

Characterization of the USDA Cucurbita pepo, Cucurbita moschata, and Cucurbita maxima Collections

This manuscript ([permalink](#)) was automatically generated from [ch728/cucurbit-usda@7dda7a1](#) on August 24, 2021.

Authors

- **Christopher Owen Hernandez**

 [XXXX-XXXX-XXXX-XXXX](#) ·  [ch728](#)

Department of Plant Breeding and Genetics, Cornell University · Funded by Grant XXXXXXXX

Abstract

The *Cucurbita* genus is home to a number of economically and culturally important species. We present the analysis of genotyping-by-sequencing data generated from sequencing the USDA germplasm collections of *Cucurbita pepo*, *Cucurbita moschata*, and *Cucurbita maxima*. These collections include a mixture of wild, landrace, and cultivated specimens from all over the world. Roughly 4,000 - 40,000 quality SNPs were called in each of the collections, which ranged in size from 314 to 829 accessions. Genomic analyses were conducted to characterize the diversity in each of the species and revealed extensive structure corresponding to a combination of geographical origin and morphotype/market class. GWAS was conducted for each data set using both historical and contemporary data, and signals were detected for several traits, including the bush gene (*Bu*) in *C. pepo*. These data represent the largest collection of sequence *Cucurbita* and can be used to direct the maintenance of genetic diversity, develop breeding resources, and to help prioritize whole-genome re-sequencing for further GWAS and other genomics studies aimed at understanding the phenotypic and genetic diversity present *Cucurbita*.

Introduction

The *Cucurbitaceae* (Cucurbit) family is home to a number of vining species mostly cultivated for their fruits. This diverse and economically important family includes cucumber (*Cucumis sativa*), melon (*Cucumis melo*), watermelon (*Citrullus lanatus*), and squash (*Cucurbita* spp.). Like other cucurbits, squash exhibit diversity in growth habit, fruit morphology, metabolite content, disease resistance, and have a nuanced domestication story. The genomes of *Cucurbita* spp. are small (roughly 500 Mb), but result from complex interactions between ancient genomes brought together through an allopolyploidization event. These factors make squash an excellent model for understanding the biology of genomes, fruit development, and domestication. Within *Cucurbita*, five species are recognized as domesticated. Three of these are broadly cultivated: *Cucurbita maxima*, *Cucurbita moschata*, and *Cucurbita pepo*. Few genomic resources have been available for working with these species; although, draft genomes and annotations, along with web-based tools and other genomics data are emerging. Already, these resources have been used to elucidate the genetics of fruit quality, growth habit, disease resistance, and to increase the efficiency of cucurbit improvement; however, there has yet to be a comprehensive survey of the genetic diversity in large diverse *Cucurbita* germplasm panels, such as those maintained by the USDA within the Germplasm Resources Information Network (GRIN) system.

Germplasm collections play a vital role in maintaining and preserving genetic variation. These collections can be mined by breeders for valuable alleles and can be used by geneticists for mapping studies. Many of the collections of The Cucurbit Coordinated Agricultural Project (CucCap project) have been established to help close the knowledge gap in Cucurbits. This collaborative project aims to provide genomics resources and tools that can aid in both applied breeding and basic research. The genetic and phenotypic diversity present in the USDA watermelon and cucumber collections has already been explored as part of the CucCap project, partially through the sequencing of USDA germplasm collections and development of core collections for whole-genome sequencing.

The classification system used in squash is complex. Squash from each species can be classed as winter or summer squash depending on whether the fruit is consumed at an immature or mature stage, the latter is a winter squash. Squash are considered ornamental if they are used for decoration, and some irregularly shaped, inedible ornamental squash are called gourds; however, gourds include members of *Cucurbita* as well as some species from *Lagenaria*—not all gourds are squash. Many squash are known as pumpkins; the pumpkin designation is a culture dependent colloquialism that can refer to jack O' lantern types, squash used for desserts or, in some Latin

American countries, to eating squash from *C. moschata* known locally as Calabaza . Cultivars deemed as pumpkins can be found in all widely cultivated squash species. Unlike the previous groupings, morphotypes/market classes are defined within species. For example, a Zucchini is reliably a member of *C. pepo* and a Buttercups are from *C. maxima*. Adding to the complexity of their classification, the *Cucurbita* species are believed to have arisen from independent domestication events and the relationships between cultivated and wild species remains poorly understood .

is the most economically important of the species and is split into two different subspecies: subsp. and subsp. . Evidence points to Mexico as the center of origin for and southwest/central United States as the origin of . The progenitor of is considered by some to be subsp. var. , whereas subsp. is a candidate progenitor for . Europe played a crucial role as a secondary center of diversification for , but not . Important morphotypes of include Zucchini, Spaghetti squash, Cocozelle, Vegetable marrow, and some ornamental pumpkins. subsp. includes summer squash from the Crookneck, Scallop, and Straightneck group, and winter squash such as Delicata and Acorn .

The origin of *C. moschata* is more uncertain than *C. pepo*; it is unclear whether *C. moschata* has a South or North American origin . Where and when domestication occurred for this species is also unknown; however it is known that *C. moschata* had an India-Myanmar secondary center of origin where the species was further diversified . *C. moschata* plays an important role in squash breeding as it cross-fertile to various degrees with *C. pepo* and *C. maxima*, and can thus be used as a bridge to move genes across species . Popular market classes of *C. moschata* include Cheese types like Dickenson, which is widely used for canned pumpkin products, Butternut (neck) types, Japonica, and tropical pumpkins known as Calabaza .

C. maxima contains many popular winter squash including Buttercup/Kobocha types, Kuri, Hubbard, and Banana squash . This species also sports the world's largest fruit, the giant pumpkin whose fruit are grown for competition and can reach well over 1000 Kg . Although this species exhibits a wide range of phenotypic diversity in terms of fruit characteristics, it appears to be the least genetically diverse of the three species described . *C. maxima* is believed to have a South American origin, and was likely domesticated near Peru, with a secondary center of domestication in Japan/China.

In this study, we set out to characterize the genetic diversity present in the USDA germplasm collections for *C. pepo*, *C. moschata*, and *C. maxima*. We present genotyping-by-sequencing data from each of these collections, population genomics analysis, results from genome-wide association using historical and contemporary phenotypes, and develop a core panel for re-sequencing.

Material and Methods

References
