

Article

Critical Evaluation of the Cgrain Value™ as a Tool for Rapid Morphometric Phenotyping of Husked Oat (*Avena sativa* L.) Grains

David Evershed , Eamon J. Durkan, Rachel Hasler, Fiona Corke , John H. Doonan 
and Catherine J. Howarth 

National Plant Phenomics Centre, Institute of Biological, Environmental & Rural Sciences (IBERS),
Aberystwyth University, Aberystwyth SY23 3EE, UK; eamondurkan@gmail.com (E.J.D.); fic5@aber.ac.uk (F.C.);
jhd2@aber.ac.uk (J.H.D.)

* Correspondence: dae51@aber.ac.uk (D.E.); cnh@aber.ac.uk (C.J.H.)

Abstract: Mechanised non-contact, non-destructive imaging methodologies have revolutionised plant phenotyping, increasing throughput well beyond what was possible using traditional manual methods. Quantifying the variation in post-harvest material such as seeds and fruits, usually the economically important part of the crop, can be critical for commercial quality assessment as well as breeding and research. Therefore, reliable methods that gather metrics of interest, quickly and efficiently, are of widespread interest across sectors. This study focuses on evaluating the phenotyping capabilities of the Cgrain Value™, a novel grain imaging machine designed for quality and purity assessment and used primarily in commercial cereal production and processing. The performance of the Cgrain Value™ in its generation of high-throughput quantitative phenotypic data is compared with a well-established machine, MARVIN, assessing repeatability and reproducibility across a range of metrics. The findings highlight the potential of the Cgrain Value™, and some shortcomings, to provide detailed three-dimensional size, shape, and colour information rapidly, offering insights into oat grain morphology that could enhance genome-wide association studies and inform the breeding efforts in oat improvement programmes.

Keywords: oats; *Avena sativa*; seed phenotyping; morphometric analysis; seed geometry



Citation: Evershed, D.; Durkan, E.J.; Hasler, R.; Corke, F.; Doonan, J.H.; Howarth, C.J. Critical Evaluation of the Cgrain Value™ as a Tool for Rapid Morphometric Phenotyping of Husked Oat (*Avena sativa* L.) Grains. *Seeds* **2024**, *3*, 436–455. <https://doi.org/10.3390/seeds3030030>

Received: 4 July 2024

Revised: 14 August 2024

Accepted: 15 August 2024

Published: 22 August 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

High-throughput, non-destructive precision technologies to assess the variation present in breeding populations and diversity collections are essential to accelerate the development of new varieties that address emerging challenges such as disease and climate change as well as to identify the underlying genes controlling traits. Linking the variation in phenotypic traits with explanatory genetic and environmental factors typically requires large datasets. The construction of such datasets is costly, particularly in terms of labour. Mechanised and other systematic approaches can reduce the cost per data point. Thus, high-throughput methodologies for the collection of plant phenotypic data can improve the quality and efficacy of genome association studies over a range of species and measured characters [1].

The measurement of cereal grain size and shape traits via image-based methodologies have been found to be effective in successful genome-wide association mapping. For example, data collected by the MARVIN Seed analyzer (GTA Sensorik GmbH) (hereafter “MARVIN”) have been used to find QTLs associated with the variation in grain geometry in wheat [2–4], barley [5,6], and triticale [7,8]. MARVIN has also been used to determine the effect of the environment on grain size parameters [9]. Other more bespoke imaging equipment have been used to reveal the genetic control of grain size and shape traits in sorghum [10], using the SeedCount SC5000 (Next Instruments, Condell Park, NSW,

Australia) and the Crop Grain Appearance Quality Scanning Machine (SC-E, Wanshen Technology Company, Hangzhou, China) in wheat [11]. Open-source software packages such as “GrainScan” [12] and “SmartGrain” [13], developed for use with “off the shelf” imaging equipment have also been used in successful QTL mapping in, for example, durum wheat [14] and rice [15]. These examples by no means constitute an exhaustive list, but the predominant method for high-throughput cereal grain size and shape data collection at time of writing utilises MARVIN. The data collected by MARVIN in these studies is predominantly two-dimensional (2-D), and in many cases, also lacks colour information. While MARVIN analysis has been very successful, the data provided limit a more nuanced understanding of the genetic origins of shape. Some mid-throughput equipment and methodologies have been implemented to capture three-dimensional (3-D) size and shape data of seeds, such as the “phenoSeeder” [16], which is based on a pick-and-place robot with high precision scales and 3-D imaging stations, uCT scanning [17–19], and point cloud analysis [20]. However, the relatively low throughput of these methodologies render them less suitable for genetic analysis.

The grain colour of oats is industrially significant, but difficult to measure accurately and consistently. The current systems are based on human eye perception and classifying into one group of named colours including “white”, “yellow”, “red”, “tan”, “grey”, “black”, and various combinations thereof [21,22]. Oat husk colour varies by genotype and despite being largely unrelated to the appearance and quality of the valuable groat within, there are industrial market preferences for certain colours, particularly for “white” or “bright” oats for both human and animal consumption [23–25]. Oat grain colour is also commonly used in industry to detect problems—e.g., immature grains, mould, insects, or drying damage [26–28].

The Cgrain Value™ (Cgrain AB, Uppsala, Sweden) (hereafter “Cgrain”), principally developed for the rapid assessment of cereal quality and purity in end-use industrial applications such as mills and maltings, provides 3-D metrics [29,30]. By means of a right-angled, mirrored imaging surface (Figure 1) and AI-based modelling by an integrated computer, the machine is capable of predicting the 3-D geometry, weight, and colour information from a single image-capture step at a rate of several grains per second. The machine therefore has the potential to offer an enhanced dataset at little or no cost in terms of sample processing time when compared to other mainstream grain phenotyping equipment.

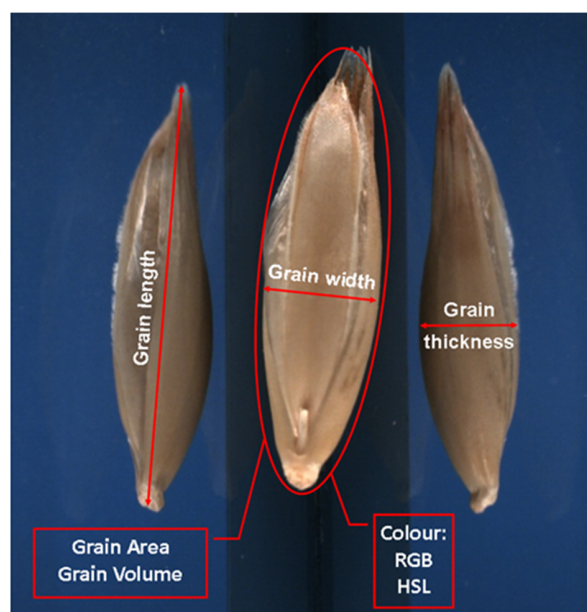


Figure 1. Example of an image of an oat grain as captured by the Cgrain Value™. The angled mirror allows for almost complete visualisation of the grain surface, from which predictions of 3-D structure and colour can be made.

However, the Cgrain system has not previously been compared to other seed measuring equipment in a research setting. Here, we report a benchmarking study of Cgrain against MARVIN for the traits that both devices measure, using a large mapping population of oats that is currently under development within the breeding programme in Aberystwyth. Additionally, we report the high-throughput collection of predicted 3-D grain size and shape data as well as grain colour. We compare the data on a sample-by-sample basis (or plant-by-plant in the context of this study) to that collected by MARVIN, and the repeatability and reproducibility of the Cgrain results are also reported. Table 1 compares the relevant capabilities of the machines used in the test.

Table 1. Description of the relevant measurement capabilities of the machines under investigation.

Measurement	Units	Definition	Machine MARVIN	Capability Cgrain
Whole Sample				
Grain count	n/a	Total number of grains detected in the sample.	Y	Y
TGW	g	Estimation of the total weight of 1000 representative grains of the sample.	Y	Y
Individual Grains				
Grain weight	mg	Estimation of the weight of the grain.	N	Y
Grain length	mm	Longest continuous linear distance between two points on the outermost edge of the grain.	Y	Y
Grain width	mm	Longest continuous linear distance between two points on the outermost edge of the grain, perpendicular to the length measurement.	Y	Y
Grain area	mm ²	Area of background obscured by a single image of the grain.	Y	Y
Grain thickness	mm	The shortest maximum width of the grain in any orientation about the length axis, i.e., the smallest possible width slot (of infinite length) through which the grain can fully pass.	N	Y
Grain volume	mm ³	Total apparent volume in space occupied by the grain, which is assumed continuously solid.	N	Y
Grain colour	RGB/HSL	Average pixel colour of segmented grain image	N	Y

2. Materials and Methods

2.1. Plant Material and Growing Conditions

Gerald and Mascani, two popular UK varieties of husked winter oats, and 138 F5 progeny from a cross between these 2 varieties were used for this study. A total of 3 replicates of 138 individual F5 plants along with the parents were grown from September 2022 to April 2023. F4 seeds were sown in 6 cm pots in Levington F2 compost and germinated for 7 days before vernalisation for 4 weeks (5 °C, 12 h day length), watering as required. Vernalised seedlings were transplanted to 3.5 L pots of the same compost and allowed to establish for 10 days before they were transferred in a randomised block design to the LemnaTec automated phenotyping platform (LemnaTec, Wuerstli, Germany) in glasshouses at IBERS, Aberystwyth University. Plants were retained on the system to maturity and harvest ripeness with 16 h daylength provided with top-up LED lighting, as required, to 450 µM/m²/s under a temperature regime of 18 °C by day and 15 °C by night. Daily automated watering maintained compost at 65% field capacity until GS91, where watering ceased. The ripe panicles were harvested by hand and threshed using a laboratory threshing machine (Hans-Ulrich Hege Saatuchtmaschinen GmbH, Hohebuch, Germany). The seed samples were cleaned of debris and chaff by hand. Care was taken to manually pull apart any remaining multiple-kernel spikelets into individual grains. Only samples containing sufficient seed for the next generation of plants (≥25 seeds) were included in the analysis. In total, 78,259 seeds from 372 plants were analysed.

2.2. Seed Counting and Thousand Grain Weight Analysis

For the determination of seed number and Thousand Grain Weight (TGW), seeds from each plant were counted using a Data Count S-25 seed counter (Data Technologies Ltd., Tsor'a, Israel) and weighed to 2 decimal places on a Precisa XB 220A balance (Precisa Gravimetrics AG, Dietikon, Switzerland). TGW was determined by the formula:

$$\text{TGW} = \text{Total Weight of Sample} \times \frac{1000}{\text{Total Number of Grains in Sample}}$$

2.3. MARVIN Image Analysis

Grain length (mm), width (mm), and area (mm²) data were collected from all seeds produced from each plant. These measurements are collected automatically for each “object” detected by the machine, the code for which is not available to the user and therefore cannot be further defined here. Data collection comprised several individual sample presentations and image capture events per plant: grains were tipped onto the imaging tray in batches of approximately 100 for image capture, and this was repeated until all seeds from a given plant were imaged. Seeds were left in their randomly adopted resting position. For geometric measurements, any multiple seed units and overlapping/touching seeds were removed by using an area threshold of 55 mm². This threshold successfully removes any multiple seed units across all genotypes while removing very few genuine single grains. TGW (an automatic output for each image but not for a whole sample) for each plant from the MARVIN analysis was made by the following formula:

$$\text{TGW} = \frac{\text{Total Weight Analysed}}{\text{Total Objects Identified}} \times 1000$$

Grain roundness was calculated for each grain by its width-to-length ratio, resulting in a number between 0 and 1, with a score of 1 being perfectly circular:

$$\text{Grain Roundness} = \frac{\text{Grain Width}}{\text{Grain Length}}$$

2.4. Cgrain Image Analysis

The Cgrain can hold and analyse one bulk grain sample at a time up to a maximum of approximately 500 g for oats. In each instance, the entire sample was tipped into the machine's sample inlet and the oat grains were automatically presented to the image capture area at a steady rate by means of the integrated vibrating bowl, which automatically adjusts the speed of the vibration to maintain a sampling rate of approximately 5–7 seeds per second. For each sample, the Cgrain was allowed to operate until all grains had passed over the imaging surface and into the collection drawer below, from which they were recovered before the next sample was analysed. The calibration used, “Cgrain Oats v.1”, is a non-optimised, generic oat calibration used for demonstration purposes and was provided by the manufacturer.

A representative estimation of TGW for the whole sample is made by the Cgrain—this result is a function of seed count and total weight but is also adjusted based on the image classifications and individual seed measurements. The final TGW result is an output of the integrated Artificial Neural Network (ANN), the code for which is not accessible to the user and is proprietary in nature, so it cannot be further defined here (this is the case for all automatically outputted Cgrain metrics). The Cgrain also estimates the weights of each individual grain by the same method. The Cgrain therefore outputs two different results for TGW, one based on the entire sample as discussed above (TGW_{ANN}), and another result equal to the mean estimated grain weight from the sample, multiplied by a thousand (TGW_{Calc}).

Grain roundness (not an automatic output of the Cgrain analysis) was calculated in the same way as for MARVIN using the “length” and “width” parameters.

Cross-sectional grain roundness was calculated for each seed by its thickness-to-width ratio, resulting in a number between 0 and 1, with a score of 1 being perfectly circular:

$$\text{Cross Sectional Grain Roundness} = \frac{\text{Grain Thickness}}{\text{Grain Width}}$$

All samples were analysed in duplicate and plant means were taken from the two sets of results for all measurements.

2.5. Statistical Analysis

Statistical analysis was carried out using R (version 4.3.2) in the R Studio environment. Arithmetic mean results per plant for each of the traits under investigation were calculated, except for colour. Mean grain colour per plant in the RGB colour space is estimated using the quadratic mean of the RGB (0–255) values as follows:

$$\text{Mean Red} = \sqrt{\frac{\sum r^2}{n}}, \text{Mean Green} = \sqrt{\frac{\sum g^2}{n}}, \text{Mean Blue} = \sqrt{\frac{\sum b^2}{n}}$$

Mean grain colour per plant in the HSL colour space is estimated using the arithmetic mean of saturation and light, and the circular mean of hue as follows (adapted from [31]):

$$\text{Mean hue} = \arctan2\left(\sum_{i=1}^n \sin[\text{hue}_i], \sum_{i=1}^n \cos[\text{hue}_i]\right)$$

Final genotype trait means for the measured traits are calculated as the means of the individual per-plant means.

For per-plant trait means that returned normally distributed data when tested with the Shapiro–Wilk test, significance of differences between genotypes was assessed by ANOVA, and a paired T-test was carried out to assess significance of differences between machines. For non-parametric data, significance of differences between genotypes was assessed by the Kruskal–Wallis test and significance of differences between machines were assessed using the Wilcoxon Signed Rank Test. Linear correlations and corresponding *r*-values for equivalent measurements were made using least squares linear regression. Bland–Altman analysis was carried out on equivalent measurements to show the relationship between results and the limits of agreement. Where the quantitative differences between measurements are non-parametric when tested by the Shapiro–Wilk test, the 95% limits of agreement are estimated by the 2.5th and 97.5th percentiles. For quantitative differences between measurements that are accepted the Shapiro–Wilk test, the 95% limits of agreement are set by the mean difference (bias), ± 1.96 standard deviations [32].

For repeatability assessment between the duplicate analyses through the Cgrain, a coefficient of variation (CV) for duplicate measurements is calculated as follows [33]:

$$\text{CV}(\%) = 100 \times \sqrt{\frac{\sum \left(\frac{d}{m}\right)^2}{2n}}$$

where *d* is the difference between duplicates, *m* is the duplicate mean, and *n* is the total number of plants.

To visualise trait associations, correlation plots are made for the final genotype means for both machines using the Pearson method. Biplots of the first and second principal components are made for both machines [34].

3. Results

First, we compare the TGW, obtained using a manual reference method common in the industry, versus both MARVIN and the two outputs for the TGW from the Cgrain.

We then compare the measurement types that are common to both MARVIN and Cgrain. Finally, we report the data captured only by Cgrain.

3.1. Thousand Grain Weight

Shapiro–Wilk tests suggest that the TGW data produced for all plants in the population by the manual method and Cgrain TGW_{ANN} are non-parametric ($W = 0.99$ and 0.99 , p -value = 0.04 and 0.05 , respectively) while the MARVIN and Cgrain TGW_{Calc} data can be treated as normal ($W = 0.99$ and 0.99 , p -value = 0.20 and 0.07 , respectively). The TGW varied significantly by genotype for the manual method and for all three predictive methodologies (Manual: Kruskal–Wallis Test, p -value < 0.0001 , MARVIN: ANOVA, p -value < 0.0001 , TGW_{ANN} : Kruskal–Wallis Test, p -value < 0.0001 , TGW_{Calc} : ANOVA, p -value < 0.0001).

All three prediction methods produced TGW results that showed a strong linear relationship to the manual data ($r > 0.89$) (Figure 2). The TGW estimation by MARVIN for each plant was significantly higher than the manual result (Wilcoxon Signed Rank Test, $p \leq 0.0001$) with a mean difference of $+1.5$ g. The TGW estimation data by calculation from the Cgrain mean estimated seed weight (TGW_{Calc}) was not significantly different from the manual result (Wilcoxon Signed Rank Test, $p = 0.22$). The ANN-generated TGW prediction by Cgrain (TGW_{ANN}) showed significantly lower TGW than the manual method (Wilcoxon Signed Rank Test, $p < 0.0001$) with a mean difference of -2.7 g.

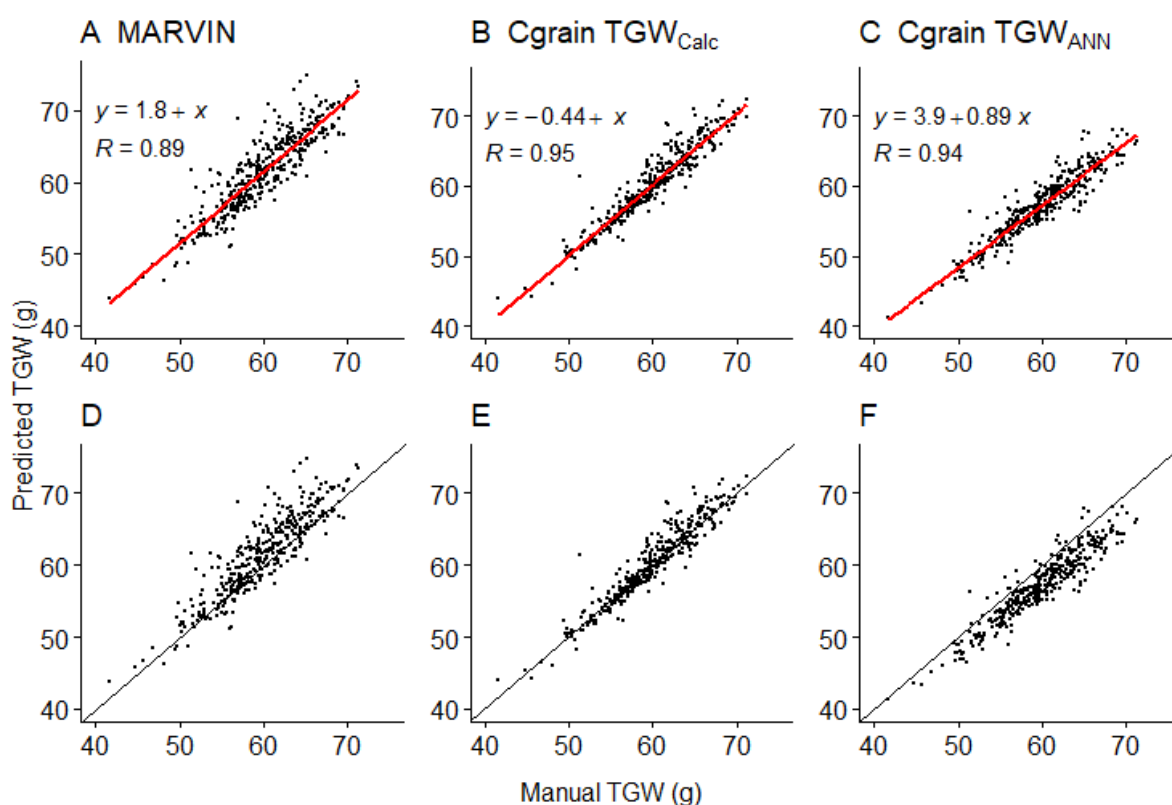


Figure 2. (A–C) Scatter plots of manual TGW (g) vs. TGW (g) as predicted by MARVIN, by Cgrain calculation from mean seed weight (TGW_{Calc}), and Cgrain ANN output (TGW_{ANN}) with linear regression lines and corresponding r values and line equations. (D–F) Scatter plots of manual TGW (g) vs. TGW (g) as predicted by MARVIN, by calculation from Cgrain mean seed weight (TGW_{Calc}), and Cgrain ANN output (TGW_{ANN}), shown with the line of equality ($y = x$) for the manual TGW (g) data.

3.2. All Seeds—Grain Length

In addition to obtaining data on the mean grain parameters of each plant, both MARVIN and Cgrain provide data on each individual seed that is analysed. This reveals the

wide range of sizes found within the grain from an individual plant. For example, the distribution of length measurements of each individual grain analysed by MARVIN and Cgrain are shown in Figure 3. The MARVIN data have more longer grains, whereas the Cgrain analysis indicates a much tighter distribution of grain lengths with no grain longer than 19.2 mm. Oat grain length data distribution for both machines is non-parametric and multi-modal. The distributions differ significantly between machines (Wilcoxon Signed Rank Test, $p < 0.0001$).

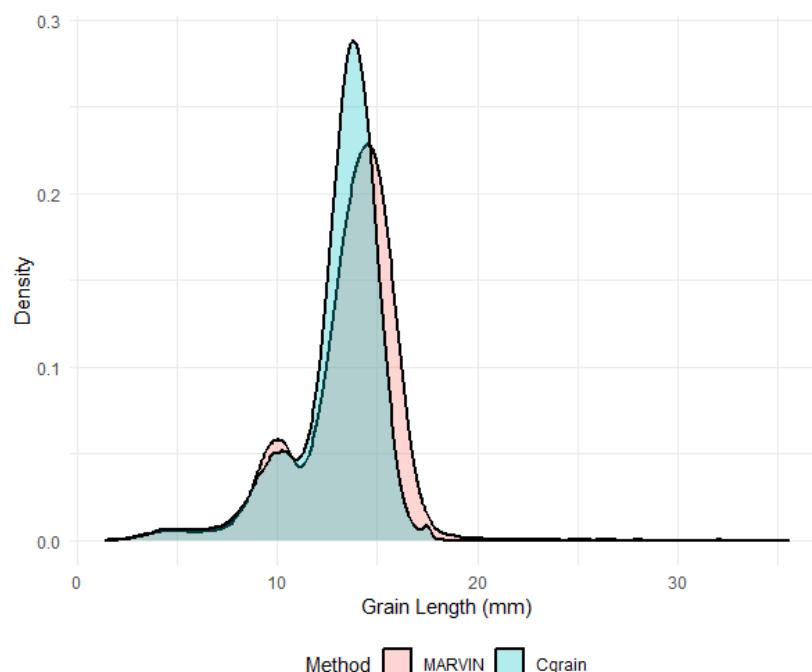


Figure 3. Density plot comparing the data distributions of grain length from all seeds from all plants in the study as predicted by the two machines.

3.3. Plant Means—Grain Length, Width, Roundness, and Area

3.3.1. Grain Length

A Shapiro–Wilk test suggested that the mean grain length for all plants in the population assumes a non-parametric distribution when measured by MARVIN ($W = 0.99$, p -value = 0.02) and is normally distributed when measured by Cgrain ($W = 1.0$, p -value = 0.32). The mean grain length per plant was significantly different between genotypes for both machines (MARVIN: Kruskal–Wallis Test, p -value < 0.0001, Cgrain: ANOVA, p -value < 0.0001).

The mean grain length per plant was significantly lower when measured by Cgrain compared to MARVIN (Wilcoxon Signed Rank Test, $p < 0.0001$) but the mean length data were strongly correlated between the two machines ($r = 0.94$) (Figure 4A).

The plot of the results, as compared to the line of equality (Figure 4B) and the Bland–Altman plot (Figure 5A), shows a proportional bias: as the grain length increases, the extent to which Cgrain apparently underpredicts the grain length also increases. For the Bland–Altman analysis, the limits of agreement are predicted by the 2.5th and 97.5th percentiles, meaning approximately 95% of the plant mean grain length results when measured by Cgrain were between equal (0.0 mm) and 1.3 mm lower than the MARVIN prediction, with a mean bias of -0.5 mm.

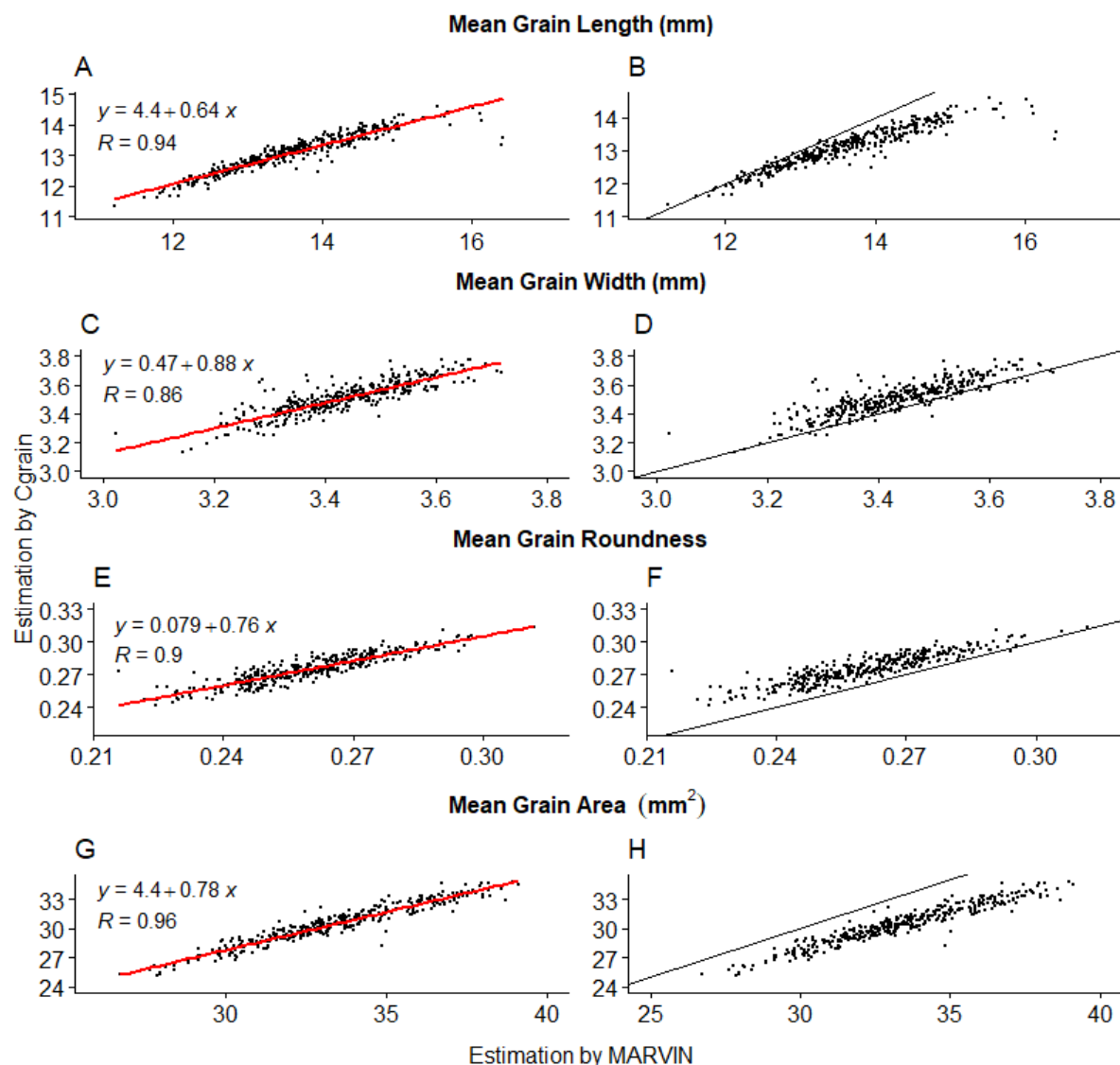


Figure 4. (A) Scatter plot of mean grain length when measured by MARVIN vs. Cgrain, shown with linear regression line and corresponding r value and line equation. (B) Scatter plot of mean grain length when measured by MARVIN vs. Cgrain, shown with the line of equality ($y = x$) for the MARVIN data. (C) Scatter plot of mean grain width when measured by MARVIN vs. Cgrain, shown with linear regression line and corresponding r value and line equation. (D) Scatter plot of mean grain width when measured by MARVIN vs. Cgrain, shown with the line of equality ($y = x$) for the MARVIN data. (E) Scatter plot of mean grain roundness when measured by MARVIN vs. Cgrain, shown with linear regression line and corresponding r value and line equation. (F) Scatter plot of mean grain roundness when measured by MARVIN vs. Cgrain, shown with the line of equality ($y = x$) for the MARVIN data. (G) Scatter plot of mean grain area when measured by MARVIN vs. Cgrain, shown with linear regression line and corresponding r value and line equation. (H) Scatter plot of mean grain area when measured by MARVIN vs. Cgrain, shown with the line of equality ($y = x$) for the MARVIN data.

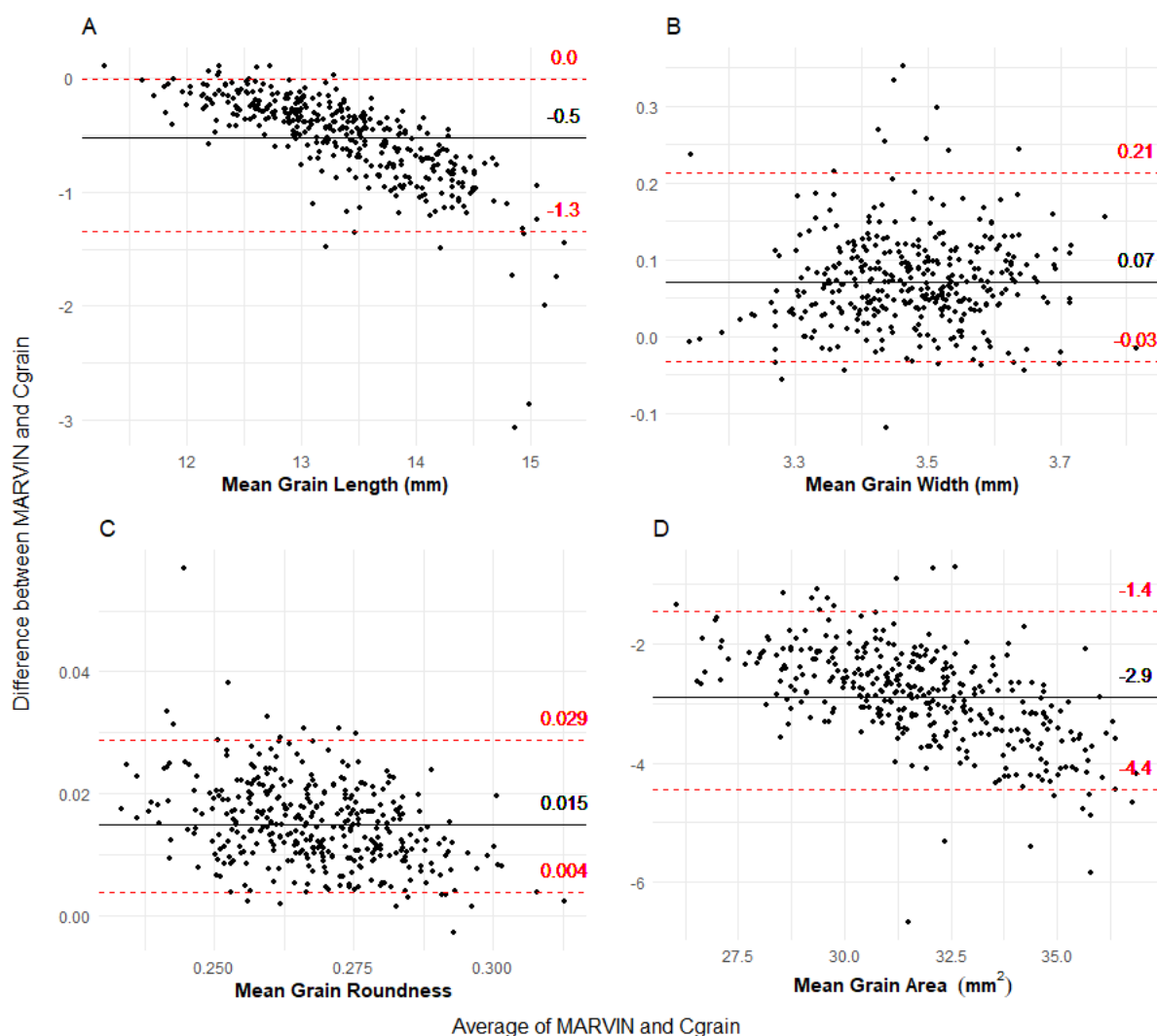


Figure 5. Bland–Altman plot showing nature of differences in measurement between machines for (A) mean grain length, (B) mean grain width, (C) mean grain roundness, and (D) mean grain area. For each plot, the horizontal axis represents the average result for each plant by the two machines and the vertical axis represents the difference in results between the two machines for each plant (Cgrain minus MARVIN). The horizontal black line and associated value show the ‘bias’, i.e., the mean difference when measured by Cgrain vs. MARVIN. Red dashed lines and associated values represent the limits of agreement as estimated by the 2.5th and 97.5th percentiles.

3.3.2. Grain Width

A Shapiro–Wilk test suggested that mean grain width for the all plants in the population assumes a normal distribution for both machines (MARVIN: $W = 1.0$, p -value = 0.71, Cgrain: $W = 1.0$, p -value = 0.89). The mean grain width per plant was significantly different between genotypes for both machines (MARVIN: ANOVA, p -value < 0.0001, Cgrain: ANOVA, p -value < 0.0001).

The mean grain width per plant was significantly higher when measured by Cgrain compared to MARVIN (Wilcoxon Signed Rank Test, $p < 0.0001$). The Cgrain mean width data were moderately well correlated to that collected by MARVIN ($r = 0.86$) (Figure 4C).

The plot of the results, as compared to the line of equality (Figure 4D) and the Bland–Altman plot (Figure 5B), does not appear to show any strong discernible proportional bias. For the Bland–Altman analysis, the limits of agreement are predicted by the 2.5th and 97.5th percentiles, meaning approximately 95% of the plant mean grain width results when

measured by Cgrain were between 0.21 mm higher and 0.03 mm lower than the MARVIN prediction, with a mean bias of +0.07 mm.

3.3.3. Grain Roundness

A Shapiro–Wilk test suggested that the mean grain roundness for the all plants in the population assumes a normal distribution for both machines (MARVIN: $W = 1.0$, p -value = 0.87, Cgrain: $W = 1.0$, p -value = 0.49). The mean grain roundness per plant was significantly different between genotypes for both machines (MARVIN: ANOVA, p -value < 0.0001, Cgrain: ANOVA, p -value < 0.0001).

The mean grain roundness per plant was significantly higher when measured by Cgrain compared to MARVIN (Wilcoxon Signed Rank Test, $p < 0.0001$). The Cgrain mean roundness data were strongly correlated to that collected by MARVIN ($r = 0.90$) (Figure 4E).

The plot of the results, as compared to the line of equality (Figure 4F) and the Bland–Altman plot (Figure 5C), does not appear to show any strong discernible proportional bias. For the Bland–Altman analysis, the limits of agreement are predicted by the 2.5th and 97.5th percentiles, meaning approximately 95% of the plant mean grain roundness results when measured by Cgrain were between 0.004 lower and 0.029 higher than the MARVIN prediction, with a mean bias of +0.015.

3.3.4. Grain Area

A Shapiro–Wilk test showed that the mean grain area for the all plants in the population assumes a normal distribution for both machines (MARVIN: $W = 0.99$, p -value = 0.085, Cgrain: $W = 0.99$, p -value = 0.11). The mean grain area per plant was significantly different between genotypes for both machines (MARVIN: ANOVA, p -value < 0.001, Cgrain: ANOVA, p -value < 0.001).

The mean grain area per plant was significantly lower when measured by Cgrain compared to MARVIN (Wilcoxon Signed Rank Test, $p < 0.0001$). The Cgrain mean area data were strongly correlated to that collected by MARVIN ($r = 0.96$) (Figure 4G).

The plot of the results, as compared to the line of equality (Figure 4H) and the Bland–Altman plot (Figure 5D), suggests a weak proportional bias: as grain area increases, the extent to which Cgrain apparently underpredicts grain area also increases. For the Bland–Altman analysis, the limits of agreement are predicted by the 2.5th and 97.5th percentiles, meaning approximately 95% of plant mean grain area results when measured by Cgrain were between 1.4 mm² lower and 4.4 mm² lower than the MARVIN prediction, with a mean bias of −2.9 mm².

3.4. Cgrain Specific Traits: Grain Weight, Thickness, Cross-Sectional Roundness, and Volume

Additional per-grain morphometric traits that are not obtainable from MARVIN are measured by Cgrain. These include grain weight, 3-D estimations of thickness (and the derivative cross-sectional roundness), and volume.

The mean grain weight as measured by Cgrain has been separately discussed as TGW_{Calc} , (which is the same number as mean grain weight in milligrammes): the distribution can be assumed normal and there are significant differences between genotypes. (Shapiro–Wilk test: $W = 0.99$, p -value = 0.07, ANOVA: p -value < 0.0001) (Figure 6A).

A Shapiro–Wilk test suggested that the mean grain thickness for the all plants in the population as determined by Cgrain assumes a non-parametric distribution ($W = 0.99$, p -value < 0.01) (Figure 6B). The mean grain thickness per plant varied significantly between genotypes (Kruskal–Wallis Test, p -value < 0.0001).

A Shapiro–Wilk test suggested that the mean grain cross-sectional roundness for all plants in the population as determined by Cgrain assumes a non-parametric distribution ($W = 0.99$, p -value = 0.03) (Figure 6C). The mean grain cross-sectional roundness per plant varied significantly between genotypes (Kruskal–Wallis Test, p -value < 0.0001).

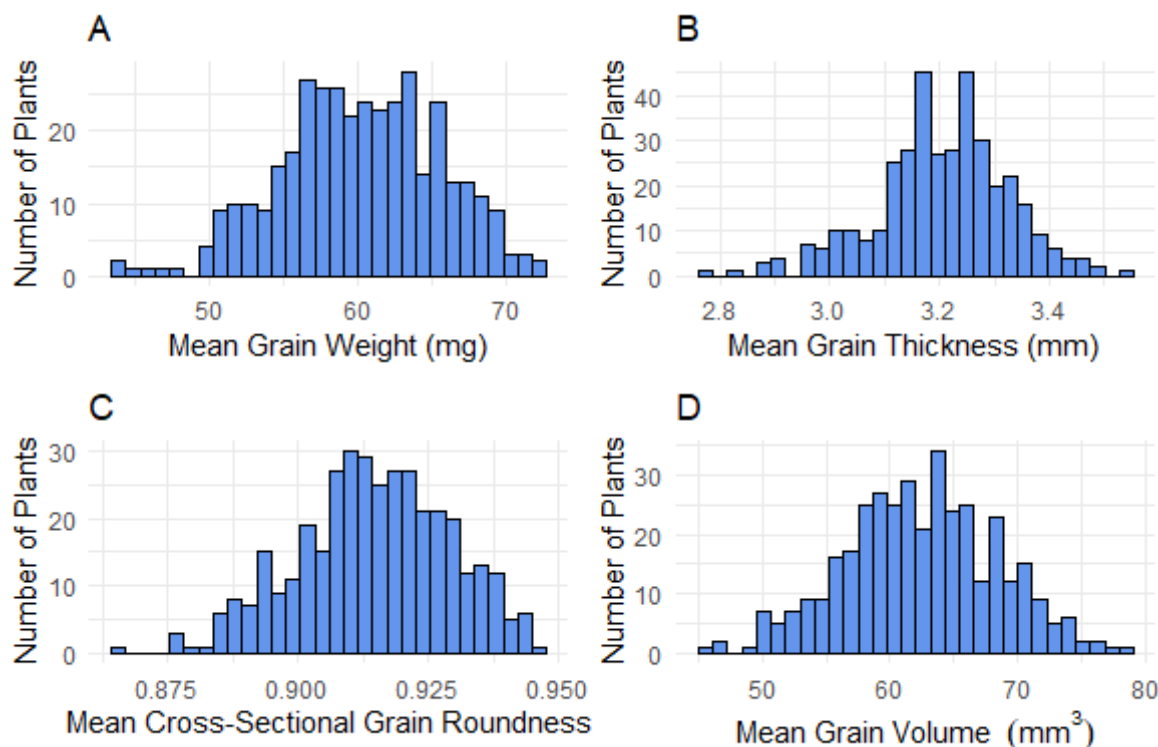


Figure 6. Histograms showing the distributions of per-plant means for the additional morphometric traits estimated by Cgrain: (A) mean grain weight, (B) mean grain thickness, (C) mean cross-sectional grain roundness, and (D) mean grain volume.

A Shapiro–Wilk test suggested that the mean grain volume for the all plants in the population as determined by Cgrain is normally distributed ($W = 1.0$, p -value = 0.90) (Figure 6D). The mean grain volume per plant varied significantly between genotypes (ANOVA, p -value < 0.0001).

3.5. Cgrain: Repeatability of Geometric Measurements

The seeds from all plants were analysed twice with Cgrain and a comparison between the two sets of results obtained was made. With the exception of the mean grain length, geometric measurements were significantly different between passes. Repeat analyses correlated strongly ($r \geq 0.94$) with the exception of the mean grain cross-sectional roundness, which was relatively poor ($r = 0.82$). Table 2 summarises the differences/agreement between the two replicates for the overall plant means of grain morphometric measurements.

Table 2. Comparison of the 1st and 2nd passes of each plant sample through the Cgrain.

Measurement	r	Wilcoxon Signed Rank Test p -Value	Median Paired Difference ¹	Median Difference 95% Confidence Interval ²	95% Limits of Agreement ³	Coefficient of Variation (CV) (%)
Mean grain weight (mg)	0.94	<0.0001	−0.35	−0.48, −0.23	−3.8, +3.8	2.2
Grain length (mm)	0.96	0.13	−0.01	−0.02, +0.00	−0.4, +0.3	1.0
Grain width (mm)	0.95	<0.0001	−0.016	−0.020, −0.014	−0.04, +0.10	0.9
Grain roundness	0.94	<0.0001	−0.0012	−0.0016, −0.0008	−0.008, +0.010	1.3
Grain thickness (mm)	0.96	<0.0001	−0.009	−0.012, −0.006	−0.05, +0.08	0.7
Cross-sectional grain roundness	0.82	<0.001	+0.0015	+0.0007, +0.0023	−0.016, +0.015	0.7

Table 2. Cont.

Measurement	<i>r</i>	Wilcoxon Signed Rank Test <i>p</i> -Value	Median Paired Difference ¹	Median Difference 95% Confidence Interval ²	95% Limits of Agreement ³	Coefficient of Variation (CV) (%)
Grain area (mm ²)	0.95	<0.0001	−0.14	−0.18, −0.09	−1.0, +1.2	1.4
Grain volume (mm ³)	0.98	<0.0001	−0.34	−0.44, −0.24	−1.5, +3.0	1.4

¹ Median of 2nd pass minus 1st pass for each plant. ² From Wilcoxon signed rank test. ³ The 2.5th and 97.5th percentiles of paired differences.

3.6. Association between Grain Traits

A correlation plot of the genotype means of morphometrical traits from each of the machines is given in Figure 7, with their respective correlation coefficients. Biplots of the first (PC1) and second (PC2) principle components for the morphometrical traits as measured by Marvin and Cgrain are shown in Figure 8. The length of the arrows indicates the relative contribution of that trait to the overall variation between genotypes. The angle between any pair of arrows describes the relationship between those traits with acute or obtuse angles, respectively, indicating a positive or negative correlation. For the data from MARVIN, grain weight was highly correlated with grain width and area. This was also found with the data from Cgrain but this analysis revealed that the 3-D traits not measured by MARVIN, grain thickness and grain volume, were also highly correlated with grain weight. Grain roundness was negatively correlated with grain length for both machines.

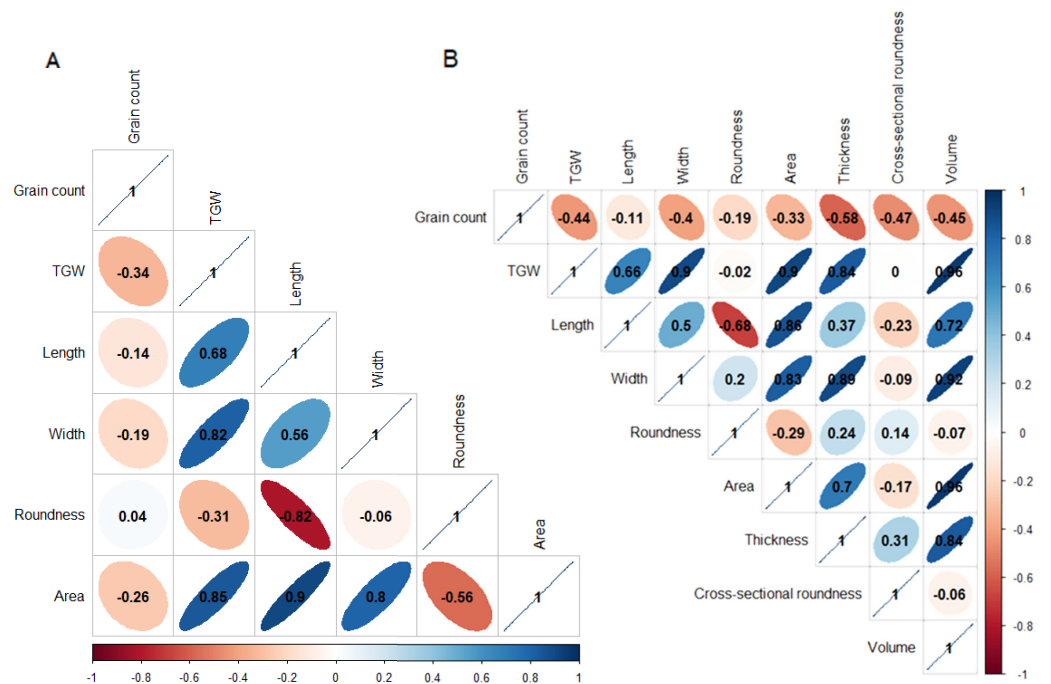


Figure 7. Correlation plots for the genotype means of MARVIN (A) and Cgrain (B) grain count and geometric measurements with Pearson correlation coefficients.

3.7. Cgrain: Grain Colour

The Cgrain outputs a colour measurement per grain, representing the average pixel colour from the segmented image in both RGB and HSL formats.

In the RGB colour space, the quadratic means for grain red (R), green (G), and blue (B) values (per plant) all formed non-parametric distributions (Shapiro–Wilk test, $W = 0.99, 0.99, 0.98$, p -value $< 0.01, < 0.001, < 0.0001$, respectively). R, G, and B, when tested independently, varied significantly by genotype (Kruskal–Wallis Test, p -value < 0.0001 for all) (Figure 9A–C).

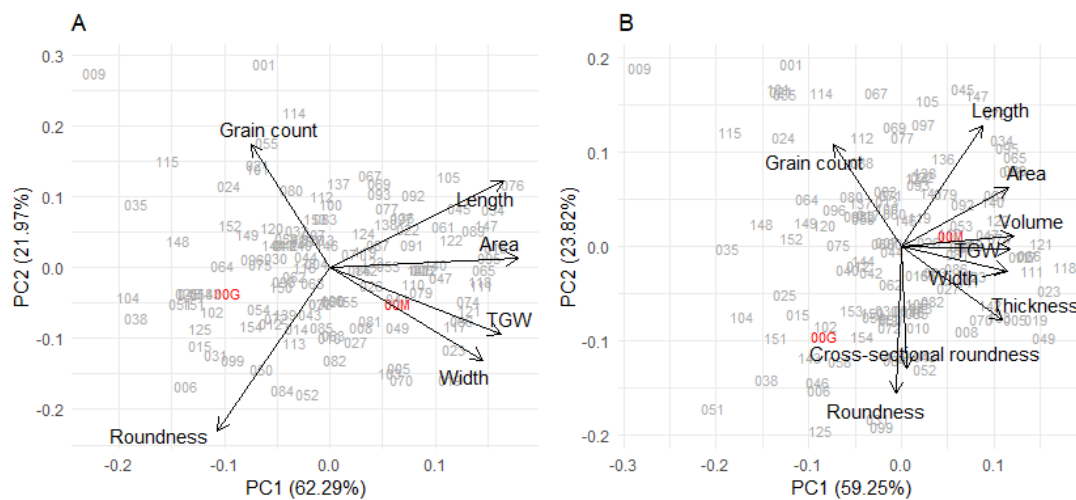


Figure 8. Biplots for genotype means of geometric traits from MARVIN (A) and Cgrain (B). The individual genotypes are plotted by their 3-digit identifier with the parents, “00G”—Gerald and “00M”—Mascani, in red.

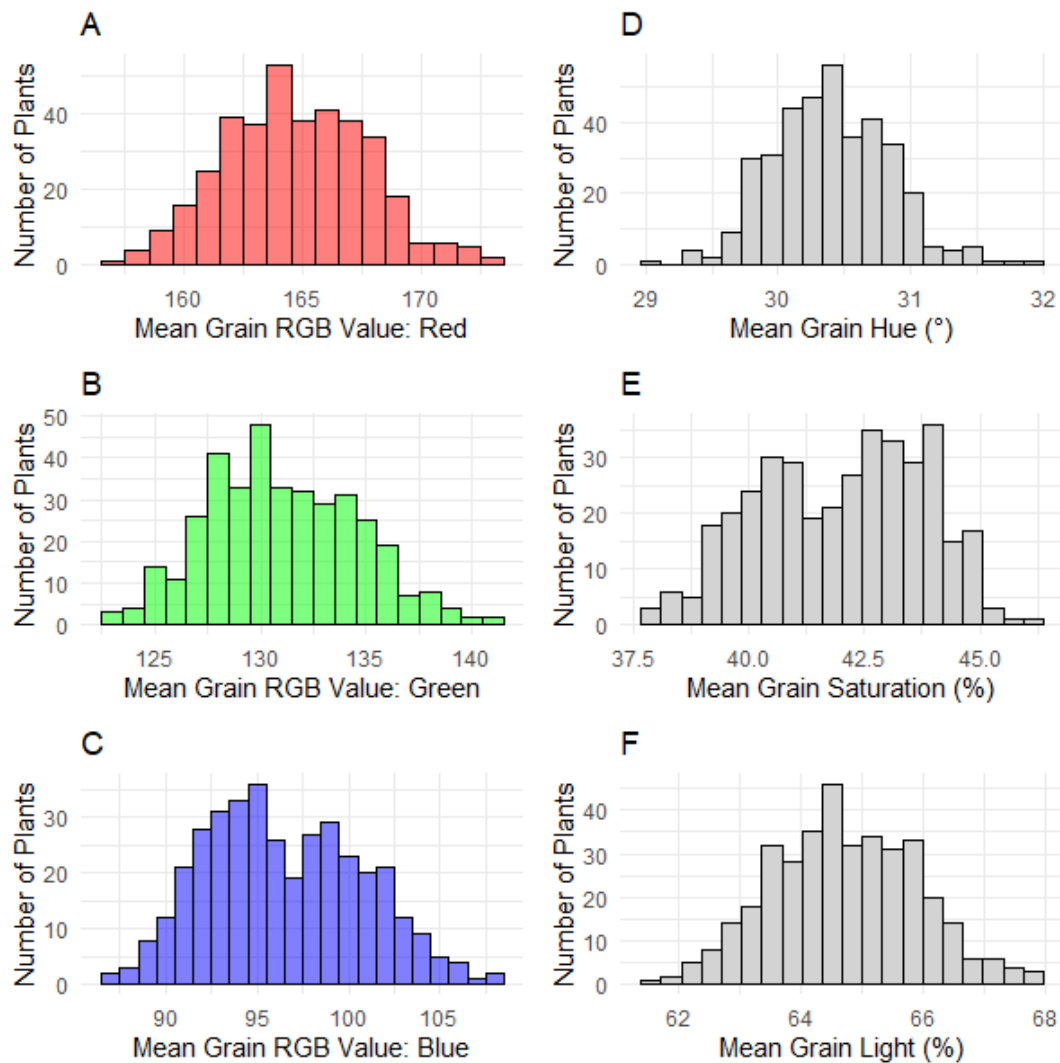


Figure 9. RGB and HSL colour space data, as imaged by the Cgrain. Histograms showing approximate distribution of mean segmented grain red intensity (A), green intensity (B), blue intensity (C), hue (D), saturation (E), and light (F) for all plants in the experiment.

In the HSL colour space, the mean grain hue and light formed normal distributions accepted by a Shapiro–Wilk test ($W = 0.99$, 1.0 , p -value < 0.07 , 0.84 , respectively). The distribution for saturation was rejected by the same test ($W = 0.98$, p -value < 0.0001). The hue and light values varied significantly by genotype (ANOVA, p -value < 0.0001 for both). The saturation also varied significantly by genotype (Kruskal–Wallis Test, p -value < 0.0001) (Figure 9D–F).

4. Discussion

4.1. Summary of Findings

For all the shared plant trait measurements between the two machines used (mean grain length, width, roundness, and area), the two machines produced data that was well correlated but significantly different. The Cgrain overall predicted shorter, wider grains than MARVIN. For grain length, and the length-derivative grain area, Cgrain showed a proportional bias relative to MARVIN. Grain roundness did not appear to show a proportional bias. The morphometric plant trait measurements estimated only by Cgrain—mean grain weight, thickness, cross-sectional roundness, and volume showed significant differences between genotypes. The mean grain colour also showed significant differences between genotypes.

The Cgrain two-pass analysis showed small but statistically significant differences between the first and second passes for the plant means for all geometric traits except mean grain length. For all measurements, the correlation between the first and second passes were strong, apart from cross-sectional roundness, which showed relatively poor correlation.

The two machines produced data with similar resulting trait associations.

4.2. Cgrain vs. MARVIN

The TGW data, the only character in this study that we could compare against manual data, showed statistically significant differences between machines. TGW_{ANN} —the Cgrain automatic “black box” output of TGW (this is the value that the machine operator sees on the summary screen immediately after analysis) had the highest paired differences from the manual value. These differences (mean = -2.7 g) would not be tolerable in an industrial context but given that the relationship is linear and strongly correlated, a simple bias adjustment in the model would likely transform the TGW_{ANN} output to an acceptable degree of agreement. The calculated TGW_{Calc} result from the Cgrain analysis was not significantly different from the manual result. The MARVIN TGW results were significantly higher than their paired manual results; the differences were modest (mean = $+1.5$ g) but the correlation was the weakest of the methods tested. The larger TGW prediction when measured by MARVIN is likely due to touching grains—these groups of two or more grains would have been counted as single objects in the analysis and, therefore, reducing the total object count for the same given sample weight would have elevated the TGW result. It is possible, through careful operation of the MARVIN, to reduce the frequency of touching grains, but this comes at the cost of a reduced throughput. In contrast, the mechanism by which the Cgrain presents the sample into the imaging area almost completely removes this issue. All three predictive methods for TGW have produced a dataset that could be legitimately used for a genome-wide association study (GWAS) or a similar trait association study. It is worth noting that the glasshouse-grown oat plants used in this study appear to have produced atypically large grains compared to field-grown material. For example, the parents of the mapping population, Gerald and Mascani, produced grains with a TGW averaging 37.4 g and 45.4 g, respectively, when grown in multi-site, multi-year field trials [9], whereas the same genotypes averaged 55.1 g and 62.1 g, respectively, in this project. The overall TGW data in this study is substantially higher than what was seen in several published multi-year, multi-site (temperate), and multi-cultivar studies [35–37], and grain length is also higher than usually reported [9,38], supporting the view that the oat grains under the test were atypically large.

Grain lengths, when looking at the entire seed collection, display a distribution characteristic of oats. The within-sample distribution of oat grain size data are of particular interest compared to other cereal grains, due to the spikelet development and multifloret structure. Spikelets may contain up to three grains, with the primary kernel being the largest. The size differential between the primary and secondary or tertiary grains give rise to bi-modal and tri-modal distributions, and the relative proportions of each from a given plant are affected by both the environment and genotype [39,40]. This particular trait is of interest to the milling industry and thence to the breeding programmes. The results from this study and the distributions, illustrated in Figure 3, further support this: there are two distinct peaks, with the largest likely to be representing primary grains, and the smaller peak likely to be representing secondary grains, hulled groats, or a combination thereof. The distributions for the two machines closely overlap for the smaller peak at approximately 10 mm but are not as strongly aligned for the larger peak, which occurs at about 14 mm for Cgrain and 15 mm for MARVIN, with the latter also having a longer tail. It is possible that this length distribution skewing of the larger grains is due to awned grains; the MARVIN output will add awns to the length parameter, whereas the Cgrain ANN model may remove them as not being part of the grain. A significant contributor to the differences in length measurement of the longer grains may be the limited size of the Cgrain's image capture area. During analysis on the Cgrain, image capture is triggered by the falling grain at a rate of several grains per second. In most cases, this millisecond-precise trigger/capture event is successful, and the resultant images look like the examples in Figure 1. In some instances, at the point of image capture, one end of the grain is outside the image capture area and the resultant image is of an incomplete oat (Figure 10). Intuitively, and from the dataset produced in this study, it appears likely that this "cutting" of oats is more likely to occur for longer grains than shorter ones. The image capture area of the Cgrain instrument is approximately 19.1 mm long, which serves as a theoretical maximum grain length, but from the entire Cgrain dataset in this study, the longest grain successfully captured fully within the image capture area was 16.7 mm. The Cgrain can classify these problem images as "cut", but exactly how they are handled by the ANN is unknown. The length data in the raw output (and length-derivative data such as area and volume) appear to represent only the oat that is seen, i.e., there is no artificial extrapolation or modelling to predict the final geometry of an incomplete oat. MARVIN, by contrast, has a significantly larger imaging area (200 mm × 200 mm) and therefore has no practical upper limit of oat grain length prediction.

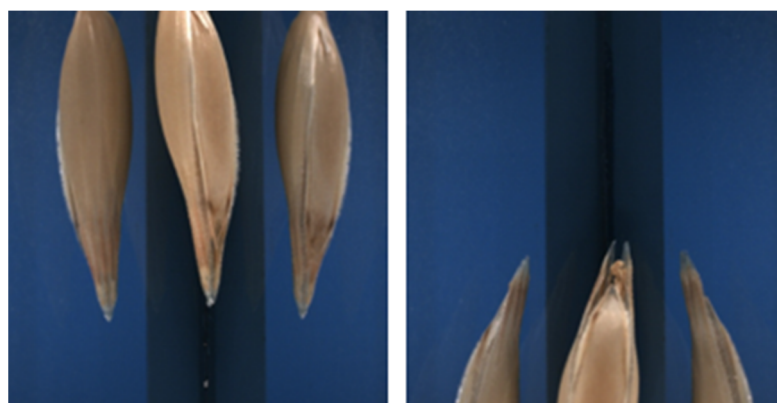


Figure 10. Typical (left) and extreme (right) examples of grain images being "cut".

Cgrain underpredicts the mean grain length per plant relative to MARVIN, and the extent worsens as the mean grain length increases. The plots in Figures 4 and 5 show a grouping of data for the Cgrain mean grain length analysis that appear non-linear relative to MARVIN and is possibly asymptotic to the maximum possible mean grain length that the Cgrain is able to predict.

This limitation of grain length appears to be a sizable disadvantage of the Cgrain when measuring oat material with larger grains. As previously discussed, the oat grains in this study are substantially larger than usually seen from field grown material. It is possible, therefore, that the material presented to the Cgrain in this study was sufficiently atypically large that some of the grains were outside its scope of operation. It is also possible that the Cgrain could be physically optimised (e.g., zooming out of the camera) to increase the maximum grain size but this feature is not currently available as an operator-level adjustment.

Despite this apparent limitation, the Cgrain produced a dataset for grain length that correlated well with MARVIN data over the range of genotypes examined. Since there was significant variation between genotypes, the Cgrain data could therefore be deemed suitable for genetic analysis.

The mean grain width per plant was significantly higher when measured by Cgrain relative to MARVIN, with a mean bias of 0.07 mm, but the data obtained from the two machines was highly significantly correlated. This relatively small bias is unlikely to be a concern at an industrial level, but the upper 95% limit of agreement at 0.21 mm would be significant. This observed difference could be due to the rotational orientation of grains in the MARVIN tray at time of image capture. In the MARVIN analysis, oat grains are unlikely to be resting with their widest width perfectly horizontal relative to the top-down camera—meaning the 2-D width prediction is likely to be slightly lower than its true value, whereas the Cgrain can make a measurement closer to the widest width of the entire grain. This problem can be mitigated in MARVIN analysis by careful positioning of the individual grains and resting them individually on the ventral furrow, but this is impractical at scale. This operational issue with MARVIN is known [41]. Both machines have produced width data that correlate well and vary significantly by genotype, with the Cgrain data possibly revealing more nuanced width information at a higher resolution, meaning that it may be suitable for GWAS.

The mean grain roundness per plant varied significantly between machines. Roundness is a function of the width and inverse length, and given that the Cgrain underpredicts length and reports wider widths relative to MARVIN, it follows that the Cgrain generates data with significantly higher roundness than MARVIN. Roundness (sometimes discussed in the literature in terms of its inverse “L/W ratio”) has wide implications for quality, storage, and processing (e.g., in the de-hulling or milling process) and is therefore an important trait [42]. It appears that neither machine’s roundness prediction method is ideal, with the apparent underprediction of length by Cgrain and width by MARVIN. However, the roundness data correlates well between machines and varies significantly by genotype, meaning that the Cgrain roundness data may be suitable for a GWAS.

The industrial relevance of the mean grain area as a trait is difficult to determine. It has served as a proxy for overall grain size in 2-D imaging but suffers in the same way as grain width in that the result is affected by the rotational orientation of the grain at the time of image capture, and, in any orientation, tells little about the geometry of the grain’s vertical plane relative to the surface. Cgrain underpredicts the area relative to MARVIN, possibly due to the discussed issue of “cut” images and MARVIN potentially incorporating awns into the total area where present. The statistically significant differences in the mean grain area between the machines are proportional in the same way as grain length, i.e., as grain area increases, the degree of disagreement also increases, suggesting the disagreement is more strongly linked to the problem of “cut” images in the Cgrain analysis. Again, the measurements of grain area by Cgrain are strongly correlated to the MARVIN data with significant differences between genotypes, making the dataset potentially valid for GWAS.

The thickness of the grain is a measurement only made by Cgrain in this study. It has industrial relevance in that the relative proportions in mechanical grain sieving fractions and screenings (the amount of grain that passes through a calibrated sieve, usually 2.0 mm for husked oats) is an important quality trait at end use [43]. The material that passes through the screen (known as “thrus”, “thins”, or “reject fraction”, among other

names) is of little value and therefore genotypes with a phenotype of high screenings are treated unfavourably in breeding and industry. The thickness parameter, being “thinner” than width, could be more likely to give a prediction of the screening trait and could therefore be considered with interest. This trait varied significantly between genotypes and therefore could be considered for GWAS. The relationship between the mean grain width and thickness is fairly strongly correlated, suggesting that the thickness of corns generally increases linearly with increased width. But there is interesting information in that specific relationship between the width and thickness of each grain, which we discuss in terms of “cross-sectional roundness”.

Cross-sectional roundness has the potential to be a character of interest in the milling industry in the same way as roundness, and is possibly a way of converting qualitative assessments of “plumpness” or “boldness” into a quantitative one. Cross-sectional roundness may also describe the depth of the ventral crease, which has implications in milling efficiency [44]. Cross-sectional roundness varied significantly by genotype and could therefore be considered suitable for GWAS.

Grain volume describes the overall 3-D space that a grain occupies and is therefore linked to several important quality traits and yield. The manner in which oat grain volume (and other 3-D measurements) are specifically related to processability traits or traits of nutrition would represent an area for valuable further study. The oat grain volume in this study showed significant variation between genotypes and could therefore be suitable for a GWAS.

The trait association by principal component analysis (PCA) for both machines (Figure 8) gives similar interpretations of the sources of phenotypic variation, with PC1 principally being grain “size” traits, essentially “smaller to bigger”, and PC2 being predominantly about grain “shape”, or “skinny to plump”. The PCA positioned the parent genotypes into similar positions observed in [9], with Mascani typically having larger grains and Gerald having smaller and slightly rounder grains. The trait associations for both machines (Figure 7) show a negative relationship between the total seed count and seed size—particularly TGW—which is a well-documented “trade-off” in seed crops [45,46]. The relatively strong negative relationship between the total seed count and the 3-D geometric traits as measured by Cgrain (thickness, cross-sectional area and volume) potentially offer valuable new insight into this significant breeding “hurdle”: in this study, it appears that the higher grain numbers produced per plant was a trade-off for plumpness, similar to what has been reported for barley, with sieving analysis used as a proxy for grain width and thickness [47].

4.3. Cgrain Repeatability

For all geometrical analyses, bar grain length, the plant means were statistically different from one another between analysis rounds. The mean grain geometries were lower on the second pass, giving overall smaller grain size results. However, the paired differences, limits of agreement, and the coefficients of variation suggest that the repeat passes have produced results that are within what would be considered a reasonable measurement error if using a manual method, e.g., vernier callipers.

The assumptions of these repeatability tests are that the “true” result does not change between replicates—but this is not necessarily the case with material such as this. It is possible that, through the mechanical agitation of sampling and passage through the machine, the grain sizes are slightly altered (especially at the fragile points at the tips of the husks). It is also possible that the grain lost a proportion of its volume via moisture loss, given that the material was in ambient storage for several days between the first and second passes. Interestingly, the grains were measured as having a higher degree of cross-sectional roundness on the second pass—if the hypothesis of lost moisture content is accurate, then this dataset could give valuable insight into the nature of post-harvest geometry changes, such as those described for maize and rice [48].

4.4. Grain Colour

The mapping population parents, Gerald and Mascani, are classed as “white” and “yellow”, respectively [49,50], meaning that colour variation in the progeny was expected. All six values for the mean grain colour that were gathered by the machine (RGB and HSL) varied significantly by genotype and may be suitable for GWAS. To increase the industrial relevance and usefulness of these measurements, a system could potentially be developed to classify samples into the colour groups currently used by plant breeders and marketers. With sufficient data, such an approach should be possible and offers advantages over screening by human eye [51], and there have been successful models for cereal colour classification, such as red and white wheat [52].

5. Conclusions

Despite being targeted primarily to the commercial sector, Cgrain produced a set of data that agrees very well with other current seed phenotyping methods, both in terms of accuracy and throughput. Thus, the machine could be highly suitable for a wide variety of research purposes. The population under test represents the phenotypic diversity typically seen in a breeding programme, which is much less than typically seen in GWAS studies. Nevertheless, we found statistically significant phenotypic differences between genotypes over a range of measurements relating to grain size, shape, and appearance. For the traits that could be measured by MARVIN, the correlation was high.

Moreover, Cgrain already provides quantitative data on a wider range of traits than MARVIN, in part due to the combination of novel imaging and its internal ANN. Thus, several additional physical traits such as colour, thickness, and volume were provided, and others might be inferred. Additional traits may require additional training, but the machine returns the raw image data to the user and bespoke data analysis routines could be developed. The precise grain-by-grain trait information could also provide a scope for different types of data analysis. One potential limitation is how the Cgrain deals with long grains: the length of these grains tends to be underestimated.

Therefore, with a potential caveat on length, we conclude the Cgrain Value™ offers an informative and high-throughput oat grain phenotyping capability.

Author Contributions: Conceptualisation, D.E.; methodology, D.E.; validation, D.E., J.H.D. and C.J.H.; formal analysis, D.E.; investigation, E.J.D., R.H. and D.E.; resources, F.C. and R.H.; data curation, E.J.D. and D.E.; writing—original draft preparation, D.E.; writing—review and editing, J.H.D., C.J.H. and D.E.; visualisation, D.E.; supervision, J.H.D. and C.J.H.; project administration, F.C., C.J.H. and J.H.D.; funding acquisition, C.J.H. and J.H.D. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by UKRI BBSRC FoodBioSystems Doctoral Training Partnership (DTP), grant number BB/T008776/1. C.J.H. and J.D. were supported by BBSRC grants BBS/E/IB/23001 and BBS/E/W/0012843. We also acknowledge financial support from the Healthy Oats project, which has been part funded by the European Regional Development Fund through the Ireland Wales Programme.

Data Availability Statement: Data are available from the Aberystwyth Data Repository: <https://doi.org/10.20391/cec25136-bdb1-4f7a-a08b-8c9abfda720e>.

Acknowledgments: We thank Ben Hutchinson of Calibre Control International Ltd. for the loan of a trial Cgrain instrument.

Conflicts of Interest: The authors declare no conflicts of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Xiao, Q.; Bai, X.; Zhang, C.; He, Y. Advanced high-throughput plant phenotyping techniques for genome-wide association studies: A review. *J. Adv. Res.* **2022**, *35*, 215–230. [CrossRef]

2. Gegas, V.C.; Nazari, A.; Griffiths, S.; Simmonds, J.; Fish, L.; Orford, S.; Sayers, L.; Doonan, J.H.; Snape, J.W. A Genetic Framework for Grain Size and Shape Variation in Wheat. *Plant Cell* **2010**, *22*, 1046–1056. [\[CrossRef\]](#)
3. Schierenbeck, M.; Alqudah, A.M.; Lohwasser, U.; Tarawneh, R.A.; Simón, M.R.; Börner, A. Genetic dissection of grain architecture-related traits in a winter wheat population. *BMC Plant Biol.* **2021**, *21*, 417. [\[CrossRef\]](#)
4. Lisker, A.; Maurer, A.; Schmutzer, T.; Kazman, E.; Cöster, H.; Holzapfel, J.; Pillen, K. A Haplotype-Based GWAS Identified Trait-Improving QTL Alleles Controlling Agronomic Traits under Contrasting Nitrogen Fertilization Treatments in the MAGIC Wheat Population WM-800. *Plants* **2022**, *11*, 3508. [\[CrossRef\]](#)
5. Xu, X.; Sharma, R.; Tondelli, A.; Russell, J.; Comadran, J.; Schnaithmann, F.; Pillen, K.; Kilian, B.; Cattivelli, L.; Thomas, W.T.B.; et al. Genome-Wide Association Analysis of Grain Yield-Associated Traits in a Pan-European Barley Cultivar Collection. *Plant Genome* **2018**, *11*, 170073. [\[CrossRef\]](#)
6. Shrestha, A.; Cosenza, F.; van Inghelandt, D.; Wu, P.-Y.; Li, J.; A Casale, F.; Weisweiler, M.; Stich, B. The double round-robin population unravels the genetic architecture of grain size in barley. *J. Exp. Bot.* **2022**, *73*, 7344–7361. [\[CrossRef\]](#)
7. Neuweiler, J.E.; Maurer, H.P.; Würschum, T. Long-term trends and genetic architecture of seed characteristics, grain yield and correlated agronomic traits in triticale (xTriticosecale Wittmack). *Plant Breed.* **2020**, *139*, 717–729. [\[CrossRef\]](#)
8. Cao, D.; Wang, D.; Li, S.; Li, Y.; Hao, M.; Liu, B. Genotyping-by-sequencing and genome-wide association study reveal genetic diversity and loci controlling agronomic traits in triticale. *Theor. Appl. Genet.* **2022**, *135*, 1705–1715. [\[CrossRef\]](#)
9. Howarth, C.J.; Martinez-Martin, P.M.J.; Cowan, A.A.; Griffiths, I.M.; Sanderson, R.; Lister, S.J.; Langdon, T.; Clarke, S.; Fradgley, N.; Marshall, A.H. Genotype and Environment Affect the Grain Quality and Yield of Winter Oats (*Avena sativa* L.). *Foods* **2021**, *10*, 2356. [\[CrossRef\]](#)
10. Tao, Y.; Zhao, X.; Wang, X.; Hathorn, A.; Hunt, C.; Cruickshank, A.W.; van Oosterom, E.J.; Godwin, I.D.; Mace, E.S.; Jordan, D.R. Large-scale GWAS in sorghum reveals common genetic control of grain size among cereals. *Plant Biotechnol. J.* **2020**, *18*, 1093–1105. [\[CrossRef\]](#)
11. Yu, H.; Hao, Y.; Li, M.; Dong, L.; Che, N.; Wang, L.; Song, S.; Liu, Y.; Kong, L.; Shi, S. Genetic architecture and candidate gene identification for grain size in bread wheat by GWAS. *Front. Plant Sci.* **2022**, *13*, 1072904. [\[CrossRef\]](#) [\[PubMed\]](#)
12. Whan, A.P.; Smith, A.B.; Cavanagh, C.R.; Ral JP, F.; Shaw, L.M.; Howitt, C.A.; Bischof, L. GrainScan: A low cost, fast method for grain size and colour measurements. *Plant Methods* **2014**, *10*, 23. [\[CrossRef\]](#)
13. 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*, 1871–1880. [\[CrossRef\]](#)
14. Alemu, A.; Feyissa, T.; Tuberosa, R.; Maccaferri, M.; Sciara, G.; Letta, T.; Abeyo, B. Genome-wide association mapping for grain shape and color traits in Ethiopian durum wheat (*Triticum turgidum* ssp. durum). *Crop. J.* **2020**, *8*, 757–768. [\[CrossRef\]](#)
15. Kabange, N.R.; Dzorkpe, G.D.; Park, D.S.; Kwon, Y.; Lee, S.B.; Lee, S.M.; Lee, J.H. Rice (*Oryza sativa* L.) Grain Size, Shape, and Weight-Related QTLs Identified Using GWAS with Multiple GAPIT Models and High-Density SNP Chip DNA Markers. *Plants* **2023**, *12*, 4044. [\[CrossRef\]](#) [\[PubMed\]](#)
16. Jahnke, S.; Roussel, J.; Hombach, T.; Kochs, J.; Fischbach, A.; Huber, G.; Scharf, H. phenoSeeder—A Robot System for Automated Handling and Phenotyping of Individual Seeds. *Plant Physiol.* **2016**, *172*, 1358–1370. [\[CrossRef\]](#)
17. Hughes, A.; Oliveira, H.R.; Fradgley, N.; Corke, F.M.K.; Cockram, J.; Doonan, J.H.; Nibau, C.; Hughes, N. µCT trait analysis reveals morphometric differences between domesticated temperate small grain cereals and their wild relatives. *Plant J.* **2019**, *99*, 98–111. [\[CrossRef\]](#)
18. Legland, D.; Alvarado, C.; Badel, E.; Guillon, F.; King, A.; Le, T.D.Q.; Rivard, C.; Paré, L.; Chateigner-Boutin, A.-L.; Grousse, C. Synchrotron Based X-ray Microtomography Reveals Cellular Morphological Features of Developing Wheat Grain. *Appl. Sci.* **2022**, *12*, 3454. [\[CrossRef\]](#)
19. Corcoran, E.; Siles, L.; Kurup, S.; Ahnert, S. Automated extraction of pod phenotype data from micro-computed tomography. *Front. Plant Sci.* **2023**, *14*, 1120182. [\[CrossRef\]](#)
20. Qin, Z.; Zhang, Z.; Hua, X.; Yang, W.; Liang, X.; Zhai, R.; Huang, C. Cereal grain 3D point cloud analysis method for shape extraction and filled/unfilled grain identification based on structured light imaging. *Sci. Rep.* **2022**, *12*, 3145. [\[CrossRef\]](#)
21. Webster, F.H. Oats. In *Cereal Grain Quality*; Springer: Dordrecht, The Netherlands, 1996; pp. 179–203. [\[CrossRef\]](#)
22. Coffman, F.A. *Oat History: Identification and Classification*; Technical Bulletins 158127; United States Department of Agriculture: Washington, DC, USA, 1977. [\[CrossRef\]](#)
23. Strychar, R. Chapter 1: World Oat Production, Trade, and Usage. *OATS: Chemistry and Technology*; AACC International, Inc.: St. Paul, MN, USA, 2011; pp. 1–10. [\[CrossRef\]](#)
24. Heuschele, D.J.; Catalano, D.; Martinson, K.; Wiersma, J. Consumer Knowledge and Horse Preference for Different Colored Oats. *J. Equine Vet. Sci.* **2018**, *71*, 6–12. [\[CrossRef\]](#)
25. Zwer, P. Oats: Characteristics and quality requirements. In *Cereal Grains: Assessing and Managing Quality*; Woodhead Publishing: Cambridge, UK, 2010; pp. 163–182. [\[CrossRef\]](#)
26. Inspecting Grain for Defects and Impurities | AHDB. Available online: <https://ahdb.org.uk/knowledge-library/inspecting-grain-for-defects-and-impurities> (accessed on 10 June 2024).
27. Marshall, A.; Cowan, S.; Edwards, S.; Griffiths, I.; Howarth, C.; Langdon, T.; White, E. Crops that feed the world 9. Oats- a cereal crop for human and livestock feed with industrial applications. *Food Secur.* **2013**, *5*, 13–33. [\[CrossRef\]](#)

28. Valentine, J. *Chapter 2: Oat Breeding. OATS: Chemistry and Technology*; AACC International, Inc.: St. Paul, MN, USA, 2011; pp. 11–30. [\[CrossRef\]](#)
29. Leiva, F.; Zakieh, M.; Alamrani, M.; Dhakal, R.; Henriksson, T.; Singh, P.K.; Chawade, A. Phenotyping Fusarium head blight through seed morphology characteristics using RGB imaging. *Front. Plant Sci.* **2022**, *13*, 1010249. [\[CrossRef\]](#)
30. Wang, K.; Taylor, D.; Ruan, Y.; Pozniak, C.J.; Izydorczyk, M.; Fu, B.X. Unveiling the factors affecting milling quality of durum wheat: Influence of kernel physical properties, grain morphology and intrinsic milling behaviours. *J. Cereal Sci.* **2023**, *113*, 103755. [\[CrossRef\]](#)
31. Jammalamadaka, S.R.; SenGupta, A. *Topics in Circular Statistics*; World Scientific Pub Co Pte Ltd.: Singapore, 2001; Volume 5. [\[CrossRef\]](#)
32. Bland, J.M.; Altman, D.G. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet* **1986**, *327*, 307–310. [\[CrossRef\]](#)
33. Hyslop, N.P.; White, W.H. Estimating Precision Using Duplicate Measurements. *J. Air Waste Manag. Assoc.* **2009**, *59*, 1032–1039. [\[CrossRef\]](#)
34. Gabriel, K.R. The biplot graphic display of matrices with application to principal component analysis. *Biometrika* **1971**, *58*, 453–467. [\[CrossRef\]](#)
35. May, W.E.; Brandt, S.; Hutt-Taylor, K. Response of oat grain yield and quality to nitrogen fertilizer and fungicides. *Agron. J.* **2020**, *112*, 1021–1034. [\[CrossRef\]](#)
36. Buerstmayr, H.; Krenn, N.; Stephan, U.; Grausgruber, H.; Zechner, E. Agronomic performance and quality of oat (*Avena sativa* L.) genotypes of worldwide origin produced under Central European growing conditions. *Field Crop. Res.* **2007**, *101*, 343–351. [\[CrossRef\]](#)
37. Finnan, J.; Burke, B.; Spink, J. The effect of nitrogen timing and rate on radiation interception, grain yield and grain quality in autumn sown oats. *Field Crop. Res.* **2019**, *231*, 130–140. [\[CrossRef\]](#)
38. Doehlert, D.C. Quality Improvement in Oat. *J. Crop. Prod.* **2002**, *5*, 165–189. [\[CrossRef\]](#)
39. Doehlert, D.C.; McMullen, M.S.; Riveland, N.R. Sources of Variation in Oat Kernel Size. *Cereal Chem.* **2002**, *79*, 528–534. [\[CrossRef\]](#)
40. Doehlert, D.C.; Jannink, J.-L.; McMullen, M.S. Size Distributions of Different Orders of Kernels within the Oat Spikelet. *Crop. Sci.* **2008**, *48*, 298–304. [\[CrossRef\]](#)
41. Kim, J.; Savin, R.; Slafer, G.A. Weight of individual wheat grains estimated from high-throughput digital images of grain area. *Eur. J. Agron.* **2021**, *124*, 126237. [\[CrossRef\]](#)
42. Shah, A.; Masoodi, F.A.; Gani, A.; Ashwar, B.A. Geometrical, functional, thermal, and structural properties of oat varieties from temperate region of India. *J. Food Sci. Technol.* **2016**, *53*, 1856–1866. [\[CrossRef\]](#)
43. Lapveteläinen, A.; Alho-Lehto, P.; Sinn, L.; Laukkanen, T.; Lindman, T.; Kallio, H.; Kaitaranta, J.; Katajisto, J. Relationships of Selected Physical, Chemical, and Sensory Parameters in Oat Grain, Rolled Oats, and Cooked Oatmeal—A Three-Year Study with Eight Cultivars. *Cereal Chem.* **2001**, *78*, 322–329. [\[CrossRef\]](#)
44. Aprodu, I.; Banu, I. Milling, functional and thermo-mechanical properties of wheat, rye, triticale, barley and oat. *J. Cereal Sci.* **2017**, *77*, 42–48. [\[CrossRef\]](#)
45. Sadras, V.O. Evolutionary aspects of the trade-off between seed size and number in crops. *Field Crop. Res.* **2007**, *100*, 125–138. [\[CrossRef\]](#)
46. Peltonen-Sainio, P.; Kangas, A.; Salo, Y.; Jauhiainen, L. Grain number dominates grain weight in temperate cereal yield determination: Evidence based on 30 years of multi-location trials. *Field Crop. Res.* **2007**, *100*, 179–188. [\[CrossRef\]](#)
47. Vahamidis, P.; Stefopoulou, A.; Kotoulas, V.; Bresta, P.; Nikolopoulos, D.; Karabourniotis, G.; Mantonakis, G.; Vlachos, C.; Dercas, N.; Economou, G. Grain size variation in two-rowed malt barley under Mediterranean conditions: Phenotypic plasticity and relevant trade-offs. *Field Crop. Res.* **2022**, *279*, 108454. [\[CrossRef\]](#)
48. Kruszelnicka, W.; Chen, Z.; Ambrose, K. Moisture-Dependent Physical-Mechanical Properties of Maize, Rice, and Soybeans as Related to Handling and Processing. *Materials* **2022**, *15*, 8729. [\[CrossRef\]](#) [\[PubMed\]](#)
49. Gerald Winter Oat—Senova. Available online: <https://www.senova.uk.com/gerald> (accessed on 10 June 2024).
50. Mascani Winter Oat—Senova. Available online: <https://www.senova.uk.com/mascani> (accessed on 10 June 2024).
51. Komyshev, E.G.; Genaev, M.A.; Afonnikov, D.A. Analysis of color and texture characteristics of cereals on digital images. *Vavilov J. Genet. Breed.* **2020**, *24*, 340–347. [\[CrossRef\]](#) [\[PubMed\]](#)
52. Brabec, D.; Guttieri, M.; Pearson, T.; Carsrud, B. Effectiveness of an image-based sorter to select for kernel color within early segregating hard winter wheat (*Triticum aestivum* L.) Populations. *Cereal Res. Commun.* **2017**, *45*, 488–499. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.