Phenotypic analysis of a European *Camelina sativa* diversity panel

Seth Polydore[a] and Noah Fahlgren[a]

[a]Donald Danforth Plant Science Center, 975 N Warson Road, St. Louis, MO 63132, USA

ORCID ID: 0000-0002-7779-7367 (S.P.)

ABSTRACT

Increased industrialization and pollution necessitate the development of new energy sources while minimizing the impact on the environment. *Camelina sativa* (L.) Crantz is being investigated as a potential source of novel biofuels, especially as a source of isoparaffin-rich jet fuel. However, currently, there is a paucity of phenotypic and genotypic data for *C. sativa*. Using a collection of 236 *C. sativa* genotypes from Eurasia, we assessed the phenotypic diversity of the collection using image data from the Bellwether Foundation Phenotyping Facility at the Donald Danforth Plant Science Center. Traits were measured using PlantCV. Based on phenotypic data, we found that most of the *C. sativa* accessions fell into 15 groups. There is strong phenotypic variance within each of these groups. Our study provides insight into defining potential *C. sativa* ideotypes.

1. INTRODUCTION

Increased industrialization and urbanization necessitate the development of new energy sources while minimizing the impact on the environment. Currently, fossil fuels are by far the most utilized energy source in the world[1]. However, burning fossil fuels releases many pollutants, including CO_2, and ultimately raises the amount of CO_2 in the atmosphere[2]. Increasing CO_2 has been directly linked to more powerful extreme weather events, such as hurricanes[3,4], shifting weather patterns leading to droughts in some areas[5] and floods in others[6–8]. Archer et al. (2020) estimated the "ultimate cost" (the cost of carbon to the global economy when all downstream effects are accounted for) of carbon to be roughly $100,000 per ton of carbon released[9].

New energy sources are needed to attenuate the rate of released carbon. One such source is biofuels from feedstocks such as *Camelina sativa* (L.) Crantz. *C. sativa* has received interest in recent years due to its potential as a low-input biofuel crop[10]. *C. sativa*'s seeds are nearly 30% oil and 50% polyunsaturated fatty acids by weight[11]; these oils can be used as a drop-in alternative fuel source[12]. For example, a strain of isoparaffin-rich *C. sativa* oil was commercial flight-tested by Japan Air-lines and KLM Royal Dutch Airlines. These test flights concluded that *C. sativa*'s oil can reduce emissions by as much as 84%[10].
C. sativa was widely cultivated throughout Europe for its oils that were used heavily in cooking and as an ingredient in paints[13–15]. However, in the 18th century, *C. sativa* was supplanted by *Brassica napus* (L.) as a major oil crop. As such, there is a paucity of genetic and phenotypic data that severely hampers the development of *C. sativa* as a societally important crop. To develop foundational tools to further the development of *C. sativa* as an industrial crop, we previously released a genetic resource for evaluating genetic variation among *C. sativa* varieties[16]. Here, we complement the genetic resources with a phenotypic analysis of the same population. We took roughly 100,000 images over the course of six weeks using the Bellwether Foundation Phenotyping Facility at Donald Danforth Plant Science Center[17] of these same accessions. These data can be used in the development of *C. sativa* cultivars.

2.   MATERIALS AND METHODS

A diversity panel representing the European *C. sativa* accessions were largely retrieved from the collections in the Leibniz Institute of Plant Genetics and Crop Plant Research and the U.S National Plant Germplasm System. These accessions were grown in the Bellwether Foundation Phenotyping Facility at the Donald Danforth Plant Science Center[17]. Images were taken daily for roughly 5 - 6 weeks. The accessions were phenotyped using PlantCV v3.9.0[18]. A single top-view and four side-view pictures were taken per time point. In the side-view, the plants were turned horizontally in 90° intervals between frames. In the following analyses, only the first side-view picture ("frame 0") was used for this project. Principal component analysis was done using the prcomp function in R v3.2.2[19] on the following measurements at the final time point: whole plant areas, convex hull areas, height, width, floral area, number of flowers, hue circular mean, and solidity. Uniform Manifold Approximation and Projection (UMAP) analysis[20] was then done on these principal components using the UMAP R package v0.2.6. The Density Based Clustering of Applications with Noise (DBSCAN) algorithm[21] was then used to group accessions based on the UMAP results using the DBSCAN R package v1.1.5. The optimal maximum distance between clusters (eps) was elucidated via the kNNdistplot function.
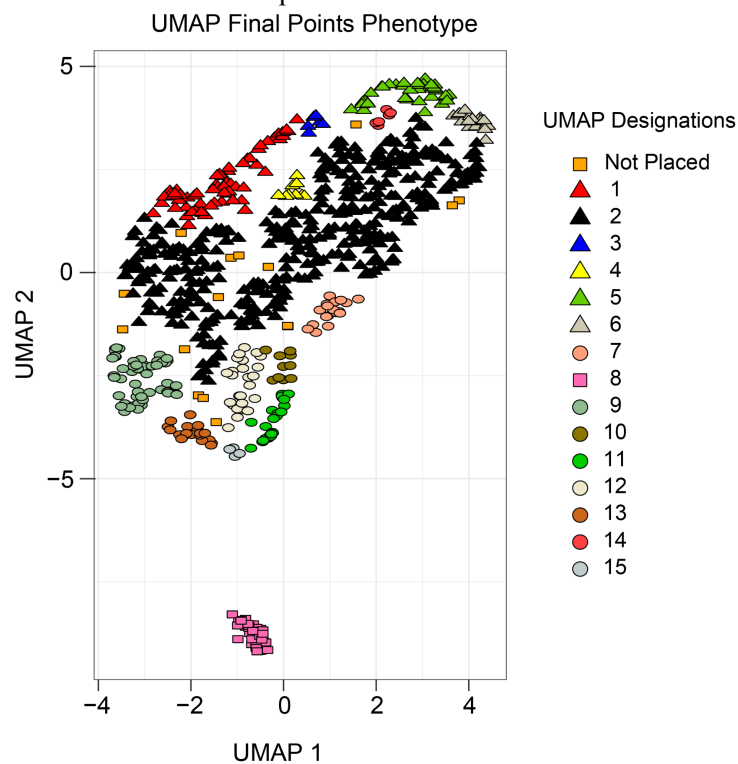
Figure 1. Uniform Manifold Approximation and Projection (UMAP) on final time point phenotypes. Colors and shapes represent the different UMAP designations.

3.   RESULTS AND DISCUSSIONS

3.1 Most accessions fall into one UMAP designation

Image data was analyzed with PlantCV, resulting in phenotypic data for total shoot area, convex hull area, solidity, maximum plant width, plant height, total leaf area, total flower area, hue circular mean, and the

number of flowers. We performed UMAP analysis on the plant's phenotypic data. Based on this analysis, we found that the *C. sativa* accessions could be largely sub-divided into 15 different groups (including 16 plants that couldn't be categorized). By far, group 2 had the most plants (**Figure 1**).

3.2 UMAP groups show clear phenotypic variance from each other

Utilizing PlantCV[18], we measured total plant area (**Fig 2a**), width (**Fig 2b**), area of leaves (**Fig 2c**) and flowers (**Fig 2d**), number of flowers (**Fig 2e**), solidity (total plant area divided by convex hull area) (**Fig 2f**), hue circular mean (**Fig 2g**), convex hull area (**Fig 2h**), and height (**Fig 2i**). Plants belonging to UMAP groups 8, 11, and 15 were clearly shorter (**Fig 2i**) and had less shoot and leaf area (**Fig 2a and c**) and convex hull area (**Fig 2h**) compared to the other categories. Group 8 had the greatest solidity; they were also the only group whose solidity increased after the 30th day post planting, suggesting these plants had the densest shoots (**Fig 2f**). Group 15 had the highest hue circular mean (**Fig 2g**) at the final time point, indicating that these plants were the darkest green. No floral tissues were detectable in group 8, 11, and 15 accessions (**Fig 2d-e**). The lack of flowering, short stature, and smaller size match the characteristics of the winter annual biotype of *C. sativa* plants, which do not bolt and flower without prior vernalization and, therefore, stay short and compact[22]. It appears that these three UMAP designations are *C. sativa* accessions winter biotype accessions.

Notably, in the measurements of total shoot area (**Fig 2a**), width (**Fig 2b**), leaf area (**Fig 2c**), convex hull area (**Fig 2h**), and height (**Fig 2i**), the non-winter biotype UMAP groups vary greatly in their phenotypic characteristics, especially at their final time points. For example, UMAP group 3 plants tend to be the largest plants (in the term area (**Fig 2a**), convex hull area (**Fig 2h**), and leaf area (**Fig 2c**)), but had fewer inflorescences than plants from UMAP groups 3, 5, and 14 (**Fig 2d-e**). As they approach maturity, *C. sativa* plants bolt and produce flower stalks. Oftentimes, many auxiliary stem branches from primary stem bolts, which leads to many clusters of floral tissues. During this time, the rosette leaves begin to senesce as the plants mobilize their nutrients to reproductive tissues. Manual inspection of images indicated that rosette leaves of plants belonging to UMAP group 14 senesce earlier than those of other groups, and given that these plants had the largest flower area (**Fig 2d**) it suggests that these plants might mobilize their energies to produce more floral tissue, though in smaller quantities than plants of UMAP group 5 (**Fig 2e**). This has the observable effect of lowering the width of the plants (**Fig 2b**) while increasing floral area.

Our study provides insight into potential *C. sativa* ideotypes (**Fig 2a-i**). Future research will focus on how visibly measurable traits can influence the content of oil in seeds (for example, perhaps plants with less leaves might produce oilier seeds). We will also perform Genome-Wide Association Studies on these data to link trait variations to specific genomic loci.
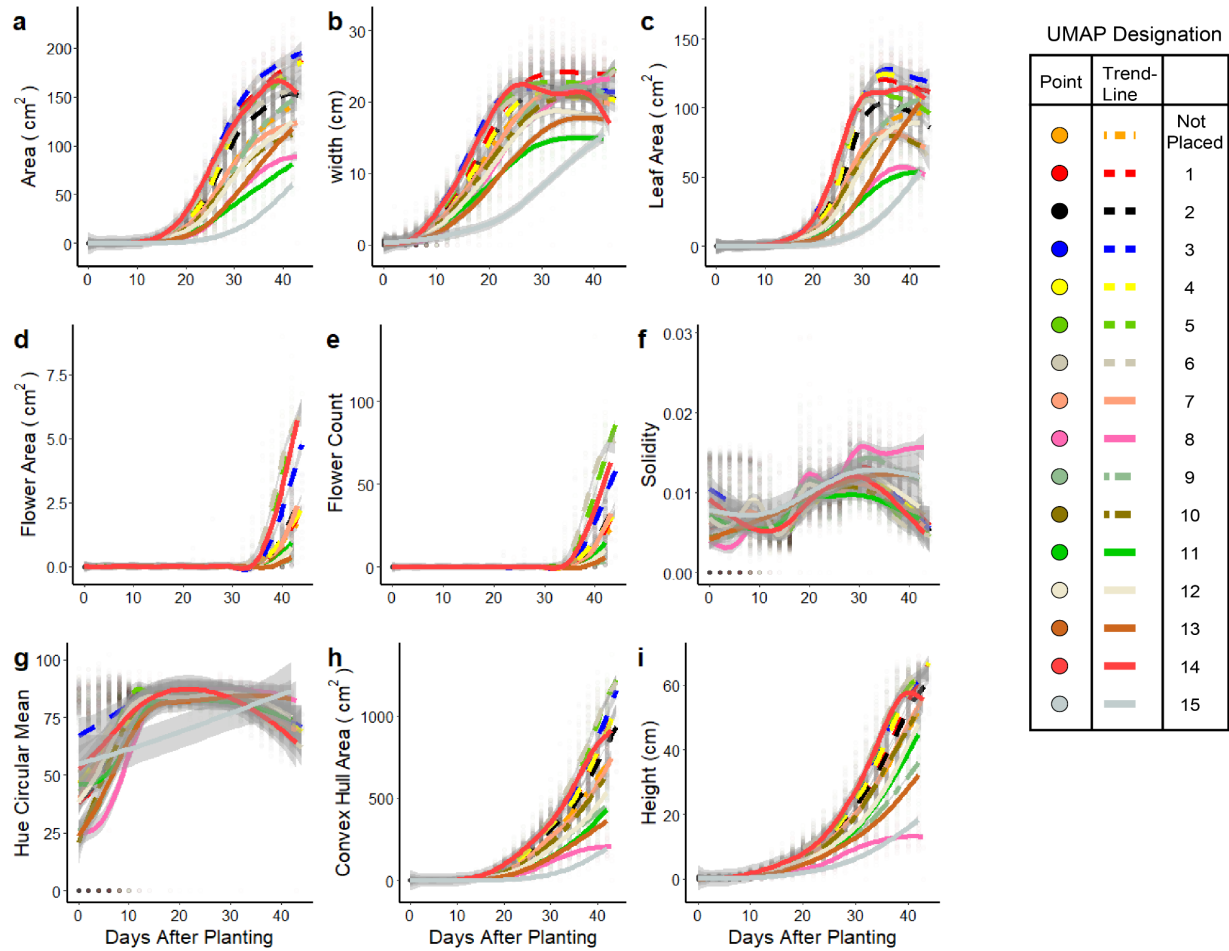
Figure 2. Trait analysis based on phenotypic grouping of *C. sativa* accessions. *C. sativa* accessions imaged and analyzed using PlantCV. Only "frame 0" images are used for phenotyping. Traits examined were a) area, b) width, c) area of leaves and d) flowers, e) number of flowers, f) solidity (area/convex hull area), g) hue circular mean, h) convex hull area, and i) height.

DATA AVAILABILITY STATEMENT

There are no shareable data for this project.

REFERENCES

1. Monthly Energy Review – February 2021. 263 (2021).
2. Marland, G., Brenkert, A. & Olivier, J. CO2 from fossil fuel burning: a comparison of ORNL and EDGAR estimates of national emissions. Environ. Sci. Policy 2, 265–273 (1999).
3. Zhang, G., Murakami, H., Knutson, T. R., Mizuta, R. & Yoshida, K. Tropical cyclone motion in a changing climate. Sci. Adv. 6, eaaz7610.
4. Knutson, T. R. & Tuleya, R. E. Impact of CO2-Induced Warming on Simulated Hurricane Intensity and Precipitation: Sensitivity to the Choice of Climate Model and Convective Parameterization. J.

Clim. 17, 3477–3495 (2004).

5. Cook, B. I., Mankin, J. S. & Anchukaitis, K. J. Climate Change and Drought: From Past to Future. Curr. Clim. Change Rep. 4, 164–179 (2018).

6. Ashley, R. M., Balmforth, D. J., Saul, A. J. & Blanskby, J. D. Flooding in the future – predicting climate change, risks and responses in urban areas. Water Sci. Technol. 52, 265–273 (2005).

7. Mirza, M. M. Q. Climate change, flooding in South Asia and implications. Reg. Environ. Change 11, 95–107 (2011).

8. Douglas, I. et al. Unjust waters: climate change, flooding and the urban poor in Africa. Environ. Urban. 20, 187–205 (2008).

9. Archer, D., Kite, E. & Lusk, G. The ultimate cost of carbon. Clim. Change 162, 2069–2086 (2020).

10. Shonnard, D. R., Williams, L. & Kalnes, T. N. Camelina-derived jet fuel and diesel: Sustainable advanced biofuels. Environ. Prog. Sustain. Energy 29, 382–392 (2010).

11. Gugel, R. K. & Falk, K. C. Agronomic and seed quality evaluation of Camelina sativa in western Canada. Can. J. Plant Sci. 86, 1047–1058 (2006).

12. Kansas State University. Camelina used to build better biofuel. ScienceDaily https://www.sciencedaily.com/releases/2014/08/140804122909.htm.

13. Berti, M., Gesch, R., Eynck, C., Anderson, J. & Cermak, S. Camelina uses, genetics, genomics, production, and management. Ind. Crops Prod. 94, 690–710 (2016).

14. Ghamkhar, K. G. et al. Camelina (Camelina sativa (L.) Crantz) as an alternative oilseed: molecular and ecogeographic analyses. Genome (2010) doi:10.1139/G10-034.

15. Moser, B. R. Camelina (Camelina sativa L.) oil as a biofuels feedstock: Golden opportunity or false hope? Lipid Technol. 22, 270–273 (2010).

16. Luo, Z. et al. Genetic Diversity and Population Structure of a Camelina sativa Spring Panel. Front. Plant Sci. 10, 184 (2019).

17. Fahlgren, N. et al. A Versatile Phenotyping System and Analytics Platform Reveals Diverse Temporal Responses to Water Availability in Setaria. Mol. Plant 8, 1520–1535 (2015).

18. Gehan, M. A. et al. PlantCV v2: Image analysis software for high-throughput plant phenotyping. PeerJ 5, e4088 (2017).

19. Bunn, A. & Korpela, M. An Introduction to dplR. 16.

20. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. ArXiv180203426 Cs Stat (2020).

21. Ester, M., Kriegel, H.-P. & Xu, X. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. 6.

22. Gesch, R. W. & Cermak, S. C. Sowing Date and Tillage Effects on Fall-Seeded Camelina in the Northern Corn Belt. Agron. J. 103, 980–987 (2011).