

Domestication Draft

NH, CN, HO, JD

August 6, 2018

1 Introduction

<Not sure on>

2 Results

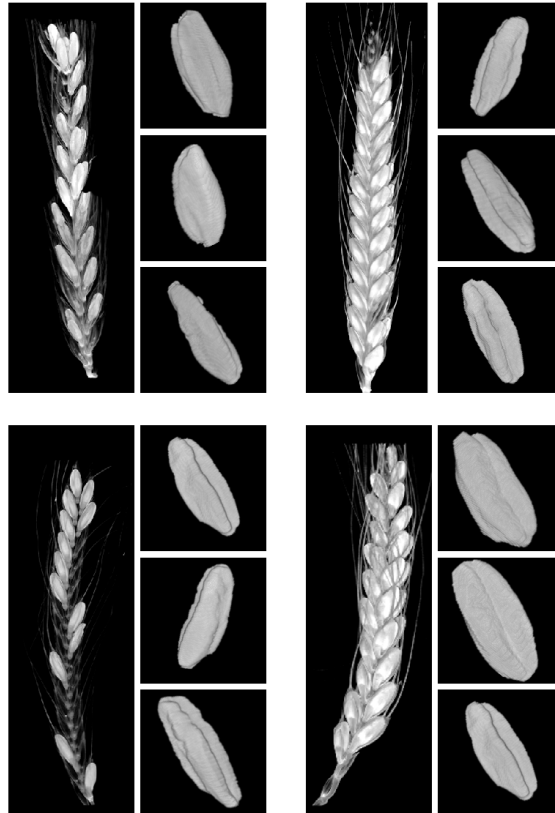


Figure 1: Wheats

2.1 Domestication

Previous studies have shown that domestication has led to an increase in grain size ([1, 2]), though the precise changes were not identified. Here we have shown that both emmer and einkorn species have both increased significantly in grain width and depth, with length remaining similar (our analysis has shown a 36% in einkorn and 49% probability in emmer that the mean values of grain length overlap over domestication state).

2.2 Ploidy

2.2.1 Domestication/Wild

Across ploidy in domesticated einkorn and emmer, 2N and 4N respectively, grain parameters are generally significantly different. The one parameter which we report almost no significant change in is grain width. A $p = 0.17$ in difference in sample mean width was calculated using the Welch T-test. A probability of mean overlap of 50% was observed for width at an individual level.

Citations needed to verify?

In wild species of einkorn and emmer no significant change is seen in individual grain length or width; Across population means non-significance is observed in grain length and width (all of which have *likelihood* > 33% of overlapping means and $p > 0.01$).

Significant changes are seen in depth, volume and surface area of grains ($p < 0.01$), indicating that the change in depth effected overall grain shape.

Citations needed to verify?

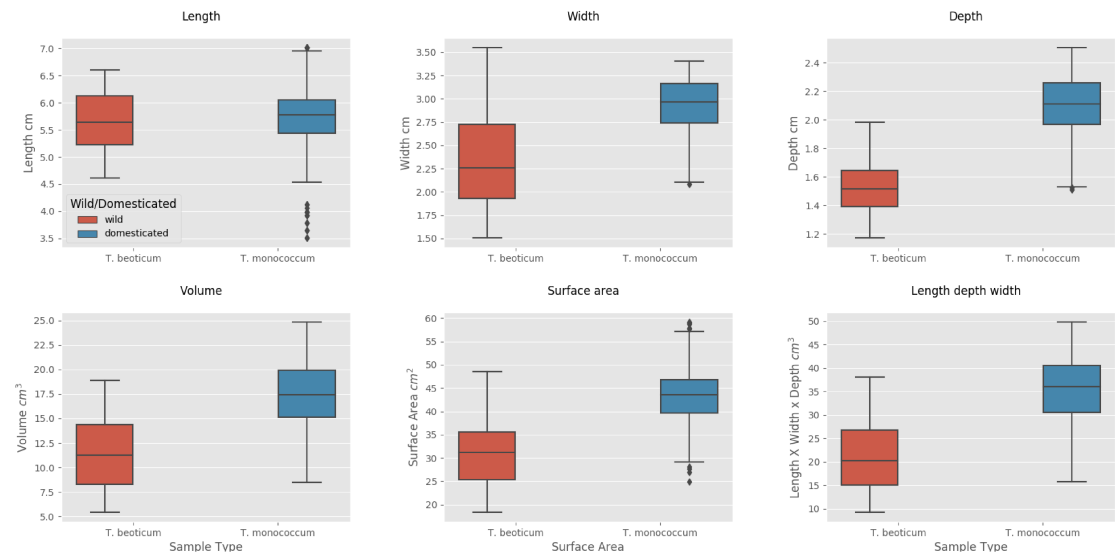


Figure 2: Einkorn Traits

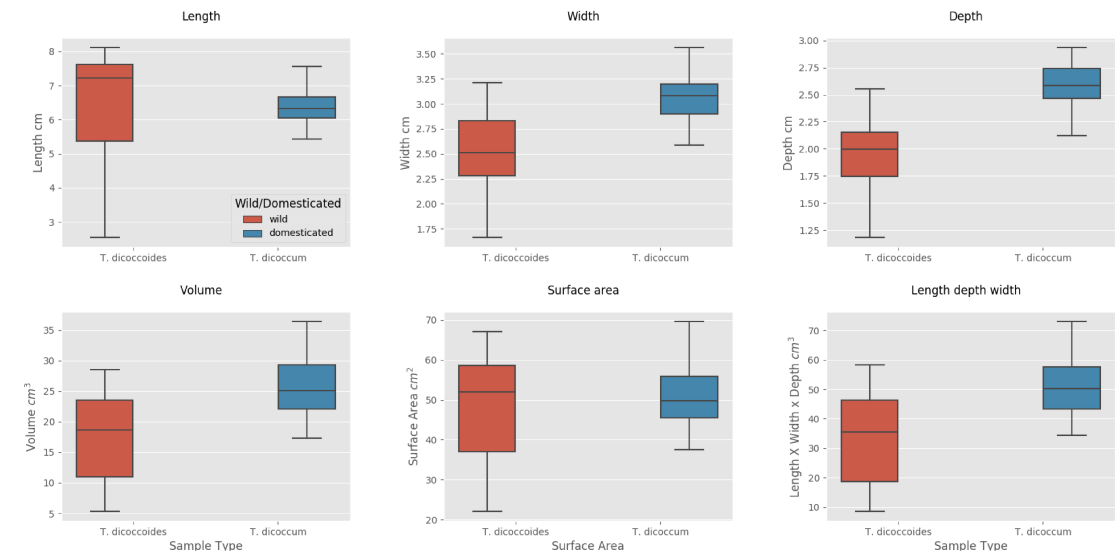


Figure 3: Emmer Traits

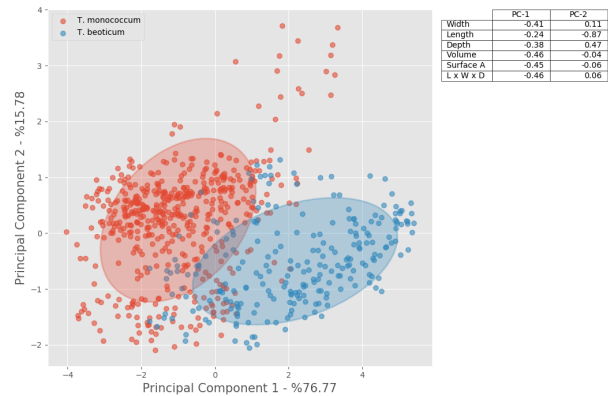


Figure 4: Einkorn PCA

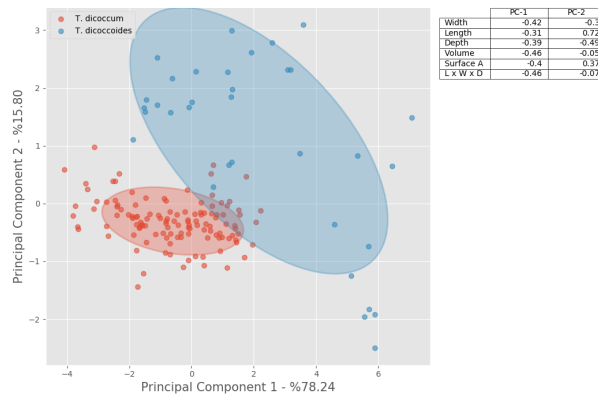


Figure 5: Emmer PCA

3 Methods

3.1 Materials

< Plants which were used > ...

3.1.1 How they were sourced

3.1.2 Any additional information

3.2 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 CT100 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.

3.3 Computational Methods

Using software developed for previous wheat studies by the National Plant Phenomics Centre ([3]). 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 (68.6µ 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 ([4]), is freely available at <insert link to NPPC>.

3.3.0.1 Pipeline The scanning and MATLAB routine pipeline:

3.3.1 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

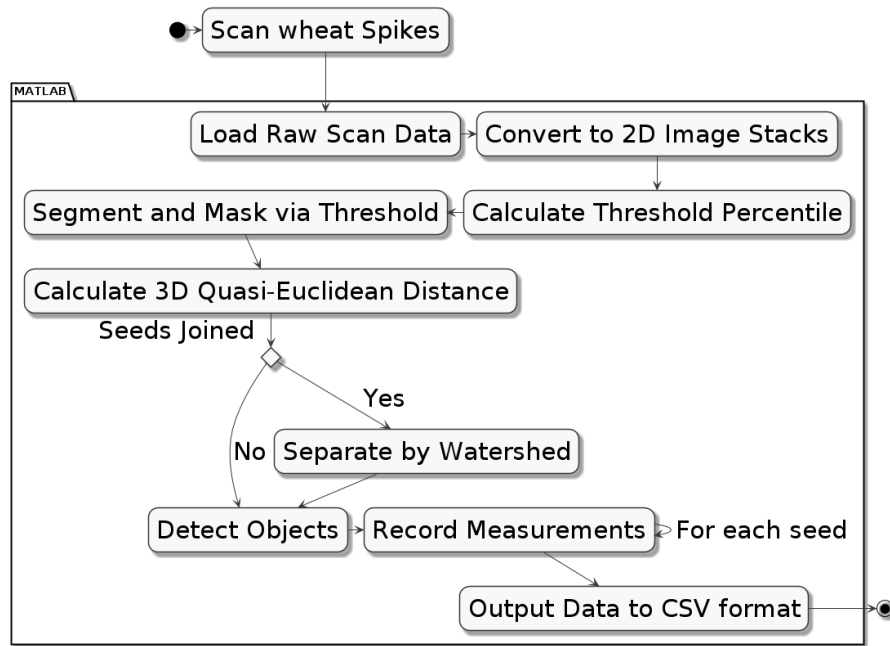


Figure 6: Pipeline

- 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 μ -CT image pixel values. The equation:1 is presented here.

$$mm = \frac{pixel \times 68.8}{1000} \quad (1)$$

3.3.2 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 [3], this adds robustness.

3.3.3 LWD

An additional phenotype is created to describe the interaction between the geometric parameters; the interaction is described in equation:2.

$$\text{geometry interaction} = length \times depth \times width \quad (2)$$

3.3.4 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.

3.4 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 ($P(A|B) \propto P(B|A) \times P(A)$) [5] 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.

3.5 Linear Modelling

A linear model allowed for an R^2 value of 0.91 in einkorn species when predicting domestication status.

3.5.1 Draft Supplemental figure for dom

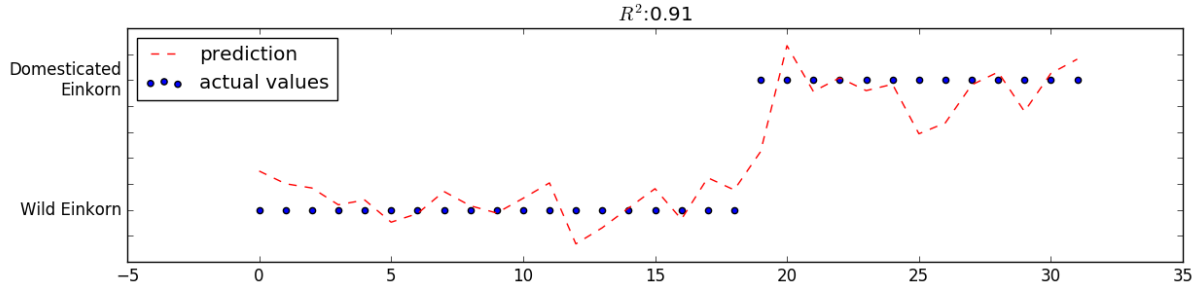


Figure 7: is showing a multiple regression of $r=0.91$ by using length, width and depth to correctly ID domestication status.

3.5.2 Draft Supplemental figure for volume

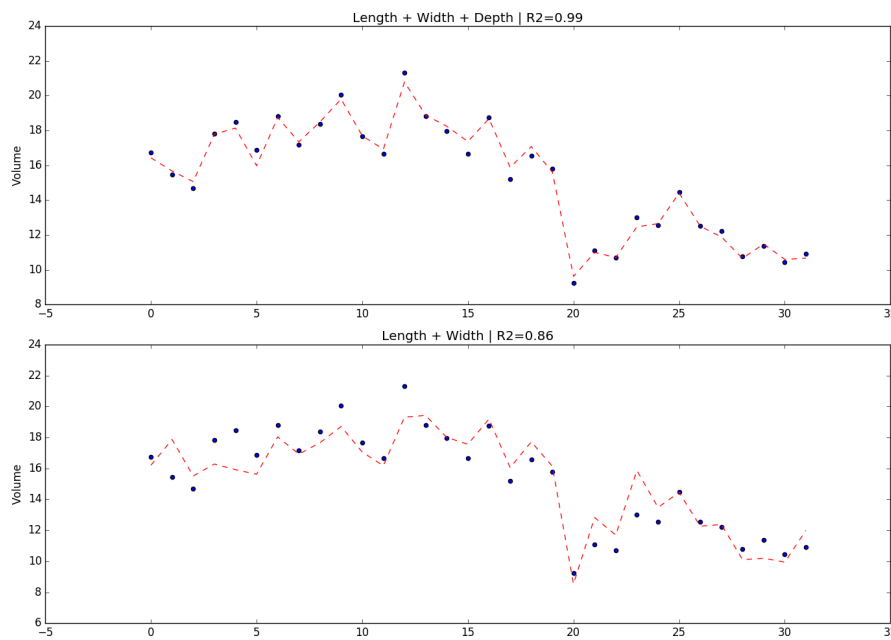


Figure 8: Showing the importance of 3rd dimension of depth

3.5.3 Least Squares Model

Equation, where the intercept (β_0) has been set by surface area

$$Y = \beta_0 + \beta_1 \text{length} + \beta_2 \text{depth} + \beta_3 \text{width} + \epsilon$$

3.5.4 Misc

The model produced an r^2 value by:

$$1 - \frac{\sum_{i=0}^n r^2}{\sum_{i=0}^n (y_i - \bar{y})^2}$$

4 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 [6] . Therefore justification for our model is very useful.

On the other hand [7] 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 observed 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

5 Data Tables

5.1 einkorn

	pval	tval	diff_mean	0.25	0.975	Bayesian % likelihood of similarity
length	0.0256	2.2389	0.094	0.0084	0.1788	0.365
width	0	15.2051	0.575	0.5154	0.634	0
depth	0	35.7882	0.562	0.5299	0.5939	0
volume	0	22.4279	6.005	5.501	6.5098	0
surface_area	0	22.3852	11.938	10.9794	12.8972	0
length_depth_width	0	25.9884	14.505	13.4528	15.5568	0
mean_length	0.1843	1.3591	0.128	-0.073	0.3286	0.49125
mean_width	0	6.9796	0.514	0.3809	0.6469	0.30915
mean_depth	0	7.4806	0.500	0.3662	0.6329	0.2112
mean_volume	0	9.1114	5.648	4.4402	6.8552	0.0657
mean_surface_area	0	12.885	11.363	9.6065	13.1195	0.17955
mean_length_depth_width	0	9.3852	13.289	10.5195	16.0577	0.02475
density	0	52.571	0.317	0.3018	0.3322	0
Einkorn						

Figure 9: einkorn table

5.2 emmer

	pval	tval	diff_mean	0.25	0.975	Bayesian % likelihood of similarity
length	0.9333	0.0843	0.024	-0.3037	0.3521	0.49165
width	0	7.7512	0.526	0.4244	0.6267	0.0638
depth	0	10.9044	0.640	0.5572	0.7229	0.00465
volume	0	6.4316	8.488	6.4431	10.5334	0
surface_area	0.1708	1.3959	3.391	0.0088	6.7725	0.22565
length_depth_width	0	6.6617	17.900	13.8293	21.9704	0
mean_length	0.9405	0.0797	0.022	-0.5466	0.5914	0.4972
mean_width	0.0437	3.4882	0.561	0.2832	0.8387	0.43615
mean_depth	0.0415	3.7346	0.689	0.3845	0.9926	0.3709
mean_volume	0.0201	3.3848	9.251	3.0649	15.4361	0.26695
mean_surface_area	0.2389	1.2933	4.131	-5.1452	13.4072	0.48195
mean_length_depth_width	0.0132	3.7335	19.299	7.4551	31.1422	0.25005
density			0.922	0.7566	1.0883	NaN
Emmer						

Figure 10: emmer table

5.3 Barley

	pval	tval	diff_mean	0.25	0.975	Bayesian % likelihood of similarity
length	0	21.4151	2.122	1.7956	2.4477	0
width	0	-7.1618	-0.329	-0.4419	-0.2157	0.14075
depth	0	-20.02	-0.501	-0.5896	-0.4127	0.0087
volume	0.1854	-1.3383	-0.782	-2.4715	0.9084	0.3003
surface_area	0.0001	4.3757	5.121	2.5149	7.7268	0.11745
length_depth_width	0.2952	-1.0552	-1.346	-5.084	2.3913	0.32525
mean_length	0	10.826	2.086	1.5037	2.6685	0.1957
mean_width	0.0039	-3.8732	-0.559	-0.7832	-0.3352	0.3283
mean_depth	0	-9.3088	-0.640	-0.8234	-0.4567	0.22105
mean_volume	0.0215	-2.5064	-3.594	-7.1728	-0.0157	0.29825
mean_surface_area	0.7265	0.3601	0.900	-3.3466	5.1467	0.4825
mean_length_depth_width	0.0249	-2.4229	-8.036	-16.5232	0.4512	0.2521
density	0	-26.6912	-0.529	-0.6338	-0.4251	0
Barley						

Figure 11: barley table

5.4 Domesticated 2N, 4N

	pval	tval	diff_mean	0.25	0.975	Bayesian % likelihood of similarity
length	0	-12.3797	-0.600	-0.7084	-0.4909	0.04945
width	0	-5.3098	-0.127	-0.1801	-0.074	0.28765
depth	0	-27.1685	-0.504	-0.5452	-0.463	0.00025
volume	0	-18.5341	-8.357	-9.0506	-7.6633	0
surface_area	0	-11.4289	-7.875	-9.0581	-6.6919	0.0015
length_depth_width	0	-17.4632	-15.495	-16.9121	-14.0787	0
mean_length	0.0077	-3.6159	-0.615	-0.9231	-0.3068	0.4557
mean_width	0.1701	-1.5109	-0.111	-0.2495	0.0276	0.4957
mean_depth	0.0001	-6.3457	-0.510	-0.6833	-0.337	0.34185
mean_volume	0.0049	-4.5541	-8.250	-10.5045	-5.9951	0.20235
mean_surface_area	0.0325	-2.8588	-8.406	-11.9889	-4.8228	0.4036
mean_length_depth_width	0.0051	-4.4248	-15.579	-20.2621	-10.8958	0.2053
density	0	-12.6033	-0.562	-0.6091	-0.5155	0
Dom Einkorn / Dom Emmer						

Figure 12: domesticated 2N, 4N table

References

- [1] Vasilis C. Gegas, Aida Nazari, Simon Griffiths, James Simmonds, Lesley Fish, Simon Orford, Liz Sayers, John H. Doonan, and John W. Snape. A Genetic Framework for Grain Size and Shape Variation in Wheat. *The Plant Cell*, 22(4):1046–1056, 2010.
- [2] Junhua H. Peng, Dongfa Sun, and Eviatar Nevo. Domestication evolution, genetics and genomics in wheat. *Molecular Breeding*, 28(3):281–301, oct 2011.
- [3] Nathan Hughes, Karen Askew, Callum P. Scotson, Kevin Williams, Colin Sauze, Fiona Corke, John H. Doonan, and Candida Nibau. Non-Destructive, high-Content analysis of wheat grain traits using X-ray micro computed tomography. *Plant Methods*, 13(1), 2017.
- [4] The MathWorks Inc. MathWorks - Support - MATLAB & Simulink, 2018.
- [5] John Kruschke. Bayesian estimation Supersedes the t-test. *Journal of Experimental Psychology: General*, 142(2):573–603, 2013.
- [6] D. Q. Fuller. Contrasting Patterns in Crop Domestication and Domestication Rates: Recent Archaeobotanical Insights from the Old World. *Annals of Botany*, 100(5):903–924, jul 2007.
- [7] George Willcox. Measuring grain size and identifying Near Eastern cereal domestication: evidence from the Euphrates valley. *Journal of Archaeological Science*, 31(2):145–150, feb 2004.