megaenvironments study of 390 flax accessions and 243 RILs [53, 97, 98].

Even it has been reported on flaxseed that the polysaccharides mucilage content does not correlate with the seed oil content [89], recent studies on *Arabidopsis* demonstrated negative correlations between them [87]. These works have been performed on the total seed oil content, without considering each FA species independently. Our pairwise correlation analyses revealed that AreaM and the ratio (AreaM/AreaS) positively correlate with the C16:0 content ($P < 0.05$ and $P < 0.01$, respectively) and the ratio (C18:3/C18:2) ($P < 0.05$), but negatively correlate with the C18:2 content ($P < 0.05$) (Fig. 6; Additional files 8, 9). Our results suggest that the relationships between the mucilage and the oil biosynthesis pathways could be more complex than as previously described [39, 87, 89]. They also suggest that the released mucilage content could be more likely linked to the FAs composition instead of the seed oil content. Finally, we cannot omit that the correlations were relatively low, which suggests a complex regulation of the mucilage biosynthesis pathway by other seed traits.

Use of MuSeeQ on various plant species

We also tested the flexibility of MuSeeQ by applying it on four others important myxospermous species (Fig. 7). The ease of use of our tool and the time required per analysis were the same for all the species, despite the great difference in the number of seeds that can be analysed by plate (Linum: ~30; Arabidopsis: ~200; Camelina: ~80; Plantago: ~60; Capsella: ~60). MuSeeQ efficiently detected and segmented mucilages and seeds (Fig. 7a). Our quantitative analysis confirmed that each species released different amounts of soluble mucilage (Fig. 7b). We also found huge differences between the species for AreaS (Fig. 7c). When AreaM was corrected by AreaS, the differences observed between species were considerably reduced (Fig. 7d). Thus, MuSeeQ can be used to screen the seeds and soluble mucilages of various species.

Conclusions

Whereas current methods used to access the soluble mucilage content are time consuming, MuSeeQ not only enables its high-throughput phenotyping for any myxospermous species, but also the simultaneous measurement of the seed shape morphometric parameters in a fast and semi-automated way.

In order to help breeders and researchers to incorporate MuSeeQ in their phenotyping strategies, genetic and functional analyses, it has been developed as a user-friendly tool. It is freely distributed to the scientific community and can be downloaded on the dedicated website (<http://MuSeeQ.free.fr>) [83].

Methods

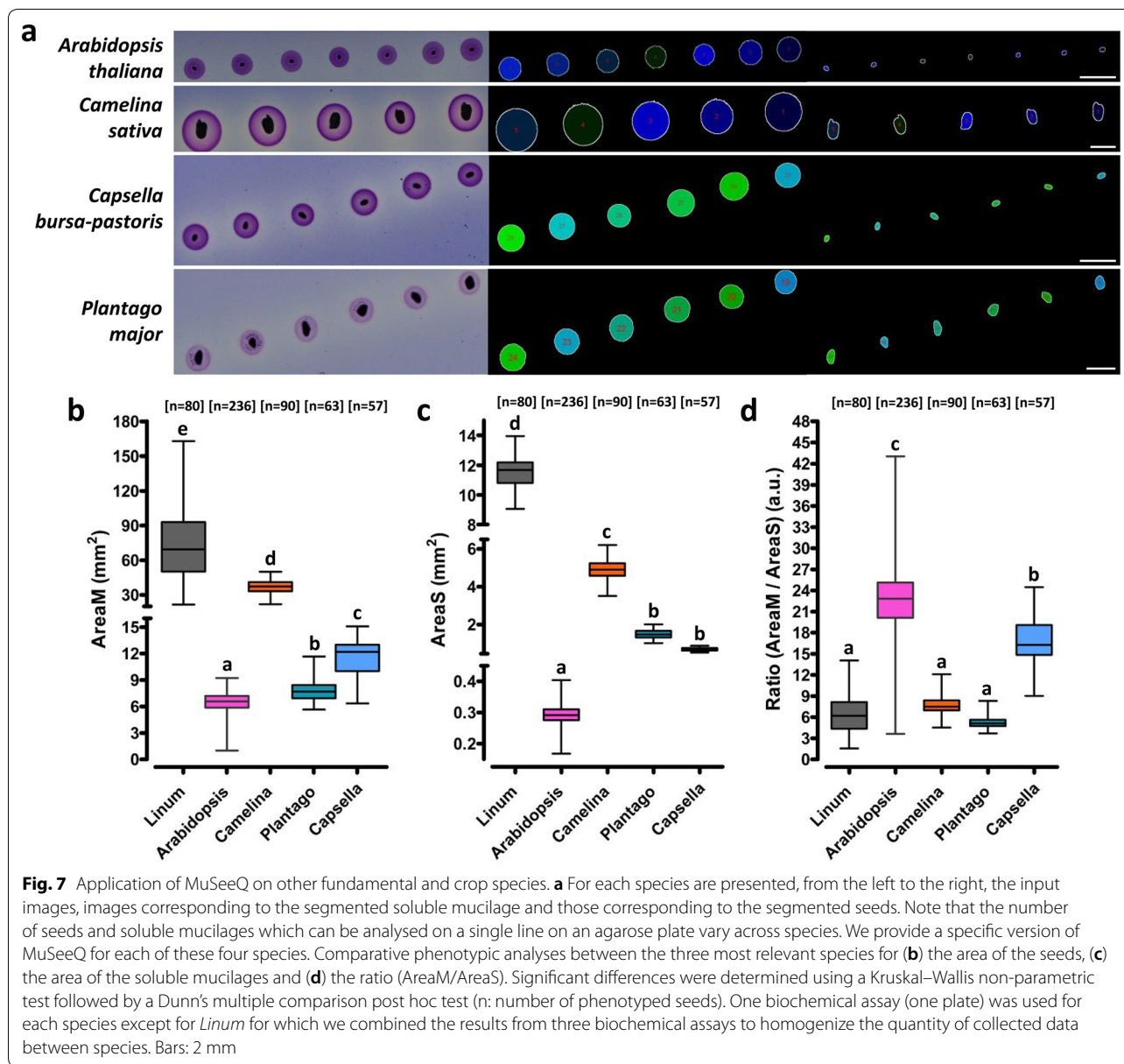
Plant material and growth conditions

Breeder seeds of *Linum usitatissimum* L. cv Oliver, a winter linseed and Viking, a spring fiber flax and 174 recombinant inbred lines (RILs) from the cross of these two varieties were kindly provided by INRA Centre de Versailles, France. Seeds were selfed in a greenhouse up to F8 under the following conditions: 60% humidity, 21 °C/15 °C day/night regime and with a 16-h photoperiod and were then multiplied in the field and used for the phenotyping of both the mucilage and FAs contents. Arabidopsis seeds (*Arabidopsis thaliana*, Columbia-0 accession) and *Camelina sativa* were from seed stocks propagated in our laboratory. *Plantago major* and *Capsella bursa-pastoris* seeds were kindly provided by Professor Michael Deyholos (The University of British Columbia IK Barber School of Arts & Sciences – Canada).

Soluble mucilage staining

The procedure required for the soluble mucilage staining on the agarose plates, i.e. the MuSeeQ's biochemical assay, is detailed in the main text, in Fig. 1 and in Additional file 10.

Validation of the technical reproducibility of MuSeeQ (Fig. 4a-c) was done using three replicates on different biochemical assays, 29 seeds of the cv Oliver for each. For the statistical validation of MuSeeQ (Fig. 4, Additional file 7), the measurements of AreaM, AreaS and the ratio (AreaM/AreaS) were performed on 21 randomly selected RILs, five seeds each. For determination of the capacities of the seeds to release the soluble and a part of the adherent mucilage [see Fig. 5a, b], flaxseeds were directly sown on the gel or after 24 h of soluble mucilage extraction in water, and MuSeeQ was applied on digital images of the gels. Analysis of the mucilage release for the RILs 84, 44 and 62 was performed by placing seeds in water containing 0.01% toluidine blue



O (Merk Millipore) without shaking. After 1 h seeds were observed using a light microscope (Axioplan 2, Zeiss, <http://www.zeiss.fr/>). For the soluble mucilages phenotyping of the RIL population (Fig. 6, Additional files 8, 9), we analysed five seeds per RIL and six RILs per assay. We performed two technical replicates and the results were combined together. Finally, the dataset used in Fig. 7 is from experiments performed with MuSeeQ on one biochemical assay per species, except for flaxseeds where we combined the analyses of three different biochemical assays to increase the size of the dataset (Additional files 11, 12).

FA phenotyping

Analysis of the FA composition was performed on mature dry seeds from all the RILs in F8. Seed oil extraction was performed on 100 mg of fresh dried seeds. Seeds were ground in liquid nitrogen. One mL of diethyl ether was added to the samples, mixed at 650 rpm for 5 min and centrifuged at 8000 rpm for 2 min. The supernatant was collected and ether evaporated. For FA composition analysis, 5 µl of n-hexane (Fisher) containing 0.005% (w/v) pentadecane were added to the pellet. Fifty µl of TMAH (TétraMéthyl Ammonium Hydroxide) were added to the samples. Samples were incubated at 20 °C at 990 rpm for

10 min and were centrifuged at 12,000 rpm at 25 °C for 10 min. One hundred fifty µl of supernatant were transferred in a glass vial. Two biological replicats were considered for each RIL of the population.

FA methyl esters analysis was performed on a TRACE GC ULTRA system coupled with a DSQ II quadrupole mass spectrometer (Thermo Scientific, France), equipped with an automated sample injector. For each sample, 1 µl was injected. The injection was performed with a 25:1 split ratio at 240 °C. For this step, the ion source was adjusted to 220 °C and the transfer line to 280 °C. For analysis, we selected the electron-impact ionization method (EI, 70 eV). The carrier gas used was Helium, with a flow rate of 1 mL min⁻¹. Separation of FA methyl esters was performed using a polar column (length: 60 m, inner diameter: 0.25 mm, 0.25 µm liquid membrane thickness, TR-FAME, Thermo Scientific, France) at 150 °C for 1 min. A gradient of 10 °C min⁻¹ was used to increase the temperature to 180 °C and the temperature maintained at 180 °C for 30 s, ramped at 1.5 °C min⁻¹ to 220 °C, 30 s at 220 °C, followed by 30 °C min⁻¹ to reach 250 °C and maintained at 250 °C for 5 min (run time: 38 min). Each mass spectrum was recorded with a scanning range of 50 to 800 m/z. X Calibur software (Thermo Scientific) was used for manual peak integration and analysis. Each FA species was identified according to its retention time by comparing it with those of pure standards (Fame mix C14-C22 Supelco-18917-1AMP). FA and standards were analysed under the same conditions. Pentadecane was applied as the internal standard. FA species concentrations for each sample were normalized against the internal control and expressed as a pourcentage of the total FA content.

Statistical analyses and data evaluation

In order to validate flaxseeds segmentation using MuSeeQ, we used SmartGrain software [45]. Linear regression analyses, scatter dot plots, Box and Whiskers plots and the Bland–Altman plot were carried out using GraphPad Prism (Version 5.0, GraphPad Software, Inc.). For linear regression analyses, the standard deviation of the residuals was computed to estimate the goodness-of-fit as:

$$S_{y,x} = \sqrt{\sum (residual^2) / (n - K)}$$

Frequency histograms were made using Matlab software (2014b, The Mathworks Inc, Natick, Massachusetts, USA). After having checked that the variables were normally distributed, Pearson's correlation analyses were performed using Statistica software v.9.1 (StatSoft, Inc., 1984–2010). The Pearson's correlation matrix and the corresponding likelihood matrix were created using Statistica (Additional files 8, 9) or after importation of the data into Matlab software (Fig. 6).

For all statistical analyses, the normality and the homocedasticity of the datasets were checked using D'Agostino & Pearson omnibus normality and Bartlett's tests, respectively. Statistical differences between samples were tested using either a one-way ANOVA followed by a Tukey's HSD (honest significant difference) post hoc test (Fig. 4, 5) or a Kruskal–Wallis non-parametric test followed by a Dunn's multiple comparison post hoc test (Fig. 7), with **p* < 0.05; ***p* < 0.01; ****p* < 0.001.

Additional files

Additional file 1. Demo video showing how to install the different versions of MuSeeQ in Fiji.

Additional file 2. Demo video showing how to use MuSeeQ to analyse your datasets.

Additional file 3. Principle of the thresholding method computed by MuSeeQ. The method is used here to dynamically detect peaks in all colour components of the HSV colour space. Spectrum areas highlighted in yellow correspond to the pixels describing as precisely as possible the soluble mucilages (left panel) and the seeds (right panel). **a** Only hue and brightness colour components are important to detect soluble mucilages. Pixels from 24 to 255 on the saturation channel were selected for minimizing the remaining background noise. **b** The brightness (value) colour component alone is important for the segmentation of the seeds.

Additional file 4. Demo video showing how to modify the script to use MuSeeQ on other species.

Additional file 5. Supports that can be used to help for the seeds sowing on the biochemical assays for *Linum usitatissimum* L.

Additional file 6. Supports that can be used to help for the seeds sowing on the biochemical assays for *Camelina sativa*.

Additional file 7. Supports that can be used to help for the seeds sowing on the biochemical assays for *Arabidopsis thaliana*.

Additional file 8. The script of the MuSeeQ macro.

Additional file 9. Correlation analysis of automated (MuSeeQ) versus manual (Fiji) measurements for **(a)** the area of the surface of soluble mucilage released in 2D (AreaM) and **(b)** the area of the seeds surface projected in 2D (AreaS). Since data are sampled from Gaussian populations (D'Agostino & Pearson omnibus normality test), Pearson's correlation coefficients were determined by linear regression analyses on *n* = 104 seeds from twenty RILs and nine agarose stained gels.

Additional file 10. Pearson's correlation matrix between the soluble mucilage, seed shape parameters and FAs-related traits. The negative correlations are depict in red whereas positive correlations are in blue. The colour intensity follows the strength of the correlation.

Additional file 11. Likelihood matrix corresponding to the Pearson's correlation matrix between the soluble mucilage, seed shape parameters and fatty acids-related traits. The negative correlations are depict in red whereas positive correlations are in blue. The colour intensity follows the strength of the correlation.

Additional file 12. Protocol for the preparation of the biochemical assay.

Abbreviations

C16:0: palmitic acid; C18:0: stearic acid; C18:1: oleic acid; C18:2: linoleic acid; C18:3: α-linolenic acid; EDTA: ethylenediaminetetraacetic acid; HCL: hydrochloric acid; NaOH: sodium hydroxide; MIV: mucilage indicator value; RGB: red, green, blue; HSV: hue, saturation, value; EM: expectation maximization; 2D:

two dimensions; 3D: three dimensions; QTL: quantitative trait loci; TMAH: tétraméthyl ammonium hydroxide; RIL: recombinant inbred lines; FA: fatty acids; ANOVA: analysis of variance; GC-FID: gaz chromatography-flame ionization detector; RIL: Recombinant inbred lines; mm: millimeter.

Authors' contributions

FMi, KP, OVW and FMe designed and conceived the research project. CP generated the RILs population. BT performed the screen of the fatty acids content. FMi designed and carried out the experiments except for the fatty acids screening. FMi performed data analysis and interpreted the results. FMi wrote the paper. KP and FMe supervised the writing. J-XF conceived the website. All authors read and approved the final manuscript.

Author details

¹ Laboratoire de Biologie des Plantes et Innovation, EA-3900, UPJV, UFR des Sciences, 33 rue St Leu, 80039 Amiens, France. ² Centre de ressources régionales en biologie moléculaire, Bâtiment Serres-Transfert, rue Dallery, 80039 Amiens Cedex 1, France. ³ Sorbonne Universités, Génie Enzymatique et Cellulaire, UMR CNRS 7025, Université de Technologie de Compiègne, CS 60319, 60203 Compiègne Cedex, France. ⁴ Present Address: Institut Jean-Pierre Bourgin, UMR1318, INRA/AgroParisTech, Saclay Plant Sciences, INRA Centre de Versailles, 78026 Versailles Cedex, France.

Acknowledgements

The authors thank Samantha Vernhettes (INRA Versailles-Grignon, France) for helpful advices and discussions. We thank Jamila Henchi for assistance in soluble mucilage extraction and quantification and all the staff of the BIOPI laboratory (University of Picardie Jules Verne) for their help in harvesting flax seeds in the greenhouse.

Competing interests

The authors declare that they have no competing interests.

Availability of data and materials

The datasets used and the MuSeeQ source codes and video tutorials are available at [78].

Consent for publication

Not applicable.

Ethics approval and consent to participate

Not applicable.

Funding

This work was performed, in partnership with the SAS PIVERT, within the frame of the French Institute for the Energy Transition (Institut pour la Transition Énergétique (ITE) P.I.V.E.R.T. (www.institut-pivert.com) selected as an Investment for the Future ("Investissements d'Avenir"). This work was supported, as part of the Investments for the Future, by the French Government under the reference ANR-001. FMi, KP, and FMe wish to thank COST action FA1006 Plant Metabolic Engineering for High Value Products (PlantEngine). The equipment utilized in this study were co-funded by the Regional Council of Picardie And European Union within CPER 2007–2020.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 12 April 2018 Accepted: 30 November 2018

Published online: 18 December 2018

References

- Gebruder B, Berlin Grubert M. Studies on the distribution of myxospermy among seeds and fruits of Angiospermae and its ecological importance. *Acta Biol Venez.* 1974;8:315–551.
- Frey-Wyssling A. The plant cell wall. In: Zimmermann W, Ozenda P, Carlquist S, Wulff HD, editors. *Handbuch der Pflanzenanatomie*, 3rd ed. Berlin: Gebrüger Borntraeger; 1976. p. 243–260.
- Fahn A. *Plant anatomy*. 3rd ed. New York: Pergamon Press; 1982.
- Boesewinkel FD, Bouman F. The seed: structure. In: Johri BM, editor. *Embryology of angiosperms*. New York: Springer; 1984. p. 567–610.
- Western TL, Skinner DJ, Haughn GW. Differentiation of mucilage secretory cells of the *Arabidopsis* seed coat. *Plant Physiol.* 2000;122:345–55.
- Macquet A, Ralet MC, Kronenberger J, Marion-Poll A, North HM. In situ, chemical and macromolecular study of the composition of *Arabidopsis thaliana* seed coat mucilage. *Plant Cell Physiol.* 2007;48:984–99.
- Naran R, Chen G, Carpita NC. Novel rhamnogalacturonan I and arabinoxylan polysaccharides of flax seed mucilage. *Plant Physiol.* 2008;148:132–41.
- Yang X, Baskin 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;14:434–42.
- North HM, Berger A, Saez-Aguayo S, Ralet MC. Understanding polysaccharide production and properties using seed coat mutants: future perspectives for the exploitation of natural variants. *Ann Bot.* 2014;114:1251–63.
- Francoz E, Ranocha P, Burlat V, Dunand C. *Arabidopsis* seed mucilage secretory cells: regulation and dynamics. *Trends Plant Sci.* 2015;20(8):515–24.
- Voiciniuc C, Yang B, Schmidt M, Günl M, Usadel B. Starting to gel: how *Arabidopsis* seed coat epidermal cells produce specialized secondary cell walls. *Int J Mol Sci.* 2015;16:3452–73.
- Arsovski AA, Popma TM, Haughn GW, Carpita NC, McCann MC, Western TL. AtBXL1 encodes a bifunctional b-D-xylidase/a-L-arabinofuranosidase required for pectic arabinan modification in *Arabidopsis* mucilage secretory cells. *Plant Physiol.* 2009;150:1219–34.
- Stork J, Harris D, Griffiths J, Williams B, Beisson F, Li-Beisson Y, et al. CELLULOSE SYNTHASE9 serves a nonredundant role in secondary cell wall synthesis in *Arabidopsis* epidermal testa cells. *Plant Physiol.* 2010;153:580–9.
- Mendu V, Griffiths J, Persson S, Stork J, Downie B, Voiciniuc C, et al. Subfunctionalization of cellulose synthases in seed coat epidermal cells mediate secondary radial wall synthesis and mucilage attachment. *Plant Physiol.* 2011;157:441–53.
- Western TL. The sticky tale of seed coat mucilages: production, genetics, and role in seed germination and dispersal. *Seed Sci Res.* 2012;22:1–25.
- Macquet A, Ralet MC, Loudet O, Kronenberger J, Mouille G, Marion-Poll A, et al. A naturally occurring mutation in an *Arabidopsis* accession affects a b-D-galactosidase that increases the hydrophilic potential of rhamnogalacturonan I in seed mucilage. *Plant Cell.* 2007;19:3990–4006.
- Sullivan S, Ralet MC, Berger A, Diatloff E, Bischoff V, Gonneau M, et al. CESAS5 is required for the synthesis of cellulose with a role in structuring the adherent mucilage of *Arabidopsis* seeds. *Plant Physiol.* 2011;156:1725–39.
- Paynel F, Pavlov A, Ancelin G, et al. Polysaccharide hydrolases are released with mucilages after water hydration of flax seeds. *Plant Physiol Biochem.* 2013;62:54–62.
- Bailey K. Cress seed mucilage. *Biochem J.* 1935;29:2477–85.
- Anderson E, Lowe HJ. The composition of flaxseed mucilage. *J Biol Chem.* 1947;168:289–97.
- Sterling C. Crystal-structure of ruthenium red and stereochemistry of its pectin stain. *Am J Bot.* 1970;57:172–5.
- Hanke DE, Northcote DH. Molecular visualization of pectin and DNA by ruthenium red. *Biopolymers.* 1975;14:1–17.
- Penfield S, Meissner RC, Shoue DA, Carpita NC, Bevan MW. MYB61 is required for mucilage deposition and extrusion in the *Arabidopsis* seed coat. *Plant Cell.* 2001;13:2777–91.
- Willats WGT, McCartney L, Mackie W, Know JP. Pectin: cell biology and prospects for functional analysis. *Plant Mol Biol.* 2001;47:9–27.
- Bergeron JA, Singer M. Metachromasy: an experimental and theoretical reevaluation. *J Biophys Biochem Cytol.* 1958;4:433–57.
- O'Brien TP, Feder N, McCully ME. Polychromatic staining of plant cell walls by toluidine blue O. *Protoplasma.* 1964;59:367–73.
- Beeckman T, De Rycke R, Viane R, Inzé D. Histological study of seed coat development in *Arabidopsis thaliana*. *J Plant Res.* 2000;113:139–48.
- Mazza G, Biliaderis CG. Functional properties of flax seed mucilage. *J Food Sci.* 1989;54:1302–5.
- Fedeniuk RW, Biliaderis CG. Composition and physicochemical properties of linseed (*Linum usitatissimum* L.) mucilage. *J Agric Food Chem.* 1994;42:240–7.

30. Usadel B, Kuschnisky AM, Rosso MG, Eckermann N, Pauly M. RHM2 is involved in mucilage pectin synthesis and is required for the development of the seed coat in Arabidopsis. *Plant Physiol.* 2004;134:286–95.
31. 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;54:466–80.
32. Huang J, Bowles D, Efsandiari E, Dean G, Carpita NC, Haughn GW. The Arabidopsis transcription factor LUH/MUM1 is required for extrusion of seed coat mucilage. *Plant Physiol.* 2011;156:491–502.
33. Walker M, Tehseen M, Doblin MS, Pettolino FA, Wilson SM, Bacic A, Golz JF. The transcriptional regulator LEUNIG_HOMOLOG regulates mucilage release from the Arabidopsis testa. *Plant Physiol.* 2011;156:46–60.
34. Ziolkovska A. Laws of flaxseed mucilage extraction. *Food Hydrocoll.* 2012;26:197–204.
35. Voiniciuc C, Dean GH, Griffiths JS, et al. FLYINGSAUCER1 is a transmembrane RING E3 ubiquitin ligase that regulates the degree of pectin methylesterification in Arabidopsis seed mucilage. *Plant Cell.* 2013;25:944–59.
36. Cui W, Mazza G, Oomah BD, Biliaderis CG. Optimization of an aqueous extraction process for flaxseed gum by response surface methodology. *Lebensmittel-Wissenschaft und Technologie.* 1994;27:363–9.
37. Kadivar M. Studies on integrated processes for the recovery of mucilage, hull, oil and protein from solin (low linolenic acid flax). Doctoral dissertation. Available from University of Saskatchewan Library Electronic Theses & Dissertations (URN etd-10212004-002848); 2001.
38. Bhatti RS. Further compositional analysis of flax: mucilage, trypsin inhibitors and hydrocyanic acid. *J Am Oil Chem Soc.* 1993;70:899–904.
39. Diederichsen A, Raney JP, Duguid SD. Variation of mucilage in flax seed and its relationship with other seed characters. *Crop Sci.* 2006;46:365.
40. Peng H. Bioimage informatics: a new area of engineering biology. *Bioinformatics.* 2008;24:1827–36.
41. Shamir L, Delaney JD, Orlov N, Eckley DM, Goldberg IG. Pattern recognition software and techniques for biological image analysis. *PLoS Comput Biol.* 2010;6:e1000974.
42. Russ JC. The image processing handbook. 6th ed. Boca Raton: CRC Press; 2011.
43. Dana W, Ivo W. Computer image analysis of seed shape and seed colour for flax cultivar description. *Comput Electr Agric.* 2008;61:126–35.
44. Herridge RP, Day RC, Baldwin S, Macknight RC. Rapid analysis of seed size in Arabidopsis for mutant and QTL discovery. *Plant Methods.* 2011;7:3.
45. Tanabata T, Shibaya T, Hori K, Ebana K, Yano M. SmartGrain: high-throughput phenotyping software for measuring seed shape through image analysis. *Plant Physiol.* 2012;160:1871–80.
46. Iva S, Oszcz G, Marie B, Gianfranco V. Phenotypic evaluation of flax seeds by image analysis. *Ind Crops Prod.* 2013;47:232–6.
47. Moore CR, Johnson LS, Kwak IY, Livny M, Broman KW, Spalding EP. High-throughput computer vision introduces the time axis to a quantitative trait map of a plant growth response. *Genetics.* 2013;195:1077–86.
48. Voiniciuc C, Heinrich-Wilhelm M, Berger A, Yang B, Ebert B, Scheller HV, North HM, Usadel B, Günl M. MUCILAGE-RELATED10 produces galactoglucomannan that maintains pectin and cellulose architecture in Arabidopsis seed mucilage. *Plant Physiol.* 2015;169:403–20.
49. Verbeken D, Dierckx S, Dewettinck K. Exudates gums: occurrence, productions and applications. *Appl Microbiol Biotechnol.* 2003;63(1):10–21.
50. Hall C, Tulbek MC, Xu Y. Flaxseed. *Adv Food Nutr Res.* 2006;51:1–97.
51. Touré A, Xueming X. Flaxseed lignans: source, biosynthesis, metabolism, antioxidant activity, bio-active components, and health benefits. *Compr Rev Food Sci Food Saf.* 2010;9:261–9.
52. Herchi W, Arraez-Roman D, Trabelsi H, Bouali I, Boukhchina S, Kallel H, et al. Phenolic compounds in flaxseed: a review of their properties and analytical methods. An overview of the last decade. *J Oleo Sci.* 2014;63(1):7–14.
53. Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S. Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. *Theor Appl Genet.* 2014;127:881–96.
54. French A, Ubeda-Tomás S, Holman TJ, Bennett MJ, Pridmore T. High-throughput quantification of root growth using a novel image analysis tool. *Plant Physiol.* 2009;150:1784–95.
55. Naeem A, French AP, Wells DM, Pridmore TP. High-throughput feature counting and measurement of roots. *Bioinformatics.* 2011;27:1337–8.
56. US National Institutes of Health; <http://rsbweb.nih.gov/ij/>. Accessed 14 July 2018.
57. Schindelin J, et al. Fiji: an open-source platform for biological-image analysis. *Nat Methods.* 2012;9:676–82.
58. Hemming J. Computer vision for identifying weeds in crops. *Gartenbautechnische Informationen, Heft 50.* Institut für Technik in Gartenbau und Landwirtschaft, Universität Hannover. 2000.
59. Ford A, Roberts A. Colour space conversions. London: Westminster University. 1998. p. 1–31.
60. Agoston MK. Computer graphics and geometric modeling: implementation and algorithms. London: Springer; 2005.
61. Walter A, Scharr H, Gilmer F, Zierer R, Nagel KA, Ernst M, et al. Dynamics of seedling growth acclimation towards altered light conditions can be quantified via GROWSCREEN: a setup and procedure designed for rapid optical phenotyping of different plant species. *New Phytol.* 2007;174:447–55.
62. Lin TT, Lai TM, Chen S, Fon DS. Gray-scale and colour machine vision systems for seedling detection. *Computers in agriculture.* In: Proceedings of the 5th international conference, Orlando, Florida. 1994; pp 105–110.
63. Philipp I, Rath T. Improving plant discrimination in image processing by use of different colour space transformations. *Comput Electron Agric.* 2002;35:1–15.
64. De Vylder J, Vandenbussche F, Hu Y, Philips W, Van Der Straeten D. Rosette tracker: an open source image analysis tool for automatic quantification of genotype effects. *Plant Physiol.* 2012;160:1149–59.
65. Haralick RM, Shapiro LG. Image segmentation techniques. *Comput Vis Graph Image Process.* 1985;29:100–32.
66. Jaffe MJ, Wakefield AH, Telewski F, Gulley E, Biro R. Computer assisted image analysis of plant growth, thigmomorphogenesis and gravitropism. *Plant Physiol.* 1985;77:722–30.
67. Leister D, Varotto C, Pesaresi P, Niwergall A, Salamini F. Largescale evaluation of plant growth in *Arabidopsis thaliana* by noninvasive image analysis. *Plant Physiol Biochem.* 1999;37:671–8.
68. Arvidsson S, Pérez-Rodríguez P, Mueller-Roeber B. A growth phenotyping pipeline for *Arabidopsis thaliana* integrating image analysis and rosette area modeling for robust quantification of genotype effects. *New Phytol.* 2011;191:895–907.
69. Pound MP, French AP, Atkinson JA, Wells DM, Bennett MJ, Pridmore T. RootNav: navigating images of complex root architectures. *Plant Physiol.* 2013;162:1802–14.
70. Pridmore T, French A, Pound M. What lies beneath: underlying assumptions in bioimage analysis. *Trends Plant Sci.* 2012;17:688–92.
71. Shimizu H, Heins RD. Computer-vision-based system for plant growth analysis. *Trans ASABE.* 1995;38:959–64.
72. Clément A, Vigouroux B. Unsupervised segmentation of scenes containing vegetation (*Forsythia*) and soil by hierarchical analysis of bidimensional histograms. *Pattern Recognit Lett.* 2003;24:1951–7.
73. Klukas C, Chen D, Pape JM. Integrated analysis platform: an open-source information system for high-throughput plant phenotyping. *Plant Physiol.* 2014;165:506–18.
74. Chan W, Jaitly N, Le Quoc V, Vinyals O. Attend and spell: a neural network for large vocabulary conversational speech recognition. In: IEEE international conference on acoustics, speech and signal processing, ICASSP 2016, Shanghai, China March 20–25, 2016; pp 4960–4964.
75. Pound MP, Atkinson JA, Townsend AJ, Wilson MH, Griffiths M, Jackson AS, Bulat A, Tzimiroopoulos G, Wells DM, Murchie EH, Pridmore TP, French AP. Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. *Gigascience.* 2017;6:1–10.
76. Zhou J, Troyanskaya OG. Predicting effects of noncoding variants with deep learning-based sequence model. *Nat Methods.* 2015;12:931–4.
77. Esteva A, Kuprel B, Novoa RA, et al. Dermatologist-level classification of skin cancer with deep neural networks. *Nature.* 2017;542:115–8.
78. Girshick R, Donahue J, Darrell T, Malik J. Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR). 2013; pp 580–587.
79. Ren S, He K, Girshick R, Sun J; Faster r-cnn: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems. 2015; pp 91–99.

80. Dai J, Li TY, He K, Sun J et al. R-FCN: Object detection via region-based fully convolutional networks. In: *Advances in neural information processing systems*. 2016; pp 379–387.
81. Golhani K, Balusundram SK, Vadamalai G, Pradhan B. A review of neural networks in plant disease detection using hyperspectral data. *Inf Process Agric*. 2018;5:354–71.
82. Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with Deep Convolutional Neural Networks. In: *Advances in neural information processing systems*. 2012; pp 1097–1105.
83. <http://MuSeeQ.free.fr>. Accessed 14 July 2018.
84. Pavlov A, Paynel F, Rihouey C, Porokhvinova E, Brutch N, Morvan C. Variability of seed traits and properties of soluble mucilages in lines of the flax genetic collection of Vavilov Institute. *Plant Physiol Biochem*. 2014;80:348–61.
85. Abràmoff MD, Magalhães PJ, Ram SJ. Image processing with ImageJ. *Biophotonics Int*. 2004;11(36):42.
86. Schneider CA, Rasband WS, Eliceiri KW. NIH Image to ImageJ: 25 years of image analysis. *Nat Methods*. 2012;9:671–5.
87. Shi L, Katavic V, Yu Y, Kunst L, Haughn G. *Arabidopsis glabra2* mutant seeds deficient in mucilage biosynthesis produce more oil. *Plant J*. 2012;69:37–46.
88. Wang Z, Chen T, Xuan L, Li Z, Du X, Zhou L, Zhang G, Jiang L. TRANSPARENT TESTA2 regulates embryonic fatty acid biosynthesis by targeting FUSCA3 during the early developmental stage of *Arabidopsis* seeds. *Plant J*. 2014;77:757–69.
89. Oomah BD, Kenaschuk EO, Cui W, Mazza G. Variation in the composition of water-soluble polysaccharides in flaxseed. *J Agri Food Chem*. 1995;43:1484–8.
90. Eaton SB, Eaton SB III, Sinclair AJ, Cordain L, Mann NJ. Dietary intake of longchain polyunsaturated fatty acids during the Paleolithic. In: Simopoulos AP, editor. *The return of ω 3 fatty acids into the food supply*. I. Land-based animal food products and their health effects, vol. 83. Basel: World Rev Nutr Diet; 1998. p. 12–23.
91. Simopoulos AP. The importance of the ratio of omega-6/omega-3 essential fatty acids. *Biomed Pharmacother*. 2002;56:365–79.
92. Bowman JL, Koornneef M. Seed morphology. In: Bowman JL, editor. *Arabidopsis: an atlas of morphology and development*. New York: Springer; 1994. p. 398–401.
93. Debeaujon I, Léon-Kloosterziel KM, Koornneef M. Influence of the testa on seed dormancy, germination, and longevity in *Arabidopsis*. *Plant Physiol*. 2000;122:403–14.
94. Western TL, Burn J, Tan WL, Skinner DJ, Martin-McCaffrey L, Moffatt BA, Haughn GW. Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in *Arabidopsis*. *Plant Physiol*. 2001;127:998–1011.
95. Kunieda T, Mitsuda N, Ohme-Takagi M, et al. NAC family proteins NARS1/NAC2 and NARS2/NAM1 the outer integument regulate embryogenesis in *Arabidopsis*. *Plant Cell*. 2008;20:2631–42.
96. Molina I, Ohlrogge JB, Pollard M. Deposition and localization of lipid polyester in developing seeds of *Brassica napus* and *Arabidopsis thaliana*. *Plant J*. 2008;53:437–49.
97. Cloutier S, Ragupathy R, Niu Z, Duguid S. SSR-based link-age map of flax (*Linum usitatissimum* L.) and mapping of QTLs underlying fatty acid composition traits. *Mol Breed*. 2011;28:437–51.
98. Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S. QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). *Theor Appl Genet*. 2015;128(5):965–84.

Ready to submit your research? Choose BMC and benefit from:

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

At BMC, research is always in progress.

Learn more biomedcentral.com/submissions

