Divergence of defensive cucurbitacins in independent Cucurbita pepo domestication events leads to differences in specialist herbivore preference

Lauren J. Brzozowski1 | Michael A. Gore1 | Anurag A. Agrawal2 | Michael Mazourek1

1Plant Breeding and Genetics Section, School of Integrative Plant Science, Cornell University, Ithaca, New York
2Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York

Abstract
Crop domestication and improvement often concurrently affect plant resistance to pests and production of secondary metabolites, creating challenges for isolating the ecological implications of selection for specific metabolites. Cucurbitacins are bitter triterpenoids with extreme phenotypic differences between Cucurbitaceae lineages, yet we lack integrated models of herbivore preference, cucurbitacin accumulation, and underlying genetic mechanisms. In Cucurbita pepo, we dissected the effect of cotyledon cucurbitacins on preference of a specialist insect pest (Acalymma vittatum) for multiple tissues, assessed genetic loci underlying cucurbitacin accumulation in diverse germplasm and a biparental F2 population (from a cross between two independent domesticates), and characterized quantitative associations between gene expression and metabolites during seedling development. Acalymma vittatum affinity for cotyledons is mediated by cucurbitacins, but other traits contribute to whole-plant resistance. Cotyledon cucurbitacin accumulation was associated with population structure, and our genetic mapping identified a single locus, Bi-4, containing genes relevant to transport and regulation – not biosynthesis – that diverged between lineages. These candidate genes were expressed during seedling development, most prominently a putative secondary metabolite transporter. Taken together, these findings support the testable hypothesis that breeding for plant resistance to insects involves targeting genes for regulation and transport of defensive metabolites, in addition to core biosynthesis genes.

KEYWORDS
Acalymma vittatum, plant-herbivore interactions, squash, secondary metabolite, terpenoid

1 | INTRODUCTION

Plants produce a diverse array of secondary metabolites, yet many are restricted to certain plant populations or lineages, complicating efforts to characterize their specific roles in ecological interactions and the genetic basis of production and accumulation (Moore, Andrew, Kühlheim, & Foley, 2014). In agricultural plants, lineage-specificity is amplified by the genetic bottlenecks associated with domestication and subsequent improvement (Ladizinsky, 2012). This process has often been associated with a decrease in resistance to insect herbivores, but is not consistently correlated with plant-wide loss of defensive metabolites (Whitehead, Turcotte, & Poveda, 2017). Thus, a challenge for modern plant breeding is understanding the degree to which past selective pressures shaped the genetic bases of...
intraspecific variation in defensive metabolites, and how this can be leveraged to augment resistance in future breeding efforts.

Intraspecific variation in defensive chemistry in crop plants can have major effects on interactions with herbivores both when there is a complete loss of biosynthesis (e.g., Rasmann et al., 2005) or a reduction of metabolites (Poelman, Van, Van, Vet, & Dicke, 2009). Connecting the ecological phenotype of resistance in the field to the genetic mechanisms can be straightforward in the case of complete pathway knockouts. In the context of plant breeding, naturally occurring mutants are identified, selected, and widely introgressed through intensive plant breeding (e.g., loss of fruit pungency in Capsicum annuum; Stewart et al., 2005), or specifically developed by genetic engineering (e.g., nicotine in Nicotiana attenuata; Steppuhn, Gase, Krock, Halitschke, & Baldwin, 2004). However, variation can also stem from spatially and temporally dynamic processes like induction by insect feeding or other stressors (Karban & Baldwin, 1997), transport (Erb, Lenk, & Turlings, 2009), or changes associated with ontogeny (Barton & Boege, 2017; Boege & Marquis, 2005). Thus, investigations into the genetic basis of intraspecific variation in defensive metabolites should involve study of how biosynthesis interacts with spatiotemporally-sensitive processes like regulation, storage and transport (Burow & Halkier, 2017; Shitan, Hayashida, & Yazaki, 2015).

Cucurbitacins, bitter triterpenoids predominately found in the Cucurbitaceae, are an ideal class of compounds for investigating the relationship between herbivory and the dynamic accumulation of defensive metabolites. Eighteen cucurbitacins (named by letters A-T) have been described (Chen, Chiu, Nie, Cordell, & Qiu, 2005), and abundance varies within and between species, tissues (Metcalf et al., 1982; Rehm, Enslin, Meeuse, & Wessels, 1957; Theis, Barber, Gillespie, Hazzard, & Adler, 2014) and developmental stages (Rehm & Wessels, 1957). Disruption of cucurbitacin accumulation has strong effects on herbivore preference: mutants that have complete pathway disruption of cucurbitacins have been invaluable in demonstrating that the absence of species-specific cucurbitacin C from cucumber (Cucumis sativus) confers susceptibility to generalist herbivores, and resistance to specialist Diabroticite beetles (Coleoptera: Chrysomelidae) that prefer and sequester cucurbitacins (Agrawal, Gorski, & Tallamy, 1999; Da Costa & Jones, 1971; Metcalf, Metcalf, & Rhodes, 1980). However, this finding is restricted to a single lineage (Cucumis sativus), and despite discovery of the mutant phenotype over half a century ago (Andeweg & De Bruyn, 1959), the genetic mechanism remains unconfirmed.

Studies of genetic mechanisms of cucurbitacin accumulation and herbivore preference have been largely independent. The genetic pathway of cucurbitacin accumulation in aboveground tissues has been facilitated by the characterization of cucurbitacin C biosynthesis in cucumber (Shang et al., 2014). In cucumber, melon (Cucumis melo), and watermelon (Citrullus lanatus), biosynthesis occurs locally in tissues through activation of the first committed step of a single oxidosqualene cyclase (followed by numerous oxidations and acetylation) by leaf-, root- or fruit-specific transcription factors (Shang et al., 2014; Zhou et al., 2016). Mutations in fruit-specific regulation (Guo et al., 2019; Shang et al., 2014; Zhao et al., 2019) were fixed in selective sweeps during domestication of multiple species, but selection for loss of oxidosqualene cyclase function was also implicated in loss of fruit bitterness in Cucumis melo ssp. melo (Zhao et al., 2019). In contrast, there is a paucity of research on genetic mechanisms differentiating lineages (especially intraspecific lineages, but see Zhao et al., 2019) and in tissues relevant for herbivores.

Building upon this work in cultivated Cucurbitaceae, Cucurbita pepo provides an excellent opportunity to connect the genetic mechanisms of cucurbitacin accumulation and herbivory because of the evolutionary association between C. pepo and cucurbitacin-sequestering specialized Diabroticite beetles (Metcalf & Lampman, 1989; Tallamy & Krischik, 1989). Both the Diabroticite beetles (Metcalf et al., 1980) and Cucurbita spp. (Sanjur, Piperno, Andres, & Wessel-Beaver, 2002) are native to the Americas, where the two cultivated C. pepo subspecies, C. pepo ssp. pepo ("CPP", e.g., zucchini) and C. pepo ssp. ovifera ("CPO", e.g., summer squash, syn. C. pepo ssp. texana), were independently domesticated (Decker, 1988; Sanjur et al., 2002). After domestication, the subspecies further diverged as CPP cultivars were largely developed in Europe, free of specialist herbivores, while CPO cultivars were bred in the Americas (Paris, 2000).

Cucurbitacins are absent in true leaves and fruits of cultivated C. pepo (Metcalf et al., 1982). At the seedling stage, however, at least a few cultivars of CPP are rich in four types of cucurbitacins (B, D, E, I) in cotyledons and roots, while CPO lack cucurbitacins plant-wide (Ferguson, Metcalf, & Metcalf, 1983; Hirsh & Barbercheck, 1996; Rehm & Wessels, 1957; Tallamy & Gorski, 1997). While this pattern suggested pathway disruption in CPO, likely by a single gene (Sharma & Hall, 1971), the genetic mechanisms differentiating the subspecies are unknown. The specialized herbivore, Acalymma vittatum (Coleoptera: Chrysomelidae), feeds on all plant tissues (cotyledons, leaves, flowers, fruit and roots), but foliar preference is furthermore associated with subspecies: both the cotyledons (high cucurbitacins) and leaves (no cucurbitacins) of CPP are preferred over CPO (Brzozowski, Leckie, Gardner, Hoffmann, & Mazourek, 2016; Ferguson et al., 1983; Hoffmann, Robinson, Kyle, & Kirkwyland, 1996). Here we generated a new mapping population by crossing CPP and CPO to help disentangle the association between cucurbitacins and selection history, allowing us to identify genetic loci implicated in biosynthesis or accumulation, and to understand if plant-wide herbivore preference is indeed predicated on cucurbitacins.

Overall, we sought to characterize the role of cotyledon cucurbitacins in herbivore preference between C. pepo subspecies and to elucidate the genetic and mechanistic bases of cotyledon cucurbitacin production. We hypothesized that cotyledon cucurbitacins affect A. vittatum preference independently of other factors associated with the divergence between C. pepo subspecies. To address this, we selected for phenotypic extremes for multiple generations in a mapping population, and measured A. vittatum preference in the field over 2 years. We also tested whether a lack of cotyledon cucurbitacins in CPO relative to CPP is due to loss of function of a major pathway gene affecting all tissues. In addition to
conducted biparental genetic mapping and screening diverse C. pepo germplasm, we also measured gene expression and cucurbitacin concentration over seedling development to identify putative transporters or regulatory factors contributing to spatiotemporal variation in cucurbitacin accumulation. These approaches allowed us to estimate the role of different genetic mechanisms contributing to cucurbitacin variation, as well as insights into past selection processes in cucurbitacin accumulation. These approaches allowed us to estimate the role of different genetic mechanisms contributing to cucurbitacin variation, as well as insights into past selection processes in cucurbitacin accumulation.

2 | MATERIALS AND METHODS

2.1 | Plants

Seeds for selection, mapping and expression studies were Cornell University stock seed: parent lines Cucurbita pepo ssp. pepo cv. Black Beauty (hereafter “CPP”), and C. p. ssp. ovifera cv. Success PM (hereafter “CPO”), and derived F1 plant and F2 mapping population. Seeds of the C. pepo diversity panel were from the selfed progeny of USDA PI lines (Table S1). Plants were started from untreated seeds in organic potting soil, and no additional fertilizer or pest control was used. Plants used in insect bioassays were started in Cornell University Agricultural Experiment Station Guterman greenhouses (Ithaca, NY), and all others were grown in an indoor growing room (Ithaca, NY). Both environments were certified organic, and had a 16 hr day, 8 hr night photoperiod with supplemental sodium halide lighting with 27°C day and 21°C night temperatures.

2.2 | Cucurbitacin extractions

Cucurbitacins were extracted from root, cotyledon, and leaf tissue at different developmental stages using protocols similar to (Brzozowski, Mazourek, & Agrawal, 2019), and detailed in Method S1. Briefly, fresh tissue was homogenized in methanol and purified by solid phase extraction. Samples were then suspended in acetonitrile before quantification on UHPLC-MS system equipped with a C18 column (150 mm × 2.1 mm, 2.6 μm particle size).

2.3 | Cucurbitacin gene homolog identification

We identified potential homologs of known cucurbitacin biosynthetic genes in Cucurbita pepo to test for allelic variation, and to examine gene expression patterns. Gene lists from cucumber (Cucumis sativus), melon (Cucumis melo) and watermelon (Citrullus lanatus) were compiled from Shang et al. (2014) and Zhou et al. (2016), and BLAST (Altschul, Gish, Miller, Myers, & Lipman, 1990) was used to identify homologs in Cucurbita pepo. Gene function and location are shown in Table S2, and C. pepo homologs follow the naming convention of Zhou et al. (2016).

2.4 | Isolating effects of cotyledon cucurbitacins on herbivore preference – Biparental F2 population evaluation

Cucurbitacins B, D, E and I were measured in the F2 mapping population derived from high cotyledon cucurbitacin CPP and low cotyledon cucurbitacin CPO. Cucurbitacins were extracted from 188 F2 individuals in four blocks of 55–58 plants, each with three replicates of CPO, CPP and the F1 as checks in an augmented incomplete block design. The extraction blocks were staggered over 4 weeks in January 2018, but all cucurbitacins were quantified together. Summary statistics and Pearson’s correlations were calculated for all checks and the F2 population for concentration of all individual and total cucurbitacins, and the B:D, E:I and BD:EI ratios. Individuals were excluded from statistical summaries of ratios when the denominator value was below the limit of detection. Best linear unbiased predictors (BLUPs) were then calculated for all phenotypes to be used in genetic mapping. Details of data transformation, outlier removal and BLUP calculation are given in Method S2.

2.5 | Isolating effects of cotyledon cucurbitacins on herbivore preference – Selection experiment

We tested the effect of selection for cotyledon cucurbitacins on A. vittatum preference for cotyledon, leaf, and floral tissues in summer 2018 and 2019 in CPP and CPO, as well as intersubspecific families derived from their cross. This allowed us to test the effect of cotyledon cucurbitacins on A. vittatum preference phenotypes independent of C. pepo population structure.

To generate the intersubspecific families, F2 individuals from the previously described F2 population with the 10% highest and lowest extremes of total cucurbitacins (n = 14 of each extreme) were selected and self-pollinated to create F2:3 families. Then, individuals from those F2:3 families (n = 6, 7 individuals from different low and high F2:3 families, respectively) were again selfed to create F3 families. This process is summarized in Figure S1. The F2:3 families were evaluated in 2018 and the F3 families were evaluated in 2019. It was necessary to evaluate succeeding generation each year due to limited seed quantity in the F2:3 generation.

Trials were conducted on the Homer C. Thompson Organic Vegetable Research Farm (Freeville, NY), where there was naturally abundant A. vittatum infestation. Families and parents (CPO and CPP) were evaluated by being: (a) sown into 72-cell peat pots and transplanted immediately after germination to measure damage to newly emerged cotyledons (“cotyledon plots”), and (b) sown into plastic 72-cell trays and transplanted with two fully expanded leaves to independently evaluate leaf damage (“leaf plots”). Both were transplanted into raised beds with black plastic mulch and drip irrigation in staggered double rows as plots of 12 plants each. Cotyledon plots had 15 cm spacing between plants within plots and 30 cm spacing between plots, and leaf plots had 45 cm between plants and 60 cm spacing between...
plots. There were four replicates of cotyledon plots, and three replicates of leaf plots.

Damage was visually estimated as percent defoliation of individual plants 1 week after transplanting on cotyledons in cotyledon plots, and 2 weeks after transplanting on leaves in both leaf and cotyledon plots. Beetles were counted at the plot level on foliage in cotyledon plots 3 weeks after transplanting (2018), and in floral tissue 4 weeks after transplanting in cotyledon (2018, 2019) and leaf (2019) plots.

Statistical analyses were conducted separately by year, and plot means (of a maximum of \( n = 12 \) plants) were used for analysis of leaf and cotyledon damage. Plots with fewer than two flowers were excluded from analysis of floral count data, and \( F_{2:3} \) plots (2018) with less than five (of 12) individuals were excluded from all analyses. Phenotypic differences (damages and beetle counts) between CPP and CPO were evaluated with linear mixed models with a random effect of replicate and fixed effect of genotype (CPO or CPP). To evaluate phenotypic differences due to cotyledon cucurbitacins in the intersubspecific families, there was a fixed effect of selection direction (high or low cotyledon cucurbitacins), random effect of family nested within selection direction, and a random effect of replicate. The analysis method was adjusted slightly for floral beetle counts in 2019 for both comparisons between CPO and CPP, and comparisons between families: since floral counts were taken in both leaf and cotyledon plots, there was an additional fixed effect of plot type (cotyledon or leaf), along with its interaction with genotype (for parents) or selection direction (for families), and the random effect of replicate was nested within plot type. In all models, an ANOVA was conducted on the fixed effect.

2.6 | Isolating effects of cotyledon cucurbitacins on herbivore preference – Induction experiment

In the selection experiment, we measured the correlation between \( A. vittatum \) damage to cotyledons and leaves, so we thus also sought to test if \( A. vittatum \) feeding induced cucurbitacin accumulation in leaves as a mechanistic connection between preference for the two tissues. The design is detailed in Method S3. Briefly, using CPP and CPO, we tested induction by five \( A. vittatum \) feeding on cotyledons, or on leaves, and had a control of no beetle feeding and measured cucurbitacins in both tissues at 24 and 72 hr after the start of beetle feeding.

2.7 | Genetic basis of cotyledon cucurbitacins – Biparental \( F_{2:3} \) family progeny evaluations

Cucurbitacins were measured in fully expanded cotyledons of individuals from six low, and seven high cucurbitacin \( F_{2:3} \) families derived from the intersubspecific cross between CPO and CPP (Figure S1) to determine the inheritance of the trait. Seed availability limited sample sizes to 10 or fewer individuals per family.

2.8 | Genetic basis of cotyledon cucurbitacins – OSC Marker development and testing in biparental population

Prior to genetic mapping in the biparental \( F_2 \) population, we tested if polymorphisms in the homolog of the oxidosqualene cyclase ("OSC") controlling the first committed step of cucurbitacin biosynthesis in other Cucurbitaceae (Table S2) co-segregated with the \( F_2 \) phenotypes. We identified the homolog, which appears to be a single copy gene in \( C. pepo \), developed four PCR primers and protocols, and sequenced the products of CPO and CPP with Sanger sequencing, as is detailed in Method S4. Putative intronic regions were identified by aligning sequences to the annotated \( C. pepo \) genome in the web based Clustal Omega (EMBL-EBI). One primer pair ("OSC5") created products with 200 bp difference in length between CPO and CPP, and we screened nine each of high and low cucurbitacin \( F_2 \) individuals to determine if the product size co-segregated with cucurbitacin phenotype.

2.9 | Genetic basis of cotyledon cucurbitacins – Biparental \( F_2 \) genetic mapping

DNA was extracted from 184 \( F_2 \) individuals, and two replicates of each parent (CPO and CPP) and their \( F_1 \) with a DNeasy 96 Plant Kit (Qiagen, Hilden, Germany) according to the manufacturer’s instructions. A 192-plex genotyping-by-sequencing library was prepared at the University of Wisconsin-Madison Biotechnology Center (Madison, WI, USA), and sequenced at Cornell University Biotechnology Resource Center (Ithaca, NY, USA) on a NextSeq 500 (Illumina, San Diego, CA, USA) with single-end 75 bp reads.

Reads were aligned to the \( Cucurbita pepo \) genome (v4.1) (Montero-Pau et al., 2018) with the "bwa" aligner in the GBSv2 pipeline in TASSEL 5 (Glaubitz et al., 2014). In VCFtools ( Danecek et al., 2011), SNPs that were not biallelic, had extreme mean read depths (<5, or >75), low minor allele frequency (<0.05), or were missing in >20% of samples were removed. Then, only SNPs polymorphic between CPO and CPP were retained. Using LB-Impute and a window size of five (Fragoso, Heffelfinger, Zhao, & Dellaporta, 2016), SNPs were error-corrected and imputed with an accuracy of 92%. Finally, segregation ratios were evaluated in the "qtl" R package with the "geno.table" function (Broman, Wu, Sen, & Churchill, 2003) and those that deviated from the expected 1:2:1 ratio (p < .001) were removed. In all, there were 183 remaining \( F_2 \) individuals, and 8,962 remaining SNPs. This set of SNPs was retained for future multi-locus mixed-model GWAS and TWAS analysis and was further pruned to be computationally manageable for linkage map construction and analysis.

A linkage map was constructed in R/qtl using the est.rf() and est.map() functions with a Kosambi mapping function. This package, redundant SNPs were removed with the "findDupMarkers" function, and unlinked markers and those substantially increasing map length were removed (using the "dropone" function), and individuals with more crossovers than expected were also removed. Chromosome 3 was split into two linkage groups presumably because of lack of
recombination in the pericentromeric region between chromosome arms. In all, 2,927 SNPs were used in the linkage map, and total map length was 2,407.7 cM (Table S3). QTL mapping was conducted in R/qtl using “scanone” and “scantwo” functions with Haley-Knott regression, and significant LOD thresholds were determined by 1,000 permutations.

A multi-locus mixed-model genome-wide association study (GWAS) approach (Segura et al., 2012) was used to test for associations independently of a linkage map. It was implemented in R/mlmm. gwas (Bonnafous et al., 2018), with the filtered (8962) and linkage map filtered (2927) SNP sets. An additive model was used, and set at a maximum of 10 forward steps, and optimal models were selected based on lowest extended BIC (Segura et al., 2012).

2.10 Genetic basis of cotyledon cucurbitacins – Population structure in diverse germplasm

In addition to the biparental population, we extracted cucurbitacins from 117 diverse individuals from USDA Cucurbita pepo PI collection (Table S1). Cucurbitac B only was measured from 2–4 biological replicates of each genetic accession, where each replicate was pooled from two individuals to smooth heterogeneity within accessions. Extractions were conducted in an augmented incomplete block design of six blocks each with three replicates of CPO and CPP as checks. Then, BLUPs were calculated with the following model: log(cucB concentration) ~ check + (1|block) + (1|genotype:new), where “check” was the identity of the checks and “new” refers to whether the genotype was a check or not.

Publicly available GBS SNP data from the accessions (Cucurbit Coordinated Agricultural Project, 2019) was filtered with a minor allele frequency of 0.05 and 25% missingness in sample calls, leaving 24,236 SNPs. Principal component analyses (PCA) was conducted on the genome-wide SNP marker matrix, and the marker matrix of SNPs within the interval identified by QTL mapping (n = 16 SNPs) in TASSEL (Bradbury et al., 2007). The first three principal components accounted for the following percent of variance (genome-wide: PC1: 85.2%, PC2: 9.2%, PC3: 5.6%; QTL mapping interval only: PC1: 61.0%, PC2: 30.8%, PC3: 8.2%), and were used in analysis.

Using the linear model $y \sim PC1 + PC2 + PC3$, we calculated the coefficient of determination to determine the percent of variation in cotyledon cucurbitacin phenotype due to population structure. Then, a linear model was used to test the effect of region the accession was collected from (Table S1) and an ANOVA was conducted on the fixed effect. Both models were assessed with and without wild accessions ($n$ = 3 wild accessions) included.

2.11 Genetic basis of cotyledon cucurbitacins – Gene expression in the biparental $F_2$ population

Cotyledon tissue was flash frozen on liquid nitrogen from all $F_2$ plants immediately prior to removing tissue for phenotyping cucurbitacins in the mapping population. Later, 21 high (total cotyledon cucurbitacin range: 0.37–31.91 μg/g FW, mean: 11.72 μg/g FW) and 14 low (total cotyledon cucurbitacin range: 0.006–0.022 μg/g FW, mean: 0.012 μg/g FW) cucurbitacin individuals were selected for gene expression analysis. RNA was extracted from the $F_2$ individuals and four samples of each parent (CPO and CPP) and their $F_1$ using a modified hot-borate protocol. A 3’ RNA-seq library was then prepared and sequenced on an Illumina NextSeq 500 with single-end 75 bp reads at the Cornell University Biotechnology Resource Center. The sequencing reads were processed by removing the first 12 base pairs and reads less than 25 base pairs in length with Trimmomatic (Bolger, Lohse, & Usadel, 2014), adapter sequence contaminants and polyA stretches greater than 12 in length stretches with cutadapt (Martin, 2011), and base calls with phred < 5 with fastq_quality_trimmer (Hannon, 2014). Then the reads were aligned to the C. pepo v4.1 genome and counted with STAR (Dobin et al., 2013).

The counts were normalized by library size by median ratio method in R/DEseq (Love, Huber, & Anders, 2014).

A transcriptome-wide association study (TWAS) was conducted by a linear regression of normalized count of each expressed gene (genes expressed in less than five individuals or with less than one count per million were dropped from the analysis, from 27,868 to 22,088 genes) using a custom R script. Differential expression was assessed also between the same $F_2$ individuals that we qualitatively classified as having high or low cotyledon cucurbitacins (see ranges above), and separately between parents (CPO and CPP) in R/DEseq (Love et al., 2014).

2.12 Towards a mechanistic understanding of cotyledon cucurbitacin accumulation

To connect the genetic mapping results to biochemical mechanisms, cucurbitacins were measured in root, cotyledon, and leaf tissue over the course of seedling development from CPP and CPO at the stages of: (a) dry seed; (b) seeds imbibed for 24 hr; (c) seedlings with radical emerging; (d) seedlings with cotyledons emerged but not green; (e) seedlings with cotyledons emerged and green; and (f) seedlings with cotyledons fully expanded and a leaf emerging. There were three biological replicates of each sample, where each sample was pooled from two to six individuals, and most had a mass of 0.5 g (but some were lower because of tissue limitations). Cucurbitacins were extracted from developing tissues in two independent experiments conducted in April and May 2019. In both, seeds sown daily for 2 weeks in saturated paper towels in a growth chamber held at 27°C, and visually sorted into the categories for concurrent extractions. In April 2019, seedlings planted in potting soil (like those used in mapping populations) were also included with three to four replications each of cotyledon and root tissue.

Data were analysed with linear models, where there was a fixed effect of experiment (April or May 2019) in every analysis. First, within a genotype and tissue, the effect of developmental stage was tested. Then, for root cucurbitacin concentration alone, the interactive
effects of genotype and developmental stage were tested. Finally, for CPP only, the interactive effects of tissue and developmental stage were tested. In all cases, effect significance was determined with a one-way ANOVA, and differences between effect levels were determined by Tukey honest significant differences at $p < .05$.

Because of size limitations, cotyledon, root, and leaf tissue was collected for RNA from an independent set of plants in October 2019. Two time points were sampled - radical emergence, and fully expanded cotyledons - and there were three replicates of leaf tissue, and five to six replicates of all other tissues. Tissue was ground in liquid nitrogen with a mortar and pestle, and RNA was extracted using a Plant RNeasy kit (Qiagen, Hilden, Germany) according to manufacturer protocols. Library preparation, sequencing, read processing, alignment and generation of normalized counts were conducted as previously described. Genome-wide differential expression was analysed in R/DEseq between genotypes within tissue and developmental stage, and between times within genotype and tissue. Genome-wide differential expression statistics are reported for expression of key biosynthetic genes identified by homology (gene list, Table S2). Genome-wide differential expression results are also reported for genes in the interval identified in QTL mapping of total cucurbitacin concentration in cotyledons at the radical emerging stage.

2.13 | Statistics
Statistical computations were done in R (R Core Team, 2016). Basic summary statistics and linear models were evaluated in base R with the "lm" function. Linear mixed models were evaluated with the "lmer" function R/lme4 (Bates, Maechler, Bolker, & Walker, 2015). Pearson’s correlations were calculated with the “rcorr” function in R/Hmisc.

3 | RESULTS
3.1 | Isolating effects of cotyledon cucurbitacins on herbivore preference
We developed a biparental $F_2$ population between a high cotyledon cucurbitacin $C. pepo$ ssp. $pepo$ (CPP, total 8.04 ± 1.14 μg/g) and low cotyledon cucurbitacin $C. pepo$ ssp. $ovifera$ (CPO, total 0.02 ± 0.01 μg/g)

**FIGURE 1** Damage to cotyledons and leaves of parent lines and selected families in field trials from 2018 and 2019. The parents (solid) and families (dashed) are classified as “low” ($C. p. ovifera$ and derived from low cucurbitacin $F_2$ individuals, respectively) or “high” ($C. p. pepo$ and derived from high cucurbitacin $F_2$ individuals, respectively). The position of the dots is the estimated marginal mean, error bars are ±1 SE and the statistics are given in text.
parent and selected progeny with phenotypic extremes (>5.0 μg/g, or less than <0.02 μg/g) of cotyledon cucurbitacins, a highly heritable trait (full phenotypic results follow; pedigree in Figure S1). This allowed us to assess the effect of cotyledon cucurbitacins on specialist herbivore, Acalymma vittatum, preference of multiple tissues independent of the confounding population structure of the two subpopulations.

Over 2 years of field trials, CPP sustained greater cotyledon damage (2018:149% more, F1,5 = 9.17, p = .03; 2019:61% more, F1,5 = 37.07, p = .002) and leaf (2018:213% more, F1,3 = 19.856, p < .001; 2019:531% more, F1,3 = 38.10, p = .008) damage than CPO, despite neither having appreciable leaf cucurbitacin concentration (Figure 1). In contrast, while fully expanded cotyledons of the families derived from selecting for high cotyledon cucurbitacins sustained more cotyledon damage than families selected for low cotyledon cucurbitacins (2018:69% more, F1,95 = 23.02, p < .001; 2019:19% more, F1,48 = 30.73, p < .001), leaf damage when three true leaves were present was equivalent between selection directions (2018: F1,56 = 0.63, p = .43; 2019: F1,35 = 0.70, p = .41) (Figure 1). In addition, there was no induction of cucurbitacin in leaf or cotyledon tissue by A. vittatum feeding on either tissue of CPP or CPO, further indicating that leaf preference is mechanistically independent of cotyledon cucurbitacins (Table S4).

At a later developmental stage, foliar beetle density was equivalent between CPO and CPP (F1,5 = 0.01, p = .95), but there were 34% fewer on low than high cotyledon cucurbitacin families (F1,92 = 4.32, p = .04). Beetle density in flowers of reproductively mature plants varied between CPO and CPP in only one of the 2 years (2018: equivalent, F1,5 = 0.50, p = .51; 2019: 64% fewer beetles in CPO, F1,6 = 6.78, p = .04, but did not differ between high and low families (2018: F1,87 < 0.001, p = .99; 2019: F1,91 = 0.73, p = .40).

Due to the importance of cotyledon cucurbitacins on herbivore preference of cotyledon tissue, our continued efforts focused on characterizing the genetic and mechanistic bases of cotyledon cucurbitacins.

### 3.2 Genetic basis of cotyledon cucurbitacins – Biparental population mapping

We assessed cotyledon cucurbitacins in fully expanded cotyledons of the parents (CPO and CPP), F1 and F2 biparental population. The cucurbitacin phenotype of the F1 resembled CPP, and the F2 progeny cucurbitacin concentration ranged from 0.003–36.3 μg/g, where cucurbitacin B was the most abundant compound (Table S5) and all cucurbitacins were positively correlated (Table S6). The broad-sense heritability on a plot basis was high for total cotyledon cucurbitacins (0.97), as well as individual compounds (Table S5). When we categorized F2 individuals with total cucurbitacin concentration at or below the CPO mean as “low” cucurbitacins, there was a 3:1 ratio of high to low cucurbitacin individuals (Observed: 131:57, Expected: 141:47, χ² = 1.08, df = 1, p = .30). Together, this result is consistent with a single dominant Mendelian gene model for cotyledon cucurbitacin accumulation, as previously demonstrated by Sharma and Hall (1971).

We tested if different alleles for the first committed biosynthetic step identified by gene homology (oxidosqualene cyclase, OSC, Table S2) caused the observed F2 phenotypes. We sequenced OSC in CPO and CPP, identified only intronic indels, and developed a codominant size-polymorphic PCR marker (Method S4). Marker state in the F2 individuals did not co-segregate with the phenotype, but rather there was a 1:2:1 segregation ratio expected of an unlinked gene within phenotypic class (high: observed, 3:5:1; Fisher’s exact test p = 1; low: observed, 2:6:1; Fisher’s exact test p = 1).

We then conducted QTL mapping with best linear unbiased predictors (BLUPs) calculated from the quantitative phenotype and found a single locus for total cotyledon cucurbitacins on chromosome 5 (LOD 77, p < .001), explaining 87% of phenotypic variation (Figure 2; Table S7), hereafter referred to as “Bi-4” (Paris & Padley, 2014). The Bi-4 interval (recombination breakpoint markers, SCP4.1LG05_1542768, SCP4.1LG05_1802502) is a 259.7 kb, 3.5 cM region containing 47 genes, none of which are cucurbitacin biosynthetic gene homologs (Table S2). This locus was significant for cucurbitacins B, D, E and the BD:EI ratio, but other marginally significant loci were detected for the B:D and EI ratios, and none were detected for cucurbitacin I (Table S7). Three additional loci, two additive and one epistatic, were identified using a two QTL model (Table S8). The same genomic regions were associated with cucurbitacins using a multi-locus mixed-model GWAS approach (Table S9), verifying our results independently of a genetic linkage map. In addition, evaluations of F2.3 progeny where the haplotype of the F2 progenitor in the Bi-4 interval was known demonstrated phenotypes are predicted by the genotype at this single Mendelian locus (Figure 3).

#### 3.3 Genetic basis of cotyledon cucurbitacins – Diverse germplasm evaluation

We measured cotyledon cucurbitacin B in 117 diverse, globally-sourced USDA PI C. pepo accessions (Table S1) and inferred...
population structure from genetic data beyond that previously characterized in US germplasm (Ferguson et al., 1983) to evaluate our QTL mapping results more widely. First, accessions collected from the Americas had lower cotyledon cucurbitacin B concentration than those collected from Europe, Africa and Asia (including wild accessions: $F_{3,113} = 6.27, p < .001$; excluding wild accessions: $F_{3,110} = 7.56, p < .001$, Figure 4). While the USDA collection is not indexed by subspecies, CPP germplasm is widely distributed globally, but CPO germplasm is uncommon outside of the Americas (Paris, 2000). In addition, we found that population structure accounts for a substantial amount of phenotypic variance in cotyledon cucurbitacins both when population structure is inferred based on genome-wide markers, and those in the $Bi-4$ interval (Table 1), indicating that cotyledon cucurbitacin concentration is associated with population differentiation broadly in the species.

3.4 | Genetic basis of cotyledon cucurbitacins – Biparental population gene expression

We next measured gene expression in 35 select $F_2$ individuals in tissue collected concurrently with cucurbitacin extraction associated with genetic mapping (fully expanded cotyledons). In a transcriptome-wide association study (TWAS) to quantitatively associate gene expression levels and cotyledon cucurbitacin concentration, six genes met a modest genome-wide significance threshold ($p_{adj} < .20$; Table 2). Two genes were in the $Bi-4$ interval: Cp4.1LG05g02530, a multidrug and toxic compound extrusion (MATE) transporter and Cp4.1LG05g03720, an unknown protein. When we classified $F_2$ phenotypes qualitatively as present or absent for cucurbitacins, we found 41 differentially expressed (DE) genes between $F_2$ individuals (Table S10). The MATE transporter was the most significant DE gene genome-wide in $F_2$ individuals ($log2foldChange = 7.06, p_{adj} < .001$) and was also DE between CPO and CPP ($log2foldChange = 44.06, p_{adj} < .001$). The MATE transporter is thus a candidate for accumulation of cotyledon cucurbitacins. Biosynthetic genes were not expressed in cotyledons of $F_2$ individuals, CPO or CPP at the developmental stage sampled (fully expanded cotyledons).

3.5 | Towards a mechanistic understanding of cotyledon cucurbitacin accumulation

Finally, we sought to connect the genetic mapping results and observed cotyledon cucurbitacin phenotypes by profiling cucurbitacin accumulation and gene expression in multiple tissues over the course of seedling development. This allowed us to evaluate evidence for cross-tissue transport (as suggested by the MATE transporter), and
### Table 2

Results from transcriptome wide association analysis (TWAS) of \( F_2 \) individuals for cucurbitacin concentration and concurrent gene expression in fully expanded cotyledons. Genes and FDR-adjusted \( p \)-values are presented for the most significant results.

<table>
<thead>
<tr>
<th>Gene</th>
<th>( p )-value</th>
<th>In Bi-4 interval?</th>
<th>Gene description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cp4.1LG05g02690</td>
<td>.006</td>
<td>No, upstream</td>
<td>Ribosomal protein L15</td>
</tr>
<tr>
<td>Cp4.1LG05g02850</td>
<td>.110</td>
<td>No, downstream</td>
<td>SNF1-related protein kinase regulatory subunit beta-2</td>
</tr>
<tr>
<td>Cp4.1LG05g03720</td>
<td>.119</td>
<td>Yes</td>
<td>MATE transporter</td>
</tr>
<tr>
<td>Cp4.1LG04g06320</td>
<td>.135</td>
<td>No, different chromosome</td>
<td>Unknown</td>
</tr>
<tr>
<td>Cp4.1LG05g02570</td>
<td>.151</td>
<td>No, upstream</td>
<td>Oxidoreductase family protein</td>
</tr>
</tbody>
</table>

### Figure 5

Cucurbitacin accumulation and gene expression over seedling development. Each row represents a specific genotype and tissue: (a) high cotyledon cucurbitacin C. p. pepo, "CPP" cotyledons, (b) CPP roots, (c) low cotyledon cucurbitacin C. p. ovifera, "CPO" cotyledons, (d) CPO roots, (e) leaves of CPO and CPP. The leftmost column is cucurbitacin concentration (μg/g FW) over seedling development where the x-axis is developmental stage (DR: dry seed, IM: imbibed seed, RE: radical emerging, CY: cotyledons emerged, not yet green, CG: cotyledons emerged and green, EC: fully expanded cotyledons, and leaf emerging). The middle and right columns are relative gene counts (relative to highest count) on identical scales for biosynthetic homologs in the cucurbitacin pathways, ordered by importance, for the radical emerging and expanded cotyledons time points. Complete gene homolog descriptions are provided in Table S2 (CpBi: oxidosqualene cyclase, CpACT: acetyltransferase, Cp890: cucurbitadienol oxidizer, CpNNN: other potential oxidizers, CpBx: putative tissue-specific transcription factor). Statistical results of differential gene expression are given in Table S11.
for potential regulators in the Bi-4 interval that may be expressed prior to full cotyledon expansion.

Cotyledon cucurbitacins in CPP steadily rose beginning at radial emergence and diminished once cotyledons were fully expanded (developmental stage effect: $F_{3,56} = 33.92$, $p < .001$; Figure 5a), while cotyledon cucurbitacins in CPO were minimal and did not vary over development ($F_{3,16} = 0.92$, $p = .49$; Figure 5c). Unexpectedly, there were cucurbitacins in roots of CPO and CPP, but CPP had 80% greater concentration ($F_{3,38} = 29.04$, $p < .001$). Root cucurbitacin concentrations did not vary by developmental stage within or between genotypes (within CPP, $F_{3,26} = 2.49$, $p = .08$ Figure 5b; within CPO, $F_{3,11} = 0.98$, $p = .44$ Figure 5d: between, $F_{3,38} = 2.52$, $p = .07$).

Cucurbitacin composition varied between tissues: cucurbitacin B was most abundant in cotyledons, but E was highly abundant in root tissue of both CPP and CPO (Figure 5b,d). In a parallel analysis of soil-grown plants at the fully expanded time point, cucurbitacin concentration was nine-fold lower, due to the larger size, but we observed the same magnitude of contrasts (data not presented). Overall, these results indicate that both CPP and CPO can synthesize cucurbitacins but, in CPO, cucurbitacins exclusively accumulate in roots.

We evaluated expression of cucurbitacin biosynthetic genes predicted by homology at two time points of seedling development (radical emergence, fully expanded cotyledons). We found that biosynthetic genes were expressed in all tissues where cucurbitacins accumulated (CPP cotyledons, Figure 5a; CPP roots, Figure 5b; CPO roots, Figure 5d), and not in tissues lacking accumulation (CPO cots Figure 5c; leaves of both, Figure 5e). There was also a temporal shift in biosynthetic gene expression: expression was greatest when the radical was emerging but attenuated by full cotyledon expansion.

While this suggests a regulatory component of cucurbitacin biosynthesis, the homolog of the known tissue-specific regulatory bHLH transcription factor homolog from other Cucurbitaceae was not expressed (Table S11). Count data and differential expression statistical results for all tissues are presented in Table S11.

We also examined expression of the 47 genes in the Bi-4 interval in CPP and CPO during seedling development. Thirty-four of these genes were expressed in cotyledons at the time of radical emergence, 10 of which met a genome-wide DE threshold, including five that have been implicated in other species as relevant for regulation or transmembrane transport (Table 3). Of particular interest are two genes with greater expression in CPP: the MATE transporter, and Cp4.1LG05g03830, an ACT containing regulatory protein induced by abscisic acid (ABA). The ACT protein is also DE between CPO and CPP and F_{2} individuals at the fully expanded cotyledon stage (Table S10).

## DISCUSSION

Plant secondary metabolite accumulation relies on functional biosynthetic genes, regulation – often specific to tissues, developmental stage, or attack by pests – and transport and storage. Selection on any of these components leads to spatial and temporal variation in metabolite accumulation, and has implications for plant-herbivore interactions (Moore et al., 2014). We examined the ecological effects and genetic basis of cucurbitacin accumulation in two independently domesticated lineages of Cucurbita pepo across developmental stages. Overall, we found that herbivore preference is driven both by cotyledon cucurbitacins and cucurbitacin-independent chemistry in the true leaves. Divergence in cotyledon cucurbitacins was attributed to a single Mendelian locus that affects accumulation early in seedling development, not through disruption of biosynthetic gene function, but instead likely through regulation or transport.

### 4.1 Isolating effects of cotyledon cucurbitacins on herbivore preference

The specialist beetle and agricultural pest, Acalymma vittatum, preferentially consumes leaves and cotyledons of C. pepo ssp. pepo (CPP) over C. pepo ssp. ovifera (CPO) (Brzozowski et al., 2016; Ferguson et al., 1983; Hoffmann et al., 1996), despite neither domesticate

### TABLE 3

Differential expression of genes in the Bi-4 interval in cotyledons of C. p. ovifera (“CPO”) and C. p. pepo (“CPP”) at seedling radial emergence. A positive Log2FoldChange value indicates greater expression in CPP, and negative corresponds to greater expression in CPO. The hypothesized role of genes in regulation (“R”) or transport (“T”) is given in the column “Role”.

<table>
<thead>
<tr>
<th>Gene name</th>
<th>Gene function</th>
<th>Role</th>
<th>Base mean</th>
<th>Log2FoldChange</th>
<th>$p_{adj}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cp4.1LG05g03830</td>
<td>ACT containing protein, ACR8</td>
<td>R</td>
<td>34.0</td>
<td>1.25</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cp4.1LG05g02530</td>
<td>MATE transporter</td>
<td>T</td>
<td>4.4</td>
<td>5.72</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cp4.1LG05g03700</td>
<td>40S ribosomal protein S13</td>
<td></td>
<td>3,125.4</td>
<td>0.69</td>
<td>.04</td>
</tr>
<tr>
<td>Cp4.1LG05g03660</td>
<td>Amine oxidase</td>
<td></td>
<td>108.0</td>
<td>0.74</td>
<td>.03</td>
</tr>
<tr>
<td>Cp4.1LG05g03670</td>
<td>Rhomboid-like protein</td>
<td>T</td>
<td>272.6</td>
<td>−1.30</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cp4.1LG05g03910</td>
<td>Protein IQ-domain 31</td>
<td></td>
<td>120.5</td>
<td>−1.17</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cp4.1LG05g03520</td>
<td>Putative myosin heavy-chain protein</td>
<td></td>
<td>50.4</td>
<td>−2.04</td>
<td>&lt;.001</td>
</tr>
<tr>
<td>Cp4.1LG05g03620</td>
<td>Molybdenum cofactor sulferase</td>
<td></td>
<td>30.1</td>
<td>−1.84</td>
<td>.002</td>
</tr>
<tr>
<td>Cp4.1LG05g03560</td>
<td>Putative LRR receptor kinase</td>
<td></td>
<td>18.2</td>
<td>−1.40</td>
<td>.002</td>
</tr>
<tr>
<td>Cp4.1LG05g03690</td>
<td>Stay green protein</td>
<td>R</td>
<td>10.5</td>
<td>−1.85</td>
<td>.042</td>
</tr>
</tbody>
</table>
having cucurbitacins in leaf tissue (Brzozowski et al., 2019; Metcalf et al., 1980, 1982). We sought to isolate the effect of cotyledon cucurbitacins on herbivore preference from other factors associated with subspecies, like defence inducibility (Brzozowski et al., 2019) and morphological traits (Paris, Lebeda, Krístkova, Andres, & Nee, 2012). Both physical (Moore & Tracy, 2019) and chemical (Rasmann et al., 2005) traits implicated in herbivore resistance can be confounded by plant population structure; accordingly, divergent selection for phenotypic extremes in an intermated population is an important method to establish or reject causality. By selecting for high and low cotyledon cucurbitacins in our biparental population, we found that cotyledon cucurbitacins were indeed predictive of A. vittatum preference for cotyledons, but not other tissues in 2 years of field experiments. Thus, cotyledon cucurbitacins are causal for cotyledon preference independent of other (unknown) differences in cotyledons associated with subspecies, but leaf and cotyledon preference are correlated at the subspecies level because of other traits associated with population structure, not pleiotropic effects of cotyledon cucurbitacins. Feeding by A. vittatum also failed to induce cucurbitacin accumulation in leaves or cotyledons, further indicating a lack of a mechanistic connection between those tissues. Acalyymma vittatum likewise did not induce cucurbitacins in cucumber (Cucumis sativus) by root feeding (Milano, Barber, & Adler, 2015), but induction has been reported at a different developmental stage (prior to leaf expansion) by a generalist mite (Agrawal et al., 1999).

### 4.2 Genetic basis of cotyledon cucurbitacins – Genetic mapping

Cotyledon cucurbitacin accumulation in C. pepo has single Mendelian gene inheritance (Sharma & Hall, 1971), but may interact with other loci (Paris & Padley, 2014), and has not been characterized. With a biparental mapping population, we indeed identified a single Mendelian locus, Bi-4, restricted to a genomic interval containing 47 genes. Based on the most recent C. pepo genome sequence (v4.1, Montero-Pau et al., 2018), none of these genes are homologs of described cucurbitacin biosynthesis genes, indicating that selection against cotyledon cucurbitacins in CPO did not result in entire pathway disruption.

In a more diverse set of C. pepo germplasm, variation in cotyledon cucurbitacins was predominately explained by population structure, and accessions collected from the Americas had the lowest cucurbitacin concentration. This divergence could be due to the biogeographical split in cultivar development (Paris, 2000): CPO cultivars were exclusively bred within the range of cucurbitacin-sequestering specialist beetles, the Americas, and thus potentially subject to selective pressure against high cucurbitacin content. Recent work showed divergent selective sweeps for fruit cucurbitacins between subspecies of C. melo (Zhao et al., 2019) and evaluations of high resolution genomic data of independent domestication events in Cucurbita spp. would likely also identify targets of selection. Across systems, divergent past selection processes affected plant-wide secondary metabolite concentration and often herbivory; parallels can be drawn to selection for high and low glucosinolate concentrations in multiple tissues in cultivated Brassicaceae (Hopkins, van Dam, & van Loon, 2009), and to herbivore specificity of expression of benzoxazinoid biosynthetic genes in maize (Tzin et al., 2015).

### 4.3 Genetic basis of cotyledon cucurbitacins – Gene expression

We also measured gene expression in fully expanded cotyledons of the biparental population. An intriguing result was that expression of Cp4-1LG05g02530, a multidrug and toxic compound extrusion (MATE) transporter in the Bi-4 interval is quantitatively and positively correlated with the phenotype. MATE transporters facilitate movement of diverse secondary metabolites like flavonoids, alkaloids, and cyanogenic glycosides in many species (Francisco & Martinoia, 2018). Notably, MATE transporters are necessary for vascular transport of nicotine within and between biosynthetically active and inactive tissues in Nicotiana (Morita et al., 2009; Shitan, Minami, Morita, Hayashida, & Ito, 2014; Shoji et al., 2009).

The cellular localization of cucurbitacins, as well as movement within or between tissues remains uncharacterized, making it difficult to generate hypotheses about how the MATE transporter affects cucurbitacin phenotypes. Nonetheless, cucurbitacins are found in roots of multiple Cucurbita spp. (Theis et al., 2014), and recent work demonstrated that grafting non-bitter Cucumis melo onto a Cucurbita maxima rootstock rendered the fruit bitter (Zhang, Nie, Id, & Cui, 2019), suggesting cross-tissue transport.

### 4.4 Towards a mechanistic understanding of cotyledon cucurbitacin accumulation

Because our expression analysis of F2 individuals was conducted in fully expanded cotyledons, it is possible that regulation or transport were not actively occurring. To address this, we also tracked cucurbitacin accumulation and expression of biosynthetic gene homologs and genes in the Bi-4 interval to assess evidence for cucurbitacin transport or regulation during seedling development of CPO and CPP.

Previous work showed that there is a pulse of cucurbitacin accumulation in seedlings early in development in many Cucurbitaceae species (Jaworski, Gorski, Shannon, & Robinson, 1985; Rehm & Wessels, 1957), and we found that cotyledon cucurbitacins increase only in CPP starting at radical emergence, but remain steady in roots of both CPO and CPP lineages. Tissue-specific accumulation was associated with local conserved biosynthetic gene expression at the stage of radical emergence. Thus, while transport of cucurbitacins may be important in some capacity (e.g., vacuolar storage), cross-tissue transport does not appear to be required for cotyledon cucurbitacin accumulation. Instead, the coordination of biosynthesis with seedling development implicates variation in regulation as potentially mediating cucurbitacin accumulation. In other Cucurbitaceae
(Cucumis sativus, Cucumis melo, Citrullus lanatus), biosynthesis occurs locally through activation by leaf-, root- or fruit-specific bHLH transcription factors (Shang et al., 2014; Zhou et al., 2016). However, biosynthetic gene expression during radical emergence in CPP does not coincide with expression of bHLH transcription factor homologs, suggesting novel regulatory mechanisms in Cucurbita pepo, or regulation at different ontogenetic stages. Local regulation, synthesis, and transport are likewise important in secondary metabolite accumulation in other systems such as glucosinolates in the Brassicaceae (Burov & Halkier, 2017).

To address if there were candidate genes related to regulation of cucurbitacin biosynthesis, we evaluated differentially expressed genes in the Bi-4 interval in cotyledons at the radical stage between CPP and CP. We propose one additional candidate gene; Cp4.1LG05g03830 is an ACT-domain containing protein similar to ACR8 in Arabidopsis (UniProt Consortium, 2019), where it acts as a regulator and expression increases in response to abscisic acid (ABA) (Hsieh & Goodman, 2002). There is higher expression in CPP, and the connection to regulation via ABA is intriguing: ABA is important in seed germination (Finch-Savage & Leubner-Metzger, 2006), and ABA increases cucurbitacin biosynthetic gene expression in cucumber (Shang et al., 2014). If validated, this candidate gene represents a link between regulation of seed germination and cucurbitacin biosynthesis.

More broadly, examining ontogenetic trajectories of defensive metabolites early in seedling development may provide important insights into growth-defense tradeoffs. Cotyledons are a particularly vulnerable life stage (Barton & Hanley, 2013), and we observe that cotyledons of C. pepo are heavily defended by cucurbitacins, yet have contrasting ontogenetic trajectories and cucurbitacin structural diversity than root tissue, which is higher in cucurbitacin E and does not vary with development. This raises questions about the selective pressures or developmental constraints that shaped this phenotype (Barton & Boege, 2017). While our study is limited to two lineages, broad studies within the family could lead to insights about growth-defense tradeoffs across multiple species and domestication events, as we expect that tradeoffs are scale dependent (Agrawal, 2019).

5 | CONCLUSIONS

Our study demonstrates that cucurbitacins are certainly relevant for herbivore preference, even when isolated from biogeographically diverged lineages. Through genetic mapping and gene expression analysis, we showed that potential targets of selection for cotyledon cucurbitacins were regulators or transporters that contribute to dynamic variation of cucurbitacins throughout development and between tissues. The complementary role that such genes play alongside core biosynthetic genes is critical, yet perhaps understudied in genetics of secondary metabolite expression. By characterizing both the ecological effects of cucurbitacins and the genetic factors contributing to intraspecific variation, this work informs present plant breeding efforts to improve insect resistance in C. pepo.

ACKNOWLEDGEMENTS

We thank the Mazourek lab staff, especially Jonathan Vantman for assistance with field work and Paige Reeves for assistance with cucurbitacin extractions, Nicolas Baert for UHPLC analysis of cucurbitacin samples, the Cornell University Agricultural Experiment Station greenhouse staff for providing excellent care of plant material, and Christine Smart and the Agrawal Lab for helpful discussion. This work was supported by the United States Department of Agriculture (USDA) National Institute of Food and Agriculture Multi-State Hatch Project 1008470. LB was supported by a Seed Matters Graduate Student Fellowship (2015–2020).

CONFLICT OF INTEREST

A.A., L.B., and M.G. declare that they have no conflict of interest. M.M. is the co-founder of Row 7, an organic seed company.

ORCID

Lauren J. Brzozowski https://orcid.org/0000-0002-9936-7106
Michael A. Gore https://orcid.org/0000-0001-6896-8024
Anurag A. Agrawal https://orcid.org/0000-0003-0095-1220
Michael Mazourek https://orcid.org/0000-0002-2285-7692

REFERENCES

DIVERGENCE OF CUCURBITACINS IN CUCURBITA PEPO


Zhao, G., Lian, Q., Zhang, Z., Fu, Q., He, Y., Ma, S., ... Huang, S. (2019). A comprehensive genome variation map of melon identifies multiple domestication events and loci influencing agronomic traits. Nature Genetics, 51, 1607–1615.


SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Brzozowski LJ, Gore MA, Agrawal AA, Mazourek M. Divergence of defensive cucurbitacins in independent Cucurbita pepo domestication events leads to differences in specialist herbivore preference. Plant Cell Environ. 2020;43:2812–2825. https://doi.org/10.1111/pce.13844