Measuring the shape of citrus using Topological Data Analysis

Erik Amezquita[1], Michelle Quigley[1], Elizabeth Munch[1], and Dan Chitwood[1]

[1]Michigan State University

November 21, 2022

Abstract

Shape is foundational to biology. Observing and documenting shape has fueled biological understanding as the shape of biomolecules, cells, tissues, and organisms arise from the effects of genetics, development, and the environment. The vision of Topological Data Analysis (TDA), that data is shape and shape is data, will be relevant as biology transitions into a data-driven era where meaningful interpretation of large datasets is a limiting factor. We focus on quantifying the morphology of X-ray CT scans of a diverse panel of citrus and their tissues. In total, we have access to 146 different citrus samples comprising 58 different species, including samples of all fundamental citrus species. First, using the power of X-rays and image processing, we are able to compare and contrast size ratios between different tissues, such as the size of exocarp compared to the mesocarp or the endocarp. Second, using Persistent Homology we focus on characterizing the size and distribution of oil glands across different citrus species, and the relation of such distribution to other phenotypes such as exocarp surface area. This shape characterization will allow us later to link genotype with phenotype, furthering our understanding on how the physical shape is genetically coded in DNA.

Measuring the shape of citrus using Topological Data Analysis

Erik J. Amézquita[1], Michelle Y. Quigley[2], Danielle Seymour[4], Daniel H. Chitwood[1,2], and Elizabeth Munch[1,3]

[1]*Dept. of Computational Mathematics, Science, & Engineering, Michigan State University*
[2]*Dept. of Horticulture, Michigan State University*
[3]*Dept. of Mathematics, Michigan State University*
[4]*Dept. of Botany & Plant Sciences, University of California—Riverside*

October 15, 2021

Abstract

Shape is foundational to biology. Observing and documenting shape has fueled biological understanding as the shape of biomolecules, cells, tissues, and organisms arise from the effects of genetics, development, and the environment. The vision of Topological Data Analysis (TDA), that data is shape and shape is data, will be relevant as biology transitions into a data-driven era where meaningful interpretation of large datasets is a limiting factor. We focus on quantifying the morphology of X-ray CT scans of a diverse panel of citrus and their tissues. In total, we have access to 146 different citrus samples comprising 58 different species, including samples of all fundamental citrus species. First, using the power of X-rays and image processing, we are able to compare and contrast size ratios between different tissues, such as the size of exocarp compared to the mesocarp or the endocarp. Second, using Persistent Homology we focus on characterizing the size and distribution of oil glands across different citrus species, and the relation of such distribution to other phenotypes such as exocarp surface area. This shape characterization will allow us later to link genotype with phenotype, furthering our understanding on how the physical shape is genetically coded in DNA.