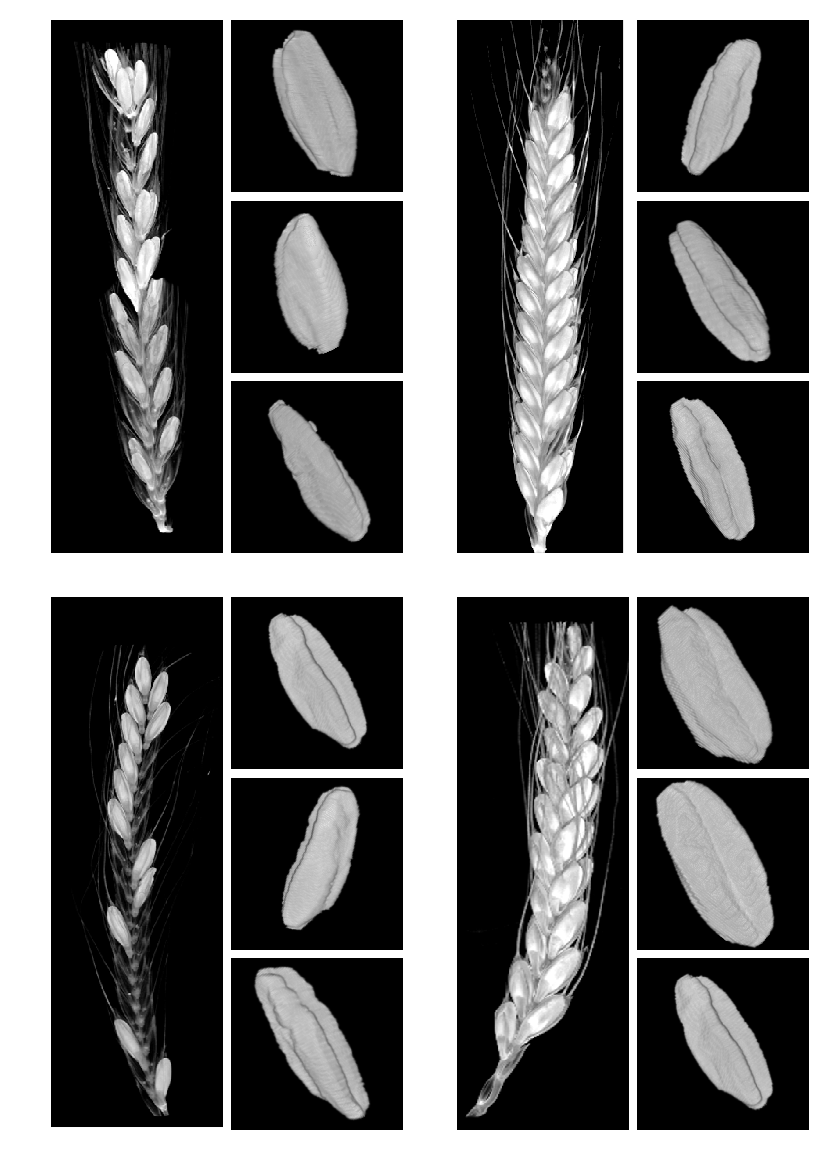
Domestication Draft

NH, CN, HO, JD

# Introduction

<Not sure on>

# Results



[fig:orgd8c7835] Einkorn (top) and Emmer (bottom) Wheat (Primitives left, Domesticated right)

Four key species, two diploid (Einkorn) and two tetraploid (Emmer) were selected to study the effect on grain morphology across domestication and polyploidisation. Domesticated *T. monococcum* and primitive *T. boeticum* were selected for being diploid; Domesticated *T. dicoccum* as primitive *T. dicoccoides* were chosen as representative species for tetraploids. (See Figure 1)

Key grain morphological traits were measured using micro-CT imaging, these include length, width, depth, volume, and surface area. From these traits, two additional descriptors were created LWD (Length X Width X Depth) and Surface area / Volume ratio these are used to provide insight to potential trait interactions.

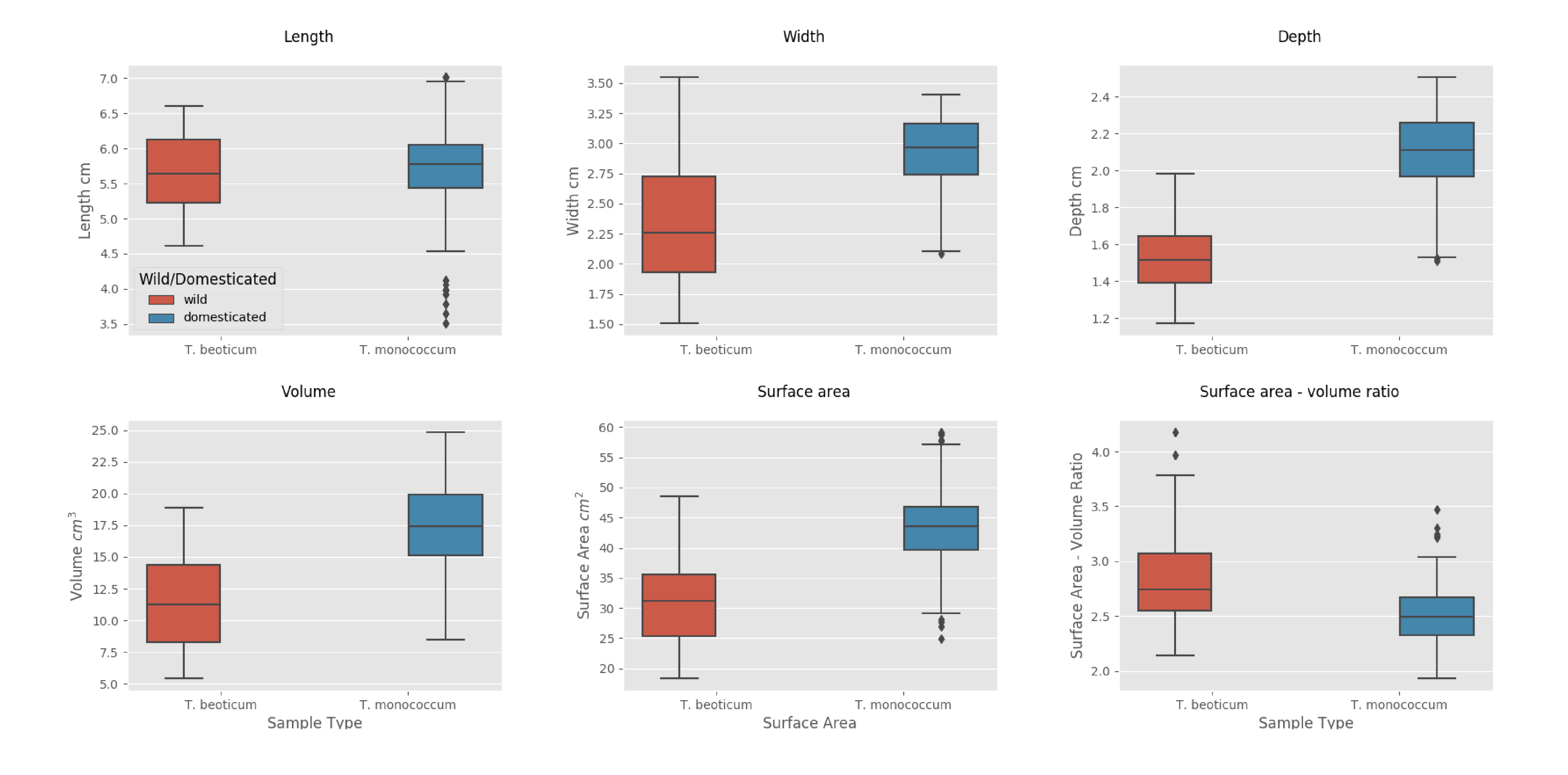
Traits were gathered using image analysis software (Hughes et al. 2017), major modifications were made to grain identification algorithms in order to adapt to the wide variety in the species studied.

Analysis was performed using a Welch T-test for significance and confidence intervals. Kruchke’s method was used for Bayesian estimation of difference (Kruschke 2013) in order to quantify differences in population means. Estimation of domestication status was achieved through multiple regression (using ordinary least squares methods).

Principle Component Analysis (PCA) was used to highlight separation of species based on grain measurements, in addition to the effect of specific traits in model though reported eigenvalues/coefficient loadings.

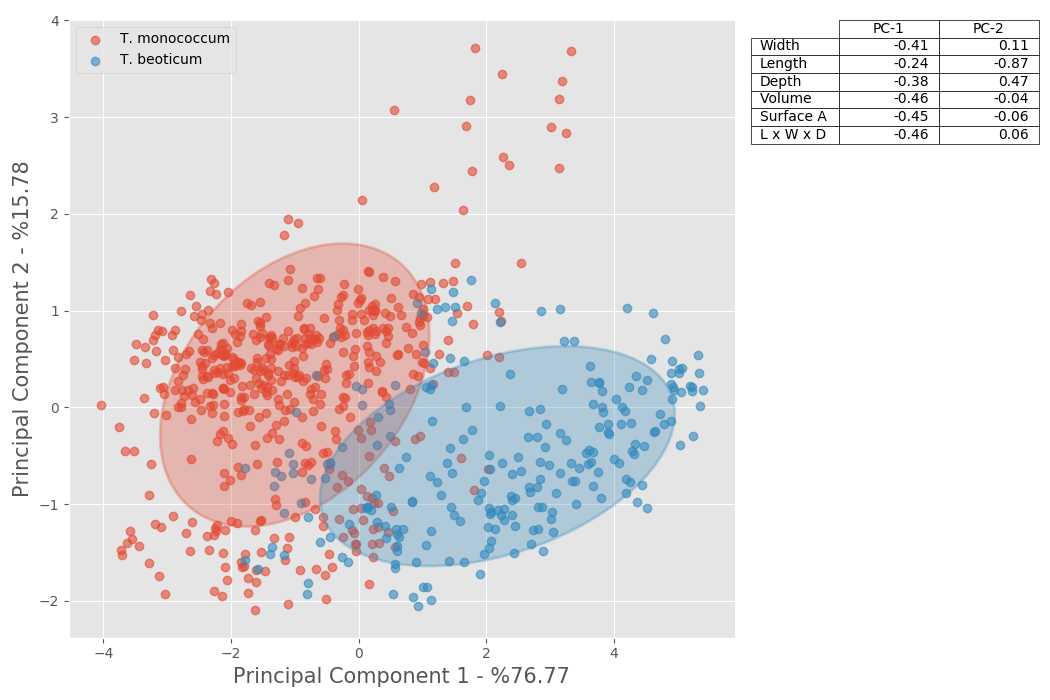
## Einkorn

No noticeable change has been observed in the length attribute between domesticated and wild einkorn ( and a predicted probability of overlapping averages). However, other traits(figure:2]: length, width, depth, volume and surface area all have a significant change between wild and domesticated ().



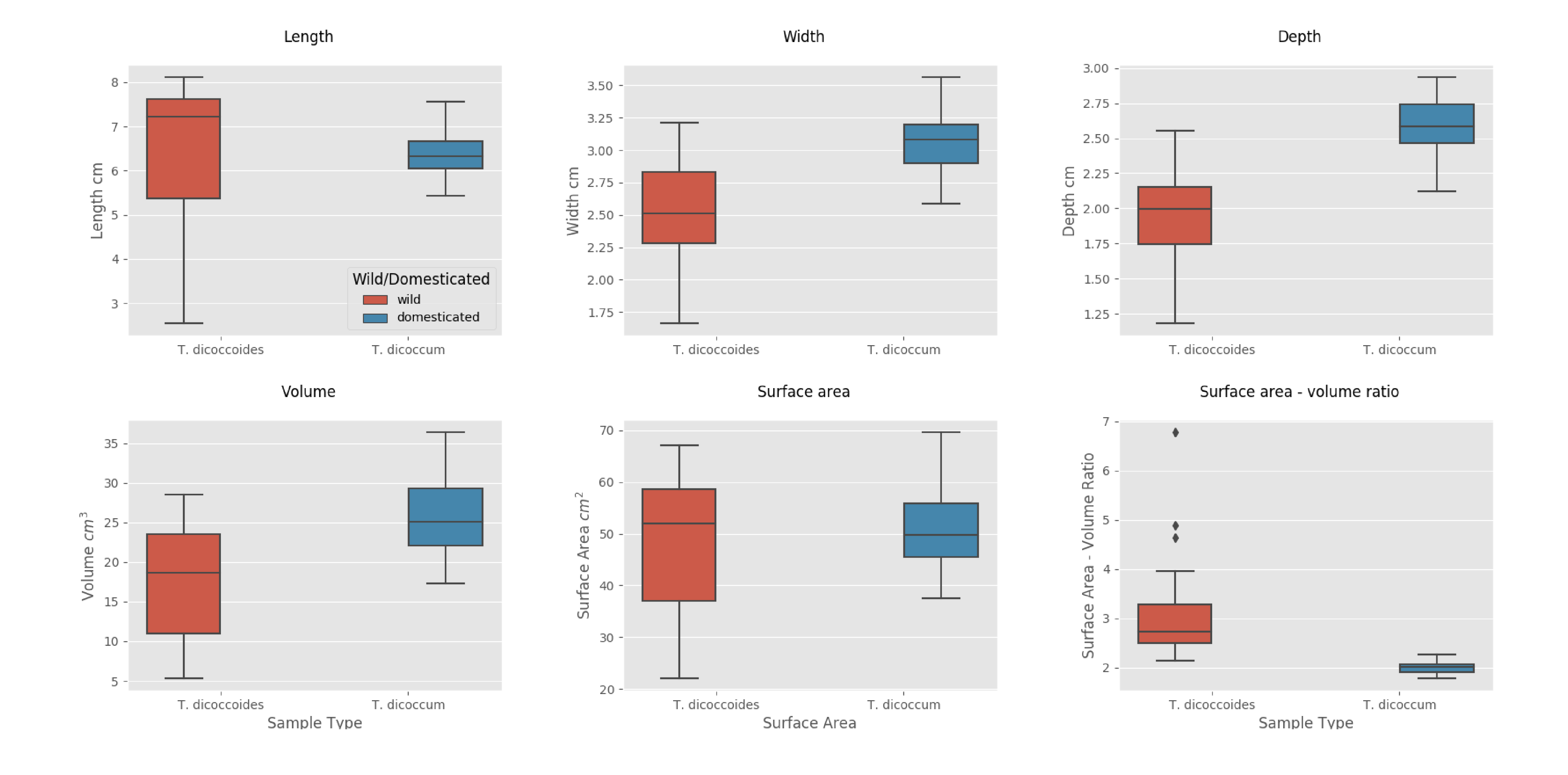
[fig:orgcea08ca] Einkorn Traits

A two component PCA reviled that no dominant trait appears to influence morphometric variation in einkorn species (figure:3). With the exception of grain length (-0.24), all other measurements (width, depth, volume, surface area and the interaction term L x W x D) have coefficients ranging -0.38 to -0.46 in the first principle component (76.77% of explained variation).

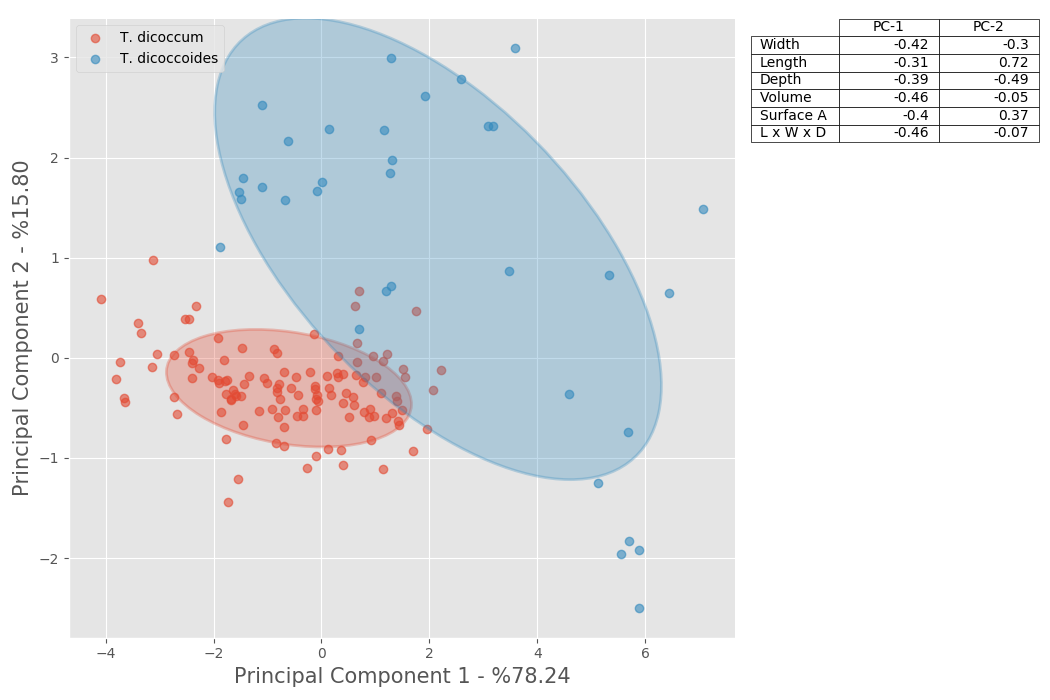


[fig:orgfcec768] Einkorn PCA

## Emmer



[fig:orgb5aab78] Emmer Traits



[fig:org169833e] Emmer PCA

# Methods

## Materials

< Plants which were used > …

### How they were sourced

### Any additional information

## 3D Scanning of Spikes

From the genotypes selected, fully dried, representative spikes were chosen for micro-CT scanning. Spikes were placed in plastic holders (34x70mm tubes) and imaged using a a μCT100 scanner (Scanco Medical, Switzerland).

The imaging system was configured with an X-ray source ranging from 20 to 100 kVp, a detector of 3072 x 400 elements. A resolution of 68.8 micro-meters per pixel was used for all scans.

## Computational Methods

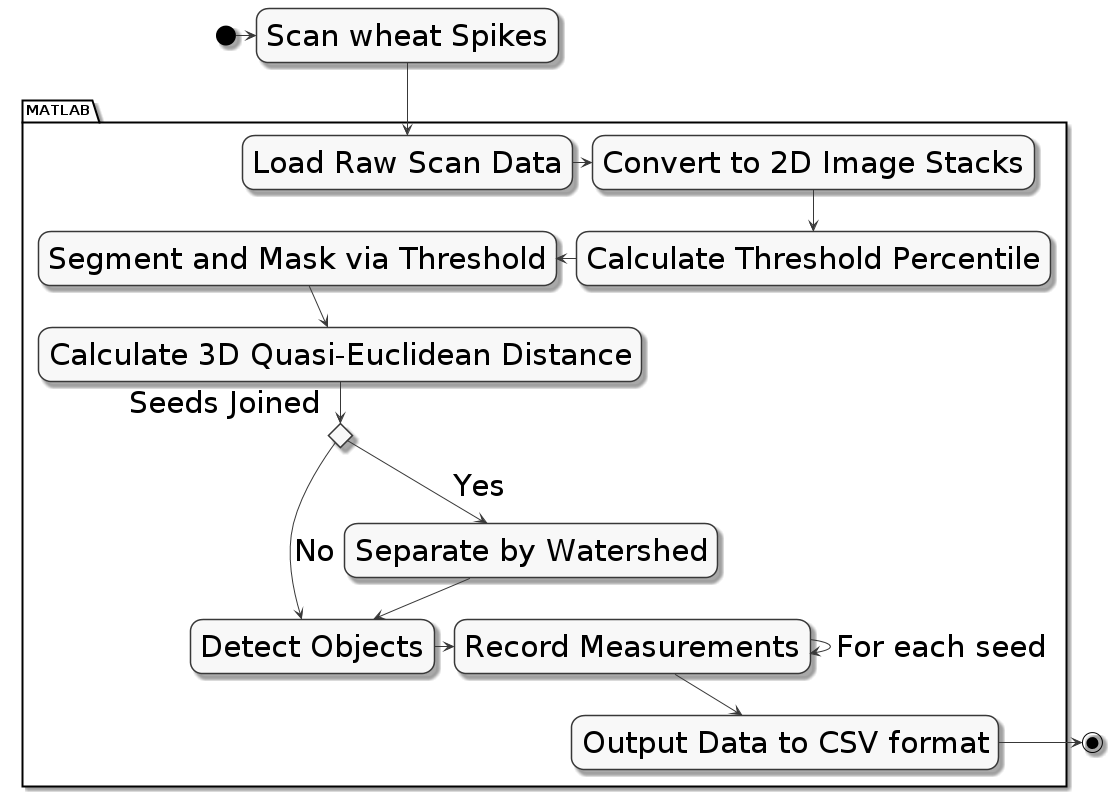
Using software developed for previous wheat studies by the National Plant Phenomics Centre ((Hughes et al. 2017)) . New and novel additions are implemented in the watershedding and segmentation processes, of the existing pipeline, in order to work with the more complex primitive species of wheat.

Due to the optimised resolution of the imaging technique (68.6 meters per pixel) objects can appear connected which are not, particularly in primitive grain. A three dimensional watershedding algorithm is used to correct any objects which appear connected when they should not be.

The software, developed in MATLAB ((The MathWorks Inc. 2018)), is freely available at <insert link to NPPC>.

#### Pipeline

The scanning and MATLAB routine pipeline:



[fig:org9ee67b0] Pipeline

### Morphometric Features

The features/phenotypes used are extracted during the imaging process.

* Length is calculated using the major axis of the whole grain.
* Width and depth are the major and minor axis of a cross section, found by selecting the grain’s midpoint.
* Volume is a complete connected pixel count per grain
* Surface area is a single pixel perimeter calculation mapped in 3 dimensions
* Length X Depth X Width is a post-image-processing value calculated by the interaction between the three dimension descriptors.

Values used in statistical functions and measurements are presented as metric units, derived from -CT image pixel values. The equation:[eq:org5f53ccd] is presented here.

### Error Removal

The data were checked for false positives, this is done by first removing outliers which are found by the 0.025 upper and lower percentiles of the data. Additionally constraints are applied to the data based on findings from previous studies (Hughes et al. 2017), this adds robustness.

### LWD

An additional phenotype is created to describe the interaction between the geometric parameters; the interaction is described in equation:[eq:orgc714724].

### Image Analysis Methods

***I could provide a lot of info on this, but weary of going off-track, more can be added (or taken away) if needed, what’s in comp methods could be enough I think.***

## Bayesian Modelling

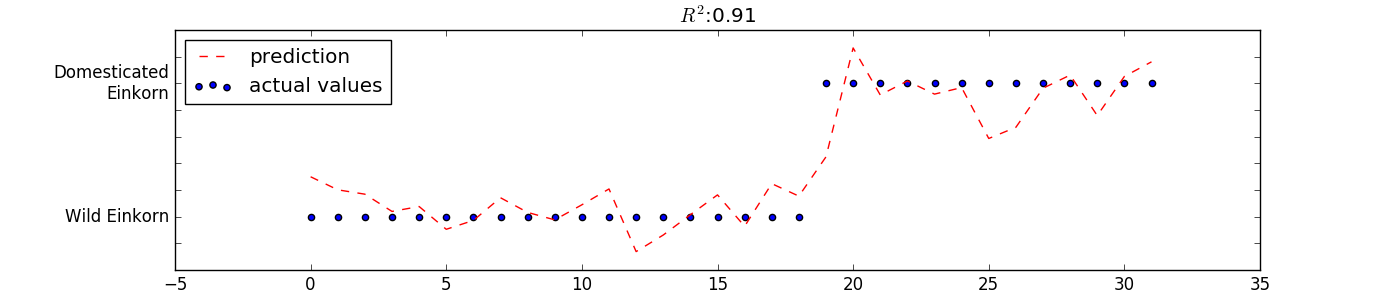
To provide deeper insight into the size of change or similarity in hypothesis testing, a Bayesian model is used. To estimate probability of two samples containing the same mean the method uses Bayes theorem () (Kruschke 2013) along with Markov-Chain-Monte-Carlo (MCMC) to draw random samples from a normal population.

From Krusches’ method a percentage likelihood of difference is produced.

## Linear Modelling

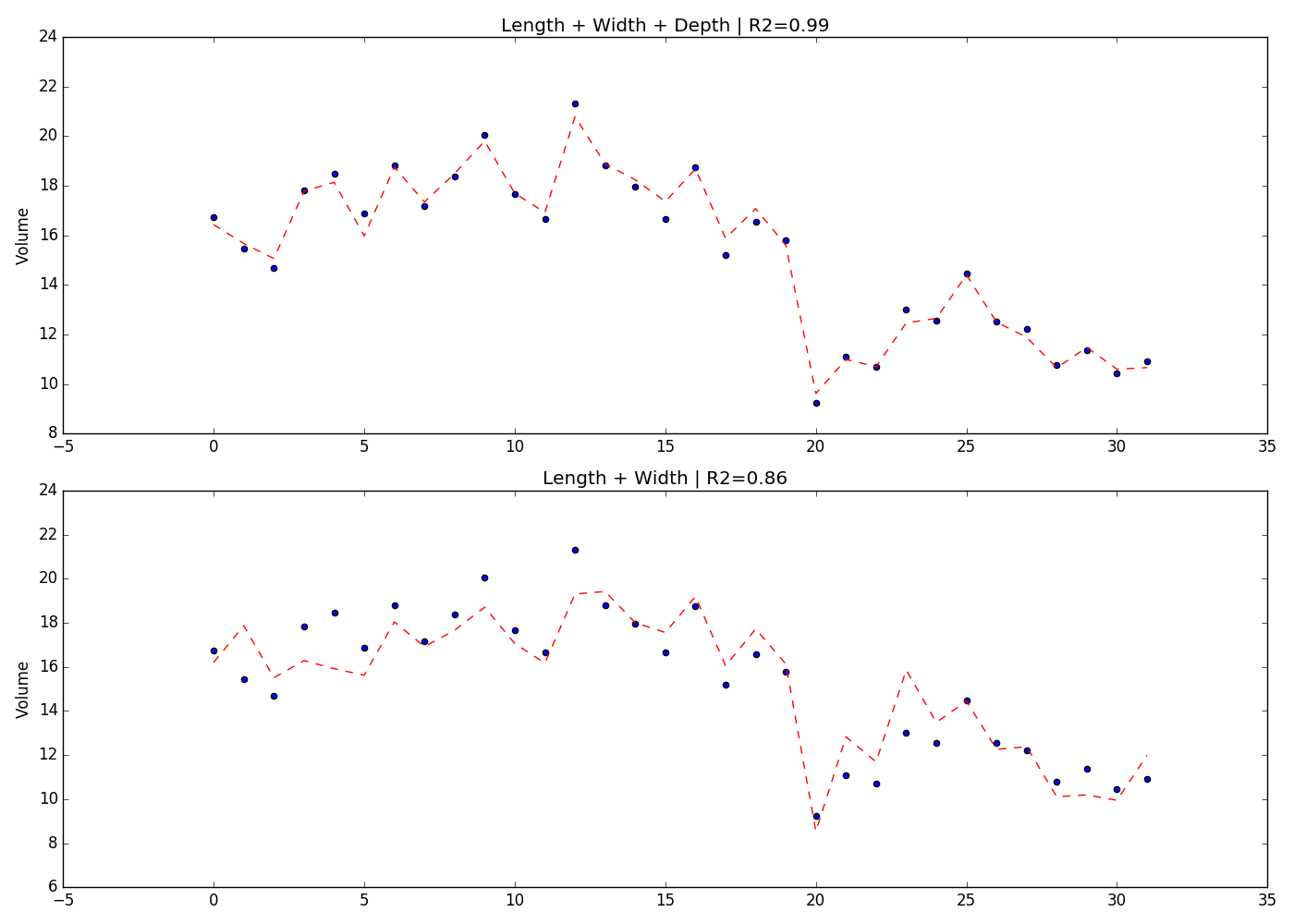
A linear model allowed for an R value of 0.91 in einkorn species when predicting domestication status.

### Draft Supplemental figure for dom



[fig:org5802f8b] is showing a multiple regression of r=.91 by using length, width and depth to correctly ID domestication status.

### Draft Supplemental figure for volume



[fig:org631db65] Showing the importance of 3rd dimension of depth

### Least Squares Model

Equation, where the intercept () has been set by surface area

### Misc

The model produced an r value by:

# Misc Information to fold into discussion?

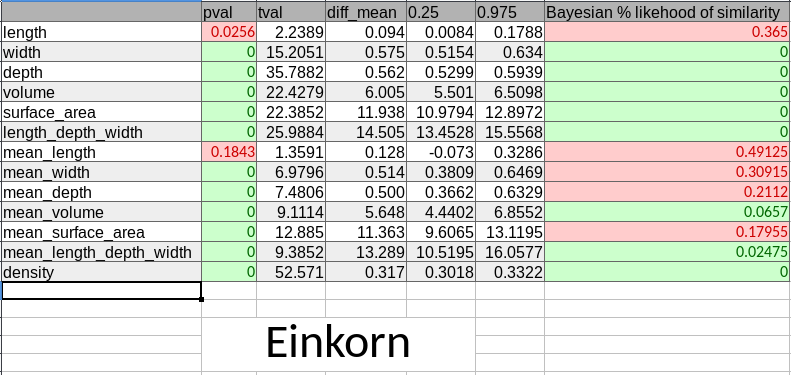
Fuller has evidenced that grain volume is used when initially identifying wheat grain which is recovered as being domesticated (Fuller 2007) . Therefore justification for our model is very useful.

On the other hand (Willcox 2004) found that barley is much harder to identify from it’s domesticated relatives. Volume is not significantly different! Which fits perfectly with our results!

Wondering outloud, with the wild 2n/4n an oberserved change in surface area. Would there be a reason for why grains would want to develop a larger surface area? Alternatively we can work off-of volume, as that changes also

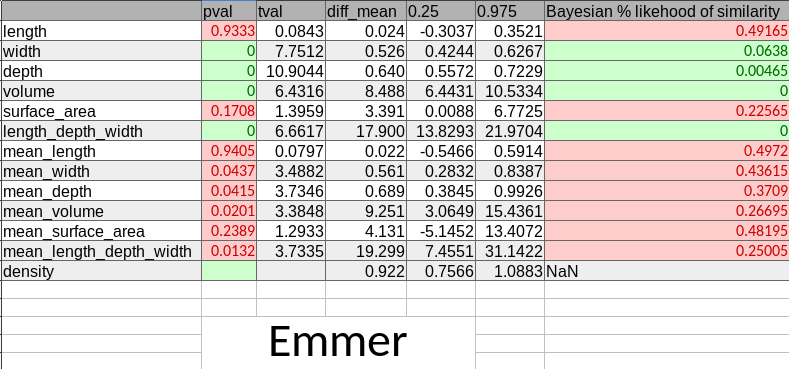
# Data Tables

## einkorn



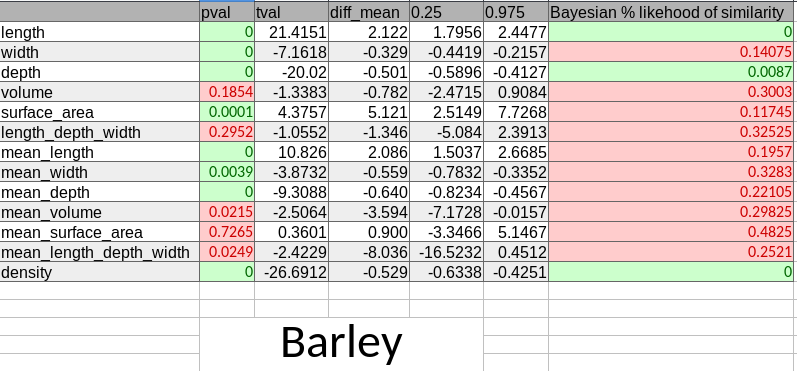
[fig:orge6342ec] einkorn table

## emmer



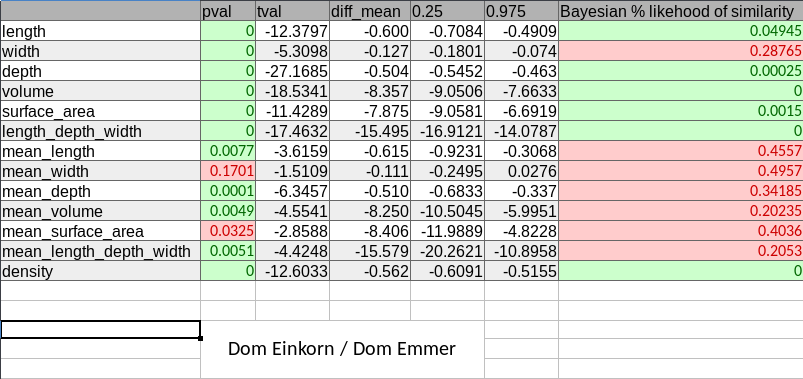
[fig:org35b3731] emmer table

## Barley



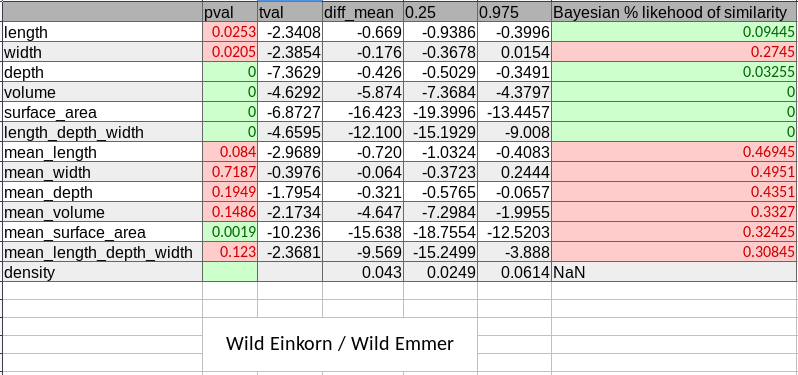
[fig:orgd55e90b] barley table

## Domesticated 2N, 4N



[fig:org4965773] domesticated 2N, 4N table

## Wild 2N, 4N



[fig:org46dcea7] wild 2N, 4N table

Fuller, D. Q. 2007. “Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World.” *Annals of Botany* 100 (5). Oxford University Press: 903–24. doi:[10.1093/aob/mcm048](https://doi.org/10.1093/aob/mcm048).

Hughes, Nathan, Karen Askew, Callum P. Scotson, Kevin Williams, Colin Sauze, Fiona Corke, John H. Doonan, and Candida Nibau. 2017. “Non-Destructive, high-Content analysis of wheat grain traits using X-ray micro computed tomography.” *Plant Methods* 13 (1). doi:[10.1186/s13007-017-0229-8](https://doi.org/10.1186/s13007-017-0229-8).

Kruschke, John. 2013. “Bayesian estimation Supersedes the t-test.” *Journal of Experimental Psychology: General* 142 (2): 573–603. doi:[10.1037/a0029146](https://doi.org/10.1037/a0029146).

The MathWorks Inc. 2018. “MathWorks - Support - MATLAB & Simulink.” <https://es.mathworks.com/support.html?s{\_}tid=gn{\_}supp>.

Willcox, George. 2004. “Measuring grain size and identifying Near Eastern cereal domestication: evidence from the Euphrates valley.” *Journal of Archaeological Science* 31 (2). Academic Press: 145–50. doi:[10.1016/J.JAS.2003.07.003](https://doi.org/10.1016/J.JAS.2003.07.003).