

Comparative Genomics of the Cucurbitaceae

Padma Nimmakayala, Thangasamy Saminathan, Venkata Lakshmi Abburi, Lav Kumar Yadav, Yan Tomason, Amnon Levi, Yiqun Weng, and Umesh K. Reddy

Abstract The genome size for watermelon (*Citrullus lanatus*), melon (*Cucumis melo*), cucumber (*Cucumis sativus*) and pumpkin (*Cucurbita pepo*) is 425, 454, 367 and 502 Mbp, respectively, and considered medium size as compared with most other crops. Whole-genome duplication is common in angiosperm plants. Research has revealed a paleohexaploidy (Y) event in the common ancestor of eudicots after the divergence of monocotyledons and dicotyledons. While analysis of published whole-genome sequences of cucumber, melon and watermelon showed traces of these ancient duplication events, there was not evidence of more recent whole-genome duplications in these species. Analysis of the syntenic relationships among watermelon, cucumber, melon and grape has identified 3543 orthologous relationships covering the watermelon, melon, and cucumber genomes. Comparison of melon and cucumber genomes synteny to detect shorter regions of rearrangements confirmed previously reported ancestral fusions of five melon chromosome pairs in cucumber, and several inter- and intra-chromosome rearrangements between the two species. Sequenced genomes of cucumber, melon and watermelon revealed a comparable range of genes from 23 k to 27 k protein coding genes with an average of four to six exons per gene. Current progress in gene mapping technologies such

P. Nimmakayala • T. Saminathan • V.L. Abburi • L.K. Yadav • U.K. Reddy (✉)
Department of Biology, Gus R. Douglass Institute, West Virginia State University,
Institute, WV 25112-1000, USA
e-mail: ureddy@wvstateu.edu

Y. Tomason
Department of Selection and Seed Production, Dnepropetrovsk State Agrarian University,
Voroshilov 25, Dnepropetrovsk 49600, Ukraine

Department of Biology, Gus R. Douglass Institute, West Virginia State University,
Institute, WV 25112-1000, USA

A. Levi
U.S. Vegetable Laboratory, USDA-ARS,
2875 Savannah Highway, Charleston, SC 29414, USA

Y. Weng
Vegetable Crops Research Unit, USDA-ARS, Horticulture Department,
University of Wisconsin, Madison, WI 53706, USA

© Springer International Publishing AG 2017

R. Grumet et al. (eds.), *Genetics and Genomics of Cucurbitaceae*,
Plant Genetics and Genomics: Crops and Models, DOI 10.1007/7397_2016_2

as optical mapping, which produces maps of large individual DNA molecules, can improve cucurbit comparative genomics to detect large structural variations and DNA rearrangements across the species.

Keywords Chromosomal rearrangement • Cucumber • Genome annotation • Melon • Synteny • Watermelon • Whole genome duplication

Introduction

The family Cucurbitaceae contains 118 genera and 825 species (Jeffrey 1980). Members are morphologically similar, which implies strong synteny at the molecular level. Cucurbits, which are herbaceous annuals, are often prostrate or climbing by means of tendrils. Stems are typically 5-angled, characterized anatomically by bicollateral vascular bundles often arranged in two concentric rings. Leaves are alternate, exstipulate, simple or occasionally palmately compound, palmately veined and usually lobed (Bates et al. 1990; Whitaker et al. 1976). A notable feature is unisexual flowers with determinate inflorescence. Calyx are symsepalous with five lobes and corolla sympetalous, usually composed of five lobes, and flowers consisting of one to five stamens (usually 3: 2 double stamens and 1 single stamen), anthers, and dehiscent longitudinally (Bates and Robinson 1995). The gynoecium is an inferior ovary with three carpels. Fruits are fleshy, often large, containing several to hundreds of seeds; the exocarp is soft leathery to hard and lignified with phytoliths. Fruit types range from a gourd-like berry or pepo, frequently containing bitter purgative cucurbitacins (bitter-tasting substances) (Nee 1990).

Cucurbitaceae or the cucurbit family is monophyletic because of morphological and biochemical distinctness and represents economically important species, particularly those with edible and medicinal fruits. The family Cucurbitaceae includes domesticated species for food: *Citrullus lanatus* (watermelon), *Cucumis sativus* (cucumber), *Cucumis melo* (melon), *Cucurbita* (five species of squash and pumpkin), *Cucumis anguria* (bur gherkin), *Momordica charantia* (bitter melon), *Sechium edule* (chayote), *Luffa* (two species of loofah), *Lagenaria siceraria* (bottle gourd), *Benincasa hispida* (wax gourd), *Trichosanthes* (two species of snake gourd), *Telfairia* (two species of oyster nut), *Sicana odorifera* (casabanana), *Coccinia grandis* (ivy gourd), *Praecitrullus fistulosus* (tinda), *Cyclanthera pedata* (slipper gourd), and *Cucumeropsis mannii* (white-seeded melon) (Bates and Robinson 1990). The most important cucurbit crops worldwide are cucumber, watermelon, melon, squash and pumpkin (McCreight 2016). In the United States, per-capita civilian utilization (farm weight) of cucurbits (watermelon 6.5 kg, cantaloupe 5.0 kg, honeydew melons 1.0 kg, cucumbers 3.0 kg) is 20% of the total vegetable consumption in the United States, contributing a farm value of \$500 million (Statistics of Vegetables and Melons).

Cucumber (*C. sativus* L.; $2n = 2x = 14$) originated in Asia and is currently a major vegetable crop in countries in Asia, Europe, and North America (Nagele and Wehner 2016). By using molecular markers, several researchers have concluded that the genetic base within each market class of cultivated cucumber is very narrow (Perl-Treves et al. 1985; Knerr et al. 1989; Knerr and Staub 1992; Qi et al. 2013; Lv et al. 2012; Pandey et al. 2013). Resequencing analyses identified four geographic groups: Indian; cultivated lines from Eurasian, Europe, and the United States; cultivated types from East Asia; and Xishuanghanna, comprised largely of landraces cultivated in tropical southwestern China (Qi et al. 2013). The Indian materials, which contained the most genomic diversity, may provide novel genetic resources.

Melon, a diploid plant species (*C. melo* L.; $2n = 2x = 24$) is an important fruit crop, with 26 million tons of melons produced worldwide in 2009 (<http://faostat.fao.org>) (Garcia-Mas et al. 2012). Melons comprise a diverse group of fresh dessert fruits that include the orange-flesh cantaloupes, green-flesh honeydew, and mixed melons (Casaba, Crenshaw, Persian, Santa Claus, Juan Canari) (Pitrat et al. 2000; Robinson and Decker-Walters 1997). Botanical diversity in melon cultivated morphotypes is of interest because of the specific biological properties and offers a unique opportunity to perform basic research for understanding various biological properties such as fruit quality, disease resistance and sex expression (Sebastian et al. 2010; Nimmakayala et al. 2016).

Watermelon ($2n = 2x = 22$) (Shimotsuma 1963) belongs to the genus *Citrullus* Schrad. Ex Eckl. et Zeyh. Its seven species thrive in dry regions throughout Africa and Asia and in semi-desert regions from the Atlantic Islands eastwards to Afghanistan and Pakistan (Jeffrey 1967). *Citrullus lanatus* Matsum. and Nakai, the common sweet watermelon, is indigenous to north Africa (Wasylikowa and Van Der Veen 2004; Chomicki and Renner 2014; Paris 2015) and may be derived from the 'egusi' melon *C. mucospermus* Fursa (Chomicki and Renner 2014). In contrast, the citron or tsamma melon (*C. amarus* Schrad.) is native to southern Africa. Genetic diversity within the *Citrullus* species provide a resources of genes conferring resistance to a numerous fungal, oomycete and viral disases, as well as resistances to nematodes and several insect pests (Levi et al. 2016).

The genus *Cucurbita* ($2n = 2x = 40$) is native to the Americas and found in the wild from the United States to Argentina (Gong et al. 2013). Five species of *Cucurbita* known as pumpkins and squash have been cultivated for millennia in the Americas, mostly for their edible fruits (Gong et al. 2013). *Cucurbita* is one of 95 genera of the gourd family Cucurbitaceae (Schaefer and Renner 2011). The five cultivated species have different native ranges and climactic adaptations (de Oliveira et al. 2016). They were distributed during cultivation differently, usually allopatrically, throughout all but the coldest parts of the Americas in pre-Columbian times, from North America to South America and from coastal lowland regions to interior highland regions (Zheng et al. 2013). Archaeological remains of *C. pepo* and *C. argyrosperma* have been found at sites in North America, and *C. moschata*, *C. maxima*, and *C. ficifolia* were found in South America (Whitaker and Cutler 1965; Fritz 1994; Lira and Montes 1994; Kong et al. 2014; Paris et al. 2015).

The completion of reference genome sequences for many important crops and the ability to perform resequencing related genomes is revolutionizing crop plant comparative genomics, including for the Cucurbitaceae, for which draft sequences are currently available for cucumber, melon and watermelon (Huang et al. 2009; Guo et al. 2013; Garcia-Mas et al. 2012). These genomes provide critical resources for comparative plant genomics. Examination of similarities and divergences among the genomes of various taxa belonging to crucial nodes of phylogenies can uncover the functional regions of genome, structural variants, inversions and translocations among the genomes (Caicedo and Purugganan 2005; Chaney et al. 2016; Gerats and Vandebussche 2005; Morrell et al. 2012). The recent increase in genomic data is also revealing an unexpected perspective of gene loss as a pervasive source of genetic variation that can cause adaptive phenotypic diversity (Albalat and Canestro 2016). Recent advances in low-cost mapping tools such as improved optics, informatics tools for optical mapping and creative innovations to resolve structural variants have made genome-mapping technology more widely available (Chaney et al. 2016) and can be used for cucurbit comparative genome studies in future.

The genome size of watermelon, melon, cucumber and pumpkin is 425, 454, 367 and 502 Mbp, respectively (Arumuganathan and Earle 1991), and considered small as compared with other crops such as wheat (15,966 Mbp), tomato (907 Mbp), cotton (2500 Mbp), onion (15,290 Mbp), pepper (3420 Mbp) and corn (2716 Mbp). Use of the melon, watermelon and cucumber genome sequences has allowed for an extensive phylogenetic comparison of cucurbit species (Huang et al. 2009; Guo et al. 2013; Garcia-Mas et al. 2012). The genome sequences and genetic maps are excellent tools for understanding the genome structure and evolution of various species with different chromosome number (melon, $2n = 2x = 24$; cucumber, $2n = 2x = 14$; watermelon $2n = 2x = 22$ and pumpkin, $2n = 2x = 40$) as will be described in the following sections.

Syntenic Relationships Among the Cucurbit Genomes

While several synteny maps are available many important plant families such as Solanaceae and grasses (refs), syntenic relationships among Cucurbitaceae remain to be resolved. However, numerous recent studies have provided insight into their relationships. Yang et al. (2012) investigated genetic differentiation between *C. sativus* var. *sativus* and the wild *C. sativus* var. *hardwickii* by comparative fluorescence *in situ* hybridization analysis of pachytene chromosomes with selected markers from the genetic map and draft genome assembly. This study revealed significant differences in the amount and distribution of heterochromatin, as well as chromosomal rearrangements, between the two taxa. In particular, six inversions, five paracentric and one pericentric, were revealed in chromosomes 4, 5 and 7. Comparison of the order of fosmid loci of selected markers on chromosome 7 of cultivated and wild cucumbers and the syntenic melon chromosome 1 suggested that the paracentric inversion in this chromosome occurred during domestication of cucumber.

These results supported the sub-species status of these two cucumber taxa and suggest that *C. sativus* var. *hardwickii* is the progenitor of cultivated cucumber.

After sequencing the cucumber genome, Huang et al. (2009) proposed that five cucumber chromosomes arose from a fusion of ten ancestral chromosomes after divergence from *C. melo*. The authors reported that 348/522 (66.7%) melon genetic markers and 136/232 (58.6%) watermelon genetic markers were aligned on the cucumber chromosomes. The comparison revealed cucumber chromosome 7 corresponds to melon chromosome 1 and watermelon group 7. Li et al. (2011a) constructed a consensus melon linkage map derived from two previous genetic maps with the largest number of cross-species cucumber molecular markers and identified that melon chromosome 1 was syntenic with cucumber chromosome 7. Furthermore, melon chromosomes 2 and 12 were syntenic with cucumber chromosome 1, melon chromosomes 4 and 6 with cucumber chromosome 3, and melon chromosomes 9 and 10 with cucumber chromosome 5. Similarly, the 3 melon chromosomes 3, 8, and 11 contained blocks that were syntenic with 2 cucumber chromosomes, 2 + 6, 4 + 6, and 2 + 6, respectively. This study further concluded that the arrangement of melon syntenic blocks across the seven cucumber chromosomes indicates that cucumber chromosome evolution is more complex than simple chromosome fusions. For instance, cucumber chromosome 7 was homologous to melon chromosome 1 along its entire length. Cucumber chromosomes 2 and 6 each contained 3 syntenic blocks detected in melon chromosomes 5 + 11 + 3, and 3 + 11 + 13, respectively, and the remaining 4 cucumber chromosomes (1, 3, 4, and 5) were syntenic with 2 melon chromosomes but differed in patterns of arrangement of melon syntenic blocks. Cucumber chromosome 1 was syntenic with melon chromosomes 2 and 12, whereas cucumber chromosome 5 was syntenic with melon chromosomes 9 and 10. In both cases, the syntenic blocks from the 2 melon chromosomes were arranged alternatively along each cucumber chromosome. In contrast, the syntenic blocks residing in melon chromosomes 6 and 4 were in a side-by-side alignment in cucumber chromosome 3. Finally, cucumber chromosome 4 housed syntenic blocks of melon chromosomes 7 and 8. Taken together, these syntenic patterns were suggestive of a complex history of chromosomal structure changes during cucumber evolution.

Garcia-Mas et al. (2012) compared an alignment of melon and cucumber genomes synteny to detect shorter regions of rearrangements that were not previously noted, to confirm most of the previously reported ancestral fusions of five melon chromosome pairs in cucumber and several inter- and intra-chromosome rearrangements. This study confirmed findings of Li et al. (2011a) that melon LGI corresponded to cucumber chromosome 7, with higher resolution of several inversions and an increase in the total chromosome size (35.8 vs. 19.2 Mb). Likewise, this study noted that melon LGIV and LGVI were 30.4 and 29.8 Mb, whereas their putative fusion in cucumber was chromosome 3 (39.7 Mb). The first distal 8.5 and 5 Mb of melon LGIV and cucumber chromosome 3, respectively, are highly collinear and melon shows a progressive increase in size toward the centromere because of transposon amplification. Garcia-Mas et al. (2012) identified 19,377 one-to-one ortholog pairs between melon and cucumber, yielding 497 orthologous syntenic

blocks. Further refinement of the physical maps and sequencing of other *Cucumis* species may shed additional light on the genome structure of the ancestor of cucumber and melon.

Guo et al. (2013) analyzed the syntenic relationships between watermelon, cucumber, melon and grape to identify 3543 orthologous relationships covering 60% of the watermelon genome. This study further resolved complicated syntenic patterns using detailed chromosome-to-chromosome relationships within the Cucurbitaceae family and identified orthologous chromosomes between watermelon, cucumber and melon. The insights of high degree of complexity of chromosomal evolution and rearrangement by using chromosome-to-chromosome orthologous relationships unveiled genomic relationships of these three important crop species of the Cucurbitaceae family. Integration of independent analyses of duplications within, and syntenies among, the four eudicot genomes (watermelon, cucumber, melon and grape) led to the precise characterization in watermelon of the seven paleotriplications identified recently as the basis for defining seven ancestral chromosomal groups in eudicots (Abrouk et al. 2010). With the ancestral hexaploidization (γ) reported for the eudicots, Guo et al. (2013) proposed an evolutionary scenario that has shaped the 11 watermelon chromosomes from the 7-chromosome eudicot ancestors through the 21 paleohexaploid intermediates. The authors suggested that the transition from the 21-chromosome eudicot intermediate ancestors involved 81 fissions and 91 fusions to reach the modern 11-chromosome structure of watermelon, represented as a mosaic of 102 ancestral blocks in the watermelon genome.

Genome Duplications

Ancient whole-genome duplications (WGDs), also referred to as paleopolyploidizations, have been reported in most evolutionary lineages. Vanneste et al. (2014) performed a Bayesian evolutionary analysis of 38 full genome sequences and three transcriptome assemblies to note clustering of angiosperm paleopolyploidizations around the Cretaceous–Paleogene (K–Pg) extinction event, about 66 million years ago. This study further demonstrated a strongly nonrandom pattern of genome duplications over time, with many WGDs clustering around the K–Pg boundary. With the increase in number of available plant genomes described, the observation of WGD events will help in understanding their evolution. In cucurbits, the description of the genome sequence of additional species will help determine whether the lack of a recent WGD is unique to this lineage (Huang et al. 2009; Guo et al. 2013; Garcia-Mas et al. 2012). Traces of duplications observed in cucumber, melon and watermelon may correspond to the ancestral paleo-hexaploidization that occurred after the divergence of monocots and dicots, with subsequent genome rearrangements and genome size reduction. Transposable elements have accumulated to a greater extent in melon than cucumber, with peak activity about 2 Mya, which suggests that the larger genome size of melon, probably to a large extent, may be due to transposon amplification. However, loss of chromosome fragments during

chromosome fusion in cucumber may also explain the larger melon genome. Melon and cucumber diverged only around 10 Mya and represent an interesting evolution relating to differences in genome size and chromosome number (450 vs. 367 Mb and $x = 12$ vs. $x = 7$).

WGD is common in angiosperm plants and produces a tremendous source of raw material for gene genesis. Previous research has revealed a paleohexaploidy (Υ) event in the common ancestor of *Arabidopsis thaliana* and grapevine after the divergence of monocotyledons and dicotyledons (Jaillon et al. 2007; Bowers et al. 2003). Subsequently, two WGDs (α and β) occurred in *Arabidopsis* and one (ρ) in poplar, with no recent WGD in grapevine or papaya (Tuskan et al. 2006). Rice underwent an ancient WGD (Yu et al. 2005). A collinear gene-order analysis of the cucumber genome revealed no recent WGD and only a few segmental duplication events (Huang et al. 2009). A distance-transversion rate at fourfold degenerate sites (4DTV method) was used to analyze paralogous gene pairs between syntenic blocks for *Arabidopsis* and cucumber (Huang et al. 2009). Two peaks (~ 0.06 and ~ 0.25) in *Arabidopsis* support the two recent WGDs. Cucumber showed ancient duplication events (peak at ~ 0.60) but not a recent WGD. This lack of recurrent WGD in the small cucumber genome provides an important complement to the grapevine and papaya genomes to study ancestral forms and arrangements of plant genes. Duplication analysis of entire phylomes has been used to confirm ancient WGD events that represent duplication peaks in the corresponding evolutionary periods (Huang et al. 2009). Melon results were consistent with the absence of WGD in the lineages leading to *C. melo* (Garcia-Mas et al. 2012). In the watermelon genome, Guo et al. (2013) identified seven major triplications that corresponded to 302 paralogous relationships covering 29% of the genome. This event would confirm a speciation event in the ancestral cucurbit genome 15–23 Mya.

Gene Prediction and Annotation

Huang et al. (2009) sequenced the whole genome of cucumber (9930V1.0) to identify 26,682 genes with a mean coding sequence size of 1046 bp and a mean of 4.39 exons per gene. Gene model prediction in this study was supported by three gene prediction methods, of which 25% had both ab initio prediction and homology-based evidence, and 7.4% had ab initio prediction supported by transcriptome datasets. In addition, 292 rRNA fragments and 699 tRNA, 238 small nucleolar RNA, 192 small nuclear RNA and 171 microRNA genes were revealed. The cucumber genes represent 15,669 families; 4362 are unique to cucumber, with 3784 single-gene families. Li et al. (2011b) improved annotation of protein coding genes from extensive RNAseq for ten different tissues of 9930 to identify 3434 lesser genes after removal of bacterial genes and corrected protein-coding structures for eight, 700 genes and identified ~ 5200 new genes to Version 2.0. The annotation of the melon genome predicted 27,427 genes with 34,843 predicted transcripts encoding 32,487 predicted polypeptides (Garcia-Mas et al. 2012). The average gene size for melon is 2776 bp, with 5.85 exons per gene, similar to

Arabidopsis, and a density of 7.3 genes per 100 Kb. A total of 16,120 genes (58.7%) had exons supported by ESTs, totaling 18,948 genes (69.1%) supported by transcript and/or a protein alignment. A total of 1253 noncoding RNA genes were identified in the melon genome. Of the 140 potential microRNA loci identified, 122 corresponded to 35 known plant miRNA families. In watermelon, among 23,440 high-confidence protein-coding genes, 85% had transcriptome support. In addition, 123 ribosomal RNA, 789 transfer RNA, 335 small nuclear RNA and 141 miRNA genes were located in the watermelon cultivated genome, which is comparable to the other sequenced genomes of cucurbits.

Transposon Annotation

By using homology and structure-based searches, Garcia-Mas et al. (2012) identified 323 transposable element representatives in the melon genome belonging to known superfamilies. With these sequences used for melon genome analysis, 73,787 copies of various superfamilies were found to occupy 19.7% of the genome space. Use of the same annotation pipeline to compare retrotransposons in the Gy14 cucumber genome revealed that retrotransposons represented 1.5% of the genome, significantly less than in melon, which suggests that retrotransposon activity was greater and more recent in the melon than cucumber lineage. Garcia-Mas et al. (2012) further compared CACTA, MULE and PIF/Harbinger, the three most represented superfamilies, to show 10× more amplification in melon for CACTA, 47× for MULE, and 3.8× for PIF, thereby confirming a divergence of 10.1 Mya between melon and cucumber. Guo et al. (2013) identified 159.8 Mb (45.2%) of the assembled watermelon genome as transposable element repeats; 68.3% could be annotated with known repeat families. Transposable element divergence rates peaked at 32%. The authors further identified 920 (7.8 Mb) full-length LTR retrotransposons in the watermelon genome. Over the past 4.5 million years, LTR retrotransposons accumulated much faster in watermelon than cucumber, so the overall difference in their genome sizes may reflect the differential LTR retrotransposon accumulation (Guo et al. 2013).

Disease Resistance Genes

Only 61 nucleotide binding site (NBS)-containing resistance (NBS-R) genes have been identified in cucumber, similar to papaya (55), but only a fraction of that found in Arabidopsis (200), poplar (398) and rice (600) (Huang et al. 2009). Distribution of NBS genes on cucumber chromosomes is non-random, with only five genes located on chromosomes 1, 6 and 7 and 20 on chromosome 2. Three-quarters of the NBS genes are located within 11 clusters, which indicates that they evolved by tandem duplication, similar to other known plant genomes. A total of 411 putative disease resistance genes were identified in the melon genome (Garcia-Mas et al.

2012); 81 represented NBS, leucine-rich repeat (LRR) and Toll-interleukin receptor (TIR) domains that were non-randomly distributed, and 45% of the NBS-LRR genes were grouped within nine clusters similar to cucumber. In watermelon, 44 NBS-LRR-TIR genes (18 TIRs and 26 coiled-coil NBS-LRR-encoding genes) were identified (Guo et al. 2013). In cucumber, 23 lipoxygenase (LOX) pathway genes were identified; such genes have an important role in defense and pest resistance by generating short-chain aldehydes and alcohols (Huang et al. 2009). In watermelon, 45 members belonging to the LOX gene family were arranged in two tandem arrays. Among the 197 receptor-like genes in the watermelon genome, 35 encode receptor-like proteins lacking a kinase domain in addition to the extracellular LRR and transmembrane domains (Guo et al. 2013). In melon, 290 transmembrane receptors, 161 receptor-like kinases (RLKs), 19 kinases containing an additional anti-fungal protein ginkbilobin-2 domain, and 110 receptor-like proteins genes were also documented (Garcia-Mas et al. 2012).

Conclusions

Use of genome sequences is becoming a strategic tool for gene expression and genome-wide association studies to accelerate plant breeding and basic biological research. Comparative cucurbit genome analysis involves examining the similarities and unique differences to shed light on the underlying genome evolution and identify economically important traits. Defining syntenic blocks by comparative mapping has shown that numerous alterations in diverse genomes contributed to genetic diversity among plants. Over time, chromosomes are broken, reassembled, partially or wholly duplicated, and even eliminated, ultimately resulting in reproductive isolation and speciation (Koenig and Weigel 2015; Hall et al. 2002). For example, comparative genome analysis to understand conserved syntenic blocks in cucurbit genomes holds promise for clarifying the selection pressures driving genetic changes. Modern genome mapping strategies such as optical mapping, which uses microscopic imaging to produce ordered restriction enzyme recognition site maps from a single linearized DNA molecule, allows for detecting DNA with resolution of 1 Kb to several mega base pairs (Chaney et al. 2016). Genomic alterations are an important source of genetic and phenotypic diversity. For example, structural variations that include insertions, deletions, duplications, inversions and translocations resolved with optical mapping strategies have been associated with stress tolerance, resistance, increase in yields, reproductive morphology, adaptation and speciation (Chaney et al. 2016). Such investigations will elucidate alterations at the level of the whole genome, for diversifying cultivars with narrow genetic backgrounds. Comparative maps of all the other cereals have been useful to bridge information from one species to the other, of immense use for breeding, ecology and molecular biology. The development of whole-genome sequence drafts has provided a foundation for widening a narrow genetic background, marker-assisted selection and to understand intricate genome rearrangements to study genetics and breed improved

varieties in less-important crops. Furthermore, by using the reference maps of various cucurbits that are anchored with several other crop genomes, we can identify major genes affecting agronomic characters found in different species.

Acknowledgments Funding support is provided by Gus R. Douglass Institute (Evans Allen Project to P.N.) and USDA-NIFA (2010-02247 2012-02511 and 2015-09260).

References

- Abrouk M, Murat F, Pont C, Messing J, Jackson S, Faraut T, et al. Palaeogenomics of plants: synteny-based modelling of extinct ancestors. *Trends Plant Sci.* 2010;15(9):479–87.
- Albalat R, Canestro C. Evolution by gene loss. *Nat Rev Genet.* 2016;17(7):379–91. doi:[10.1038/nrg.2016.39](https://doi.org/10.1038/nrg.2016.39).
- Arumuganathan K, Earle E. Nuclear DNA content of some important plant species. *Plant Mol Biol Rep.* 1991;9(3):208–18.
- Bates, D. M., Robinson, R. W., & Jeffrey, C. (Eds.). *Biology and utilization of the Cucurbitaceae.* New York, NY: Cornell University Press; (1990).
- Bates DM, Robinson RW. Cucumbers, melons and water-melons: *Cucumis* and *Citrullus* (Cucurbitaceae). In: Smartt J, Simmonds NW, editors. *Evolution of crop plants.* Harlow, UK: Longman Scientific and Technical; 1995.
- Bates DM, Robinson RW, Jeffrey C. *Biology and utilization of the Cucurbitaceae.* New York: Cornell University Press; 1990.
- Bowers JE, Chapman BA, Rong J, Paterson AH. Unravelling angiosperm genome evolution by phylogenetic analysis of chromosomal duplication events. *Nature.* 2003;422(6930):433–8.
- Caicedo AL, Purugganan MD. Comparative plant genomics. *Frontiers and prospects. Plant Physiol.* 2005;138(2):545–7. doi:[10.1104/pp.104.900148](https://doi.org/10.1104/pp.104.900148).
- Chaney L, Sharp AR, Evans CR, Udall JA. Genome mapping in plant comparative genomics. *Trends Plant Sci.* 2016;21:770–80.
- Chomicki G, Renner S. Watermelon origin solved with molecular phylogenetics including Linnaean material: another example of museomics. *New Phytol.* 2014;205:526–32.
- de Oliveira RL, Gonçalves LSA, Rodrigues R, Baba VY, Sudré CP, dos Santos MH, et al. Genetic divergence among pumpkin landraces. *Semin Ciênc Agrár.* 2016;37(2):547–56.
- Fritz GJ. Precolumbian *Cucurbita argyrosperma* ssp. *argyrosperma* (Cucurbitaceae) in the eastern woodlands of North America. *Econ Bot.* 1994;48(3):280–92.
- Garcia-Mas J, Benjak A, Sanseverino W, Bourgeois M, Mir G, González VM, et al. The genome of melon (*Cucumis melo* L.). *Proc Natl Acad Sci U S A.* 2012;109(29):11872–7. doi:[10.1073/pnas.1205415109](https://doi.org/10.1073/pnas.1205415109).
- Gerats T, Vandenbussche M. A model system for comparative research: *Petunia*. *Trends Plant Sci.* 2005;10(5):251–6. doi:[10.1016/j.tplants.2005.03.005](https://doi.org/10.1016/j.tplants.2005.03.005).
- Gong L, Paris HS, Stift G, Pachner M, Vollmann J, Lelley T. Genetic relationships and evolution in *Cucurbita* as viewed with simple sequence repeat polymorphisms: the centrality of *C. okechobeensis*. *Genet Resour Crop Evol.* 2013;60(4):1531–46. doi:[10.1007/s10722-012-9940-5](https://doi.org/10.1007/s10722-012-9940-5).
- Guo S, Zhang J, Sun H, Salse J, Lucas WJ, Zhang H, et al. The draft genome of watermelon (*Citrullus lanatus*) and resequencing of 20 diverse accessions. *Nat Genet.* 2013;45(1):51–8.. <http://www.nature.com/ng/journal/v45/n1/abs/ng.2470.html#supplementary-information>
- Hall AE, Fiebig A, Preuss D. Beyond the Arabidopsis genome: opportunities for comparative genomics. *Plant Physiol.* 2002;129(4):1439–47.
- Huang S, Li R, Zhang Z, Li L, Gu X, Fan W, et al. The genome of the cucumber, *Cucumis sativus* L. *Nat Genet.* 2009;41 doi:[10.1038/ng.475](https://doi.org/10.1038/ng.475).

- Jaillon O, Aury JM, Noel B, Policriti A, Clepet C, Casagrande A, et al. The grapevine genome sequence suggests ancestral hexaploidization in major angiosperm phyla. *Nature*. 2007;449(7161):463–7. doi:[10.1038/nature06148](https://doi.org/10.1038/nature06148).
- Jeffrey C. A review of the Cucurbitaceae. *Bot J Linn Soc*. 1980;81(3):233–47.
- Jeffrey C. Cucurbitaceae. In: Milne-Redhead E, Polhill RM, editors. *Flora of tropical East Africa*. London: Published under the Authority of the Minister for Overseas Development by the Crown Agents for Oversea Governments and Administration; 1967.
- Knerr L, Staub J. Inheritance and linkage relationships of isozyme loci in cucumber (*Cucumis sativus* L.). *Theor Appl Genet*. 1992;84(1–2):217–24.
- Knerr L, Staub J, Holder D, May B. Genetic diversity in *Cucumis sativus* L. assessed by variation at 18 allozyme coding loci. *Theor Appl Genet*. 1989;78(1):119–28.
- Koenig D, Weigel D. Beyond the thale: comparative genomics and genetics of Arabidopsis relatives. *Nat Rev Genet*. 2015;16(5):285–98.
- Kong Q, Chen J, Liu Y, Ma Y, Liu P, Wu S, et al. Genetic diversity of *Cucurbita* rootstock germplasm as assessed using simple sequence repeat markers. *Sci Hortic*. 2014;175:150–5.
- Levi A, Jarret R, Kousik S, Wechter WP, Nimakayala P, Reddy U. Genetic resources of watermelon. In: Grumet R, Katzir N, Garcia-Mas J, editors. *Genetics and genomics of the Cucurbitaceae*. New York: Springer Intl Pub AG; 2016.
- Li D, Cuevas HE, Yang L, Li Y, Garcia-Mas J, Zalapa J, et al. Syntenic relationships between cucumber (*Cucumis sativus* L.) and melon (*C. melo* L.) chromosomes as revealed by comparative genetic mapping. *BMC Genomics*. 2011a;12(1):1–14. doi:[10.1186/1471-2164-12-396](https://doi.org/10.1186/1471-2164-12-396).
- Li Z, Zhang Z, Yan P, Huang S, Fei Z, Lin K. RNA-Seq improves annotation of protein-coding genes in the cucumber genome. *BMC Genomics*. 2011b;12(1):540. doi:[10.1186/1471-2164-12-540](https://doi.org/10.1186/1471-2164-12-540).
- Lira R, Montes S. Cucurbits (*Cucurbita* spp.). In: Hernandez Bermejo JE, Leon J, editors. *Neglected crops: 1492 from a different perspective*. Rome: Plant Production and Protection Series; 1994. p. 63–77.
- Lv J, Qi J, Shi Q, Shen D, Zhang S, Shao G, et al. Genetic diversity and population structure of cucumber (*Cucumis sativus* L.). *PLoS One*. 2012;7(10):e46919.
- McCreight JD. Cultivation and uses of cucurbits. In: Grumet R, Katzir N, Garcia-Mas J, editors. *Genetics and genomics of the Cucurbitaceae*. New York: Springer Intl Pub AG; 2016.
- Morrell PL, Buckler ES, Ross-Ibarra J. Crop genomics: advances and applications. *Nat Rev Genet*. 2012;13(2):85–96.
- Nagele RP, Wehner TC. Genetic resources of cucumber. In: Grumet R, Katzir N, Garcia-Mas J, editors. *Genetics and genomics of the Cucurbitaceae*. New York: Springer Intl Pub AG; 2016.
- Nee M. The domestication of *Cucurbita* (Cucurbitaceae). *Econ Bot*. 1990;44(3):56–68.
- Nimmakayala P, Tomason YR, Abburi VL, Alvarado A, Saminathan T, Vajja VG, et al. Genome-wide differentiation of various melon horticultural groups for use in GWAS for fruit firmness and construction of a high resolution genetic map. *Front Plant Sci*. 2016;7:1437. doi:[10.3389/fpls.2016.01437](https://doi.org/10.3389/fpls.2016.01437).
- Pandey S, Ansari WA, Mishra VK, Singh AK, Singh M. Genetic diversity in Indian cucumber based on microsatellite and morphological markers. *Biochem Syst Ecol*. 2013;51:19–27.
- Paris HS, Doron-Faigenboim A, Reddy UK, Donahoo R, Levi A. Genetic relationships in *Cucurbita pepo* (pumpkin, squash, gourd) as viewed with high frequency oligonucleotide–targeting active gene (HFO–TAG) markers. *Genet Resour Crop Evol*. 2015;62(7):1095–111.
- Paris H. Origin and emergence of the sweet dessert watermelon, *Citrullus lanatus*. *Ann Bot*. 2015;116:133–48.
- Perl-Treves R, Zamir D, Navot N, Galun E. Phylogeny of *Cucumis* based on isozyme variability and its comparison with plastome phylogeny. *Theor Appl Genet*. 1985;71(3):430–6.
- Pitrat M, Hanelt P, Hammer K. Some comments on infraspecific classification of cultivars of melon. *Acta Hort*. 2000;510:29–36.
- Qi J, Liu X, Shen D, Miao H, Xie B, Li X, et al. A genomic variation map provides insights into the genetic basis of cucumber domestication and diversity. *Nat Genet*. 2013;45(12):1510–5.
- Robinson R, Decker-Walters D. *Cucurbits: crop production science in horticulture*. New York: CAB International; 1997.

- Schaefer H, Renner SS. Phylogenetic relationships in the order Cucurbitales and a new classification of the gourd family (Cucurbitaceae). *Taxon*. 2011;60:122–38.
- Sebastian P, Schaefer H, Telford IR, Renner SS. Cucumber (*Cucumis sativus*) and melon (*C. melo*) have numerous wild relatives, and the sister species of melon is from Australia. *Proc Natl Acad Sci U S A*. 2010;107:14269–73. doi:10.1073/pnas.1005338107.
- Shimotsuma M. Cytogenetic and evolutionary studies in the genus *Citrullus*. *Seiken Ziho*. 1963;15:23–4.
- Tuskan GA, Difazio S, Jansson S, Bohlmann J, Grigoriev I, Hellsten U, et al. The genome of black cottonwood, *Populus trichocarpa* (Torr. & Gray). *Science*. 2006;313(5793):1596–604.
- Vanneste K, Baele G, Maere S, Van de Peer Y. Analysis of 41 plant genomes supports a wave of successful genome duplications in association with the Cretaceous–Paleogene boundary. *Genome Res*. 2014;24(8):1334–47.
- Wasylikowa K, Van Der Veen M. An archaeobotanical contribution to the history of watermelon, *Citrullus lanatus* (Thunb.) Matsum. & Nakai (syn. *C. vulgaris* Schrad.). *Veg Hist Archaeobotany*. 2004;3:213–7.
- Whitaker T, Bemis W, Simmonds N. *Evolution of crop plants*. London: Longman Group Limited; 1976.
- Whitaker TW, Cutler HC. Cucurbits and cultures in the Americas. *Econ Bot*. 1965;19(4):344–9.
- Yang L, Koo DH, Li Y, Zhang X, Luan F, Havey MJ, et al. Chromosome rearrangements during domestication of cucumber as revealed by high-density genetic mapping and draft genome assembly. *Plant J*. 2012;71(6):895–906.
- Yu J, Wang J, Lin W, Li S, Li H, Zhou J, et al. The Genomes of *Oryza sativa*: a history of duplications. *PLoS Biol*. 2005;3(2):1.
- Zheng YH, Alverson AJ, Wang QF, Palmer JD. Chloroplast phylogeny of Cucurbita: evolution of the domesticated and wild species. *J Syst Evol*. 2013;51(3):326–34.