Engineering compact *Physalis peruviana* (goldenberry) to promote its potential as a global crop

Miguel Santo Domingo[1*], Blaine Fitzgerald[1,2*], Gina M. Robitaille[1,2], Srividya Ramakrishnan[3], Kerry Swartwood[4], Nicholas G. Karavolias[1,2], Michael C. Schatz[3+], Joyce Van Eck[4,5+], Zachary B. Lippman[1,2+]

1. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.
2. Howard Hughes Medical Institute, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, USA.
3. Department of Computer Science, Johns Hopkins University, Baltimore, MD, USA.
4. Boyce Thompson Institute, Ithaca, NY, USA.
5. Plant Breeding and Genetics Section, Cornell University, Ithaca, NY, USA

*These authors contributed equally to this work

+Corresponding author: jv27@cornell.edu, mschatz@cs.jhu.edu, lippman@cshl.edu

Present address:

Nicholas G. Karavolias, University of California, Berkeley, CA, USA

ORCID

Miguel Santo Domingo: 0000-0001-7447-9331

Blaine Fitzgerald: 0009-0004-8503-4978

Gina M. Robitaille: 0000-0002-7654-5002

Srividya Ramakrishnan: 0000-0001-8023-4752

Kerry Swartwood: 0009-0001-9700-8302

Nicholas G. Karavolias: 0000-0003-1213-771X

Michael C. Schatz: 0000-0002-4118-4446

Joyce Van Eck: 0000-0002-8005-365X

Zachary B. Lippman: 0000-0001-7668-9025

SOCIETAL STATEMENT

EN - Goldenberry (*Physalis peruviana*) produces sweet, nutritionally-rich berries, yet like many minor crops, is cultivated in limited geographical regions and has not been a focus of breeding programs for trait enhancement. Leveraging knowledge of plant architecture-related traits from related species, we used CRISPR/Cas9-mediated gene editing to generate a compact ideotype to advance future breeding efforts and agricultural production. Goldenberry growers will benefit from these compact versions because it optimizes per plot yield, facilitating larger-scale production to meet rising consumer popularity and demand.

SP - La uchuva (*Physalis peruviana*) produce frutos dulces y ricos en nutrientes, pero, igual que muchos cultivos minoritarios, se cultiva en zonas geográficas limitadas y no ha sufrido un proceso de mejora. Aprovechando conocimientos sobre rasgos relacionados con la arquitectura vegetal de especies relacionadas, hemos usado edición génica mediante CRISPR/Cas9 para generar un ideotipo compacto para promover futuros esfuerzos en su mejora y en producción agrícola. Los productores de uchuva se podrán beneficiar de estas versiones compactas ya que optimiza el rendimiento por parcela, facilitando así la producción a una mayor escala para cubrir la creciente popularidad y demanda de los consumidores.

KEYWORDS

Solanaceae, orphan crop, goldenberry, gene editing, plant architecture

INTRODUCTION

A limited number of species dominate global crop production (FAO, 2024) resulting in a fragile food supply. However, minor crops, sometimes referred to as indigenous, local, or orphan crops, are relied upon in regional contexts, where, although not fully domesticated, they provide nutritional dietary diversity. Some minor crops, such as quinoa and dragon fruit, have become popular with worldwide consumers because of increased interest in new flavors, textures, and their health benefits.

Goldenberry (*Physalis peruviana*), a member of the crop-rich Solanaceae family, is a minor crop indigenous to South America that is often touted as a superfood because of its high nutritional value (Shenstone, Lippman and Van Eck, 2020). It produces round, smooth, yellow-orange colored berries that are surrounded by a husk (inflated calyx), similar to other species in the *Physalis* genus, and its fruit flavor is reminiscent of other tropical fruits such as pineapple and mango. Depending on the region where it is grown, goldenberry is known by locally recognized names such as uchuva (Colombia) and Cape gooseberry (South Africa). While goldenberries have been historically grown in the Andean region near its evolutionary origin, they have undergone little domestication, as evidenced by its large, wild, unmanageable growth habit that complicates large-scale production (Flores Roncancio et al., 2020). Currently, Colombia is the main producer of goldenberries, producing more than 20,000 tons of fresh fruit in 2022 with 40% exported (Resilience BV, 2024); however, additional production occurs in Peru, Ecuador, South Africa and India. Improvement of its growth habit to make it more manageable in large-scale production would have a positive impact on dietary diversification with greater availability of this highly nutritious fruit.

Modification of traits in many plant species, including minor and major solanaceous crops, has been readily achievable by the application of genome editing technologies (Lemmon et al., 2018; Kwon et al., 2020). For example, we showed that CRISPR-engineered mutations of the classical stem length regulator *ERECTA* (Torii et al., 1996) in groundcherry (*Physalis grisea*) and tomato (*Solanum lycopersicum*) resulted in a more compact (determinate) growth habit (Kwon et al., 2020). Based on this outcome, engineering mutations of the two *ERECTA* genes in tetraploid goldenberry could also result in plants with a modified, and more agriculturally desirable, plant architecture.

We report here the development of high-quality genome assemblies alongside transformation and genome-editing approaches, which we used to rapidly modify goldenberry plant architecture by targeting the two sub-genome copies of *ERECTA.* To give breeders and growers a head-start with our improved, compact goldenberry without any regulatory delays, we have already acquired clearance by the US Department of Agriculture (USDA) to document that our varieties are not plant pests and no plant pests remain integrated into the plant genome, so that they are not regulated under 7 CFR Part 340. In addition, we also seek approval by the US Food and Drug Administration (FDA) to enable growers to immediately move forward with commercial production promoting rapid adoption and new market opportunities.

ECONOMIC AND BREEDING POTENTIAL OF GOLDENBERRY

Goldenberry fruits are small (4 to 10 grams) (Flores Roncancio et al., 2020) and round with a bright yellow-orange color, and have a unique flavor and nutritional profile (**Figure 1A**) (Flores Roncancio et al., 2020; Shenstone, Lippman and Van Eck, 2020). Their appeal as a new berry for consumers worldwide has expanded their accessibility to markets and grocery stores beyond South America. In particular, in the United States, goldenberries are now sold year round (**Figure 1B**), revealing its high economic potential.

Native to Peru, goldenberry grows wild between 1,500 and 3,000 m in elevation in the Andean countries of Colombia and Peru, where it has been consumed for at least several centuries dating to the Inca Empire (Legge, 1974). After European colonization in the 18th century, goldenberry was brought to South Africa and other parts of the world including India and Australia (**Figure 1C**). Its small, round fruits, are highly nutritious with health beneficial levels of antioxidants, carotenoids, and micronutrients (Ezbach et al., 2018; Haytowitz and Bhagwat, 2010; Gündeşli et al., 2019), which puts goldenberry in the superfood category, together with other commercial berries (**Figure 1D, E**). Similar to other berries, goldenberry is primarily consumed as fresh fruit, however, they are also used to make jams, condiments such as chutney, and are also available as dried fruit (raisins) (Van Eck, 2022). The global growing appeal of berries is evidenced by their expanding production and availability in markets (**Figure 1F**) (FAOSTAT, 2024), positioning goldenberries to become a highly sought after crop by farmers and distributors. However, for its

105 economic value and large-scale agricultural production to be realized, critical trait enhancements
106 (e.g. synchronous ripening, reduction of sticky acylsugars on fruits, elimination of fruit cracking,
107 and a modified plant architecture) are required through plant breeding, including new technologies
108 such as genome editing (**Figure 1G**).

109

110 **ENGINEERING A COMPACT GOLDENBERRY GROWTH HABIT**

111 To improve its growth habit, we focused on engineering mutations in the goldenberry *ERECTA*
112 gene (*Pper-ER*). We based this approach on previous work where we targeted *ERECTA* in sister
113 species tomato and groundcherry to modify plant architecture (Kwon et al., 2020). However,
114 goldenberry presented several challenges. First, as an allotetraploid (4x = 2n = 48) (Sanchez-
115 Betancourt and Núñez Zarantes, 2022), the possibility of two divergent copies of *Pper-ER* (one
116 for each subgenome) in goldenberry makes editing more challenging, as both gene copies might
117 require independent sets of guide RNA (gRNA), coupled with more complex downstream genetic
118 fixation of desired alleles compared to a diploid. Second, there was the need to establish efficient
119 plant regeneration and transformation methods for this species along with generating a high quality
120 reference genome.

121 We generated high-quality genomes for two ecotypes, *India* and *South Africa*. Using long PacBio
122 HiFi reads, we estimated the haploid genome size to be ~1.34 Gbp. Assemblies yielded ~5.3 Gbp
123 genomes, representing the four distinct haplotypes, with an average contig N50 of ~48.0 Mbp and
124 the longest contigs approaching complete chromosomes or chromosome arms (max: 226.54 Mbp)
125 (**Figure 2A**) (**Supplementary Table S1**). Genome completeness analysis showed ~98.3%
126 completeness, with 94.1% of genes duplicated, consistent with tetraploidy. The haplotype resolved
127 assemblies and their corresponding gene duplications enabled us to study haplotypic variation in
128 the genes (**Supplementary Table S2**). Importantly, we detected two highly similar *Pper-ER*
129 genes, *Pper-ER-A* and *Pper-ER-B*, allowing the design of a four-gRNA construct for simultaneous
130 targeting (**Supplementary Table S3**). In parallel, we developed plant regeneration, transformation
131 and gene-editing approaches. Regeneration efficiency was similar in both ecotypes (36% and 37%
132 of explants regenerating at least one shoot in *South Africa* and *India*, respectively), but editing was
133 more efficient in the *India* ecotype. We did not detect edits in *South Africa ERECTA* genes, but

134 we recovered, among others, two null alleles in *Pper-ER* in the *India* ecotype: one base pair
135 deletion in *Pper-er-A* and one base pair insertion in *Pper-er-B*, in the first and third exons,
136 respectively (**Figure 2B**).

137 We then introgressed both null alleles into the *South Africa* line since there is some preference for
138 their fruit flavor. After two rounds of backcrossing the edited lines to *South Africa* and a final self-
139 pollination, we recovered plants with substantially increased compactness, which we named
140 "Erecta" to denote this genotype as a new line (**Figure 2C**). In subsequent generations, we made
141 selections from the compact plant lines based on fruit flavor, resulting in stable, compact varieties
142 of *India*- and *South Africa*-like flavored fruits. Finally, we sequenced the genome of the *South*
143 *Africa* Erecta line to demonstrate absence of the transgene (T-DNA insertion) and as a non-
144 transgenic resource for future development through plant breeding (**Supplementary Table S1,**
145 **Supplementary Table S2, Supplementary Figure S1**).

146 Three-month-old *South Africa* Erecta plants were 35% shorter than *South Africa* parental line
147 plants (86.2 +- 7.5 cm and 132.0 +- 6.3 cm, respectively) (**Figure 2D**, **Supplementary Figure**
148 **S2**). This decrease resulted from a 50% reduction in internode length compared to the parental line
149 (5.0 +- 1.1 cm in Erecta 10.0 += 2.1 cm in *South Africa*), consistent with the effect of *er* mutations
150 in other species (Kwon et al., 2020). Importantly, the fruit fecundity (total amount of fruits
151 produced per plant) was not affected by the mutations (**Supplementary Figure S2D**). However,
152 similar to other *erecta*-edited crops, fruit size was also reduced proportionally; *South Africa* Erecta
153 fruit weighed ~50% of *South Africa* parental line fruit (3.3 +- 0.3 g and 6.1 +- 0.6 g, respectively).

154 Despite the reduced fruit size, the goldenberry Erecta lines are still valuable for breeding and
155 production. Goldenberries currently on the market are sold in two sizes, which we designated as
156 "large" and "small" types (**Figure 1B**). The average fruit weight of the large type is ~7.4 g, while
157 the small type is ~3.8 g (**Supplementary Figure S3**). Our Erecta plants produced, on average, 3.3
158 g fruits (**Figure 2D, Supplementary Figure S3**), only slightly less than those commercially
159 available, making them well-suited for immediate commercial evaluation. Thus, the *South Africa*
160 Erecta germplasm will provide a foundation for subsequent improvement of this flavorful and
161 highly nutritious fruit.

162 In summary, using CRISPR-Cas9-mediated gene editing of the *India* genotype followed by several
163 cycles of introgression and selection into the *South Africa* genotype because of its preferred fruit
164 flavor, we have generated two stably modified compact germplasm with fruits similar to
165 commercially available ones. With this new Erecta germplasm, planting density can be increased,
166 while also avoiding the need for laborious staking and trellising of large, bushy plants (**Figure 3**).

167

168 **IMPACT**

169 We mutated *Pper-ER* and established a pre-breeding program for compact goldenberry plants by
170 integrating information from related major and minor solanaceous crop species, and following an
171 introgression scheme. These new Erecta lines can be used as a starting point to develop new
172 varieties through traditional breeding and possibly additional gene editing to improve other traits,
173 or directly in the production of goldenberries.

174 For further improvement, fruit size is the most obvious next target. As seen in *P. grisea* (Lemmon
175 et al., 2018), the targeting of *PperCLV1* could help restore and potentially increase fruit size in the
176 Erecta lines. Additionally, other genes responsible for tomato fruit characteristics could further
177 customize fruit size and shape (Mauxion, Chevalier and Gonzalez, 2021). Additional
178 improvements include the modification of acylsugar metabolism to reduce fruit stickiness or
179 accelerated flowering for earlier harvests (**Figure 1G**).

180 Apart from goldenberry, several other minor crops could benefit from the presented scheme for
181 improvement of undesirable traits. Already, genomes and transformation approaches have been
182 developed (or are under development) for a range of minor crops, such as passion fruit (Xia et al.,
183 2021; Manders et al., 1994), fonio (Abrouk et al., 2020; Ntui et al., 2017), and groundcherry (Dale
184 et al. 2024) leveraging the potential of gene editing to expand the benefits of local (agri)culture
185 worldwide.

186

187 **AUTHOR CONTRIBUTIONS**

188 MCS, JVE, and ZBL conceived the project and designed the experiments. SR and MCS generated
189 the genomes. KS and JVE developed the transformation methods and generated the edited plants.
190 BF, GMR, and ZBL developed the Erecta genotypes. MSD and BF collected phenotypic data.
191 MSD performed the data analysis and wrote the original draft. All the authors revised and edited
192 the manuscript.

193

194 **ACKNOWLEDGEMENTS**

195 We thank members of the Lippman laboratory for discussions; B. Seman for technical support; T.
196 Mulligan, K. Schlecht, and S. Qiao for assistance with plant care. We thank G. Vogel of Cornell
197 University for guidance on terminology related to the edited lines we developed. This work was
198 supported by the National Science Foundation Plant Genome Research Program grant IOS-
199 2216612 to J.V.E., M.C.S. and Z.B.L. We thank the Indigenous peoples of South America on
200 whose ancestral lands P. peruviana grow.

201

202 **CONFLICT OF INTEREST**

203 All authors declare no competing interests.

204

205 **REFERENCES**

206 Abrouk, Michael, Hanin Ibrahim Ahmed, Philippe Cubry, et al. 2020. "Fonio Millet Genome
207 Unlocks African Orphan Crop Diversity for Agriculture in a Changing Climate." *Nature*
208 *Communications* 11 (1): 4488. https://doi.org/10.1038/s41467-020-18329-4.

209 Dale, Savanah Marie, Elise Tomaszewski, Zachary Lippman, and Joyce Van Eck. 2024.
210 "Engineering the Future of : A Focus on Agricultural Challenges, Model Species Status, and
211 Applied Improvements." *PLANTS, PEOPLE, PLANET* 6 (6): 1249–60.
212 https://doi.org/10.1002/ppp3.10536.

Etzbach, Lara, Anne Pfeiffer, Fabian Weber, and Andreas Schieber. 2018. "Characterization of Carotenoid Profiles in Goldenberry (Physalis Peruviana L.) Fruits at Various Ripening Stages and in Different Plant Tissues by HPLC-DAD-APCI-MS." *Food Chemistry* 245 (April): 508–17. https://doi.org/10.1016/j.foodchem.2017.10.120.

FAO. 2024. *World Food and Agriculture – Statistical Yearbook 2024*. FAO ; https://openknowledge.fao.org/handle/20.500.14283/cd2971en.

FAOSTAT: Food and Agriculture Organization of the United Nations. 2024. From https://www.fao.org/faostat/en/

Flórez Roncancio, Víctor Julio, Gerhard Fischer, and Ángel D. Sora R., eds. 2000. *Producción, poscosecha y exportación de la uchuva (Physalis peruviana L.)*. 1. ed. Universidad Nacional de Colombia, Fac. de Agronomía.

Gündeşli, Muhammet Ali, Nazan Korkmaz, and Volkan Okatan. 2019. "Polyphenol Content and Antioxidant Capacity of Berries: A Review." *International Journal of Agriculture Forestry and Life Sciences* 3 (2): 2.

Haytowitz, D.B., Bhagwat, S.A. 2010. USDA database for the oxygen radical capacity (ORAC) of selected foods, release 2. USDA National Nutrient Database for Standard Reference. Available: http://www.ars.usda.gov/nutrientdata

Kwon, Choon-Tak, Jung Heo, Zachary H. Lemmon, et al. 2020. "Rapid Customization of Solanaceae Fruit Crops for Urban Agriculture." *Nature Biotechnology* 38 (2): 182–88. https://doi.org/10.1038/s41587-019-0361-2.

Legge, A.P., 1974. Notes on the history, cultivation and uses of Physalis peruviana L. Journal of the Royal Horticultural Society, 99(7):310-314.

Lemmon, Zachary H., Nathan T. Reem, Justin Dalrymple, et al. 2018. "Rapid Improvement of Domestication Traits in an Orphan Crop by Genome Editing." *Nature Plants* 4 (10): 766–70. https://doi.org/10.1038/s41477-018-0259-x.

238 Manders, G., W. C. Otoni, F. B. d'Utra Vaz, N. W. Blackball, J. B. Power, and M. R. Davey. 1994.
239 "Transformation of Passionfruit (Passiflora Edulis Fv Flavicarpa Degener.) Using Agrobacterium
240 Tumefaciens." *Plant Cell Reports* 13 (12): 697–702. https://doi.org/10.1007/BF00231627.

241 Mauxion, Jean-Philippe, Christian Chevalier, and Nathalie Gonzalez. 2021. "Complex Cellular
242 and Molecular Events Determining Fruit Size." *Trends in Plant Science* 26 (10): 1023–38.
243 https://doi.org/10.1016/j.tplants.2021.05.008.

244 Ntui, Valentine Otang, Edak Aniedi Uyoh, Ikuo Nakamura, and Masahiro Mii. 2017.
245 "Agrobacterium-Mediated Genetic Transformation of Fonio (Digitaria Exilis (L.) Stapf)." *African*
246 *Journal of Biotechnology* 16 (23): 1302–7. https://doi.org/10.5897/AJB2017.15903.

247 Resilience B.V. *The Goldenberry Value Chain in Colombia: Value Chain Description, Bottlenecks*
248 *and Opportunities.* 2024. Report CST23CO04. The Hague: Netherlands Enterprise Agency, PDF.
249 https://resiliencebv.com/wp-content/uploads/2024/08/CST23CO04-Goldenberry-report-
250 V2_clean.pdf.

251 Sánchez-Betancourt, E., Núñez-Zarantes, V.M. 2022. "Nivel de Ploidía de Plantas de Uchuva,
252 Physalis Peruviana, Obtenidas Mediante Cultivo de Anteras." *Revista Temas Agrarios* 27 (1): 231–
253 45. https://doi.org/10.21897/rta.v27i1.3080.

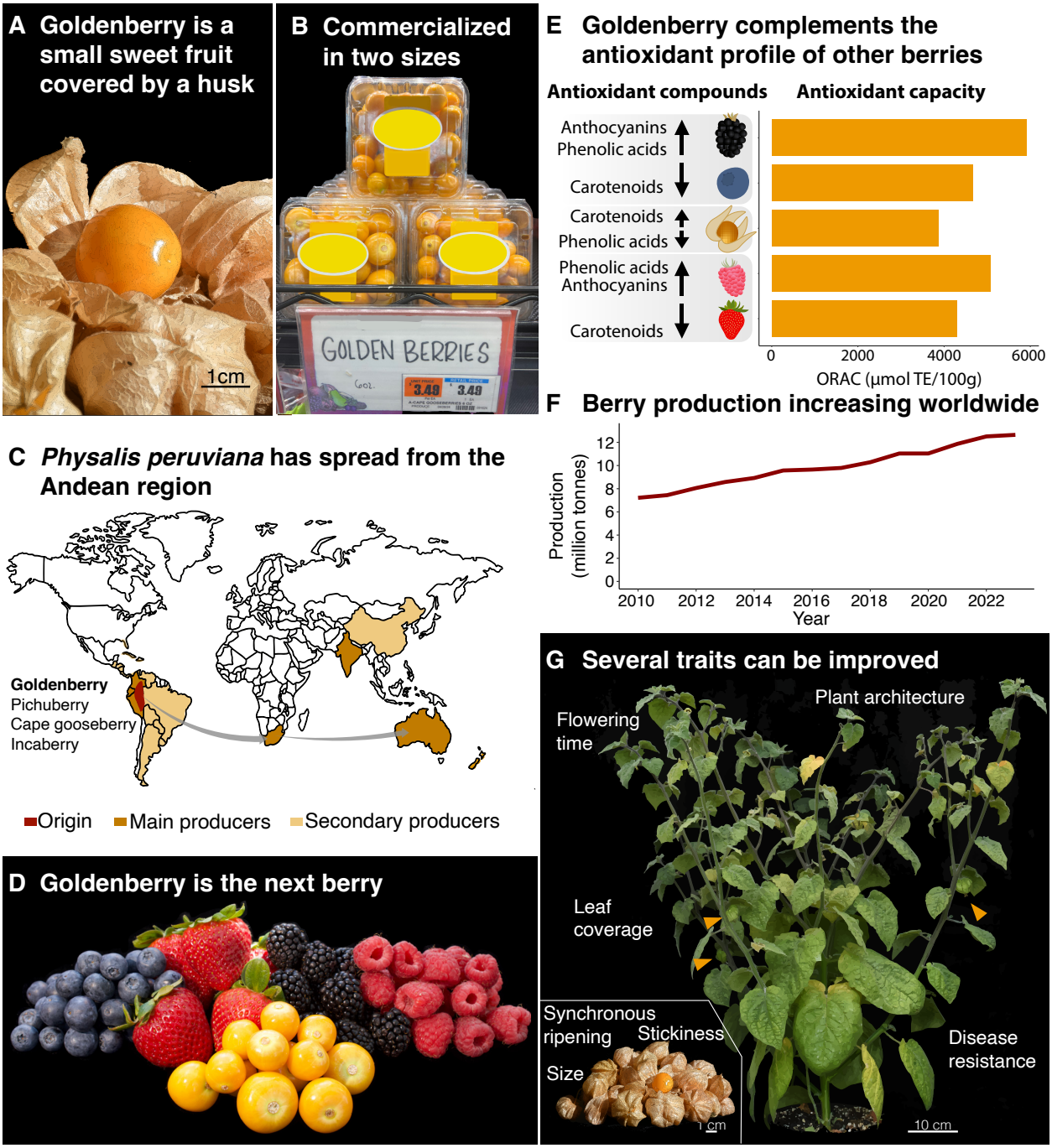
254 Shenstone, Esperanza, Zach Lippman, and Joyce Van Eck. 2020. "A Review of Nutritional
255 Properties and Health Benefits of Physalis Species." *Plant Foods for Human Nutrition* 75 (3):
256 316–25. https://doi.org/10.1007/s11130-020-00821-3.

257 Torii, K U, N Mitsukawa, T Oosumi, et al. 1996. "The Arabidopsis ERECTA Gene Encodes a
258 Putative Receptor Protein Kinase with Extracellular Leucine-Rich Repeats." *The Plant Cell* 8 (4):
259 735–46. https://doi.org/10.1105/tpc.8.4.735.

260 Van Eck, Joyce. 2022. "The Physalis Improvement Project: Blending Research with Community
261 Science." *EMBO Reports* 23 (1): e53918. https://doi.org/10.15252/embr.202153918.

262 Xia, Zhiqiang, Dongmei Huang, Shengkui Zhang, et al. 2021. "Chromosome-Scale Genome
263 Assembly Provides Insights into the Evolution and Flavor Synthesis of Passion Fruit (Passiflora
264 Edulis Sims)." *Horticulture Research* 8 (January): 14. https://doi.org/10.1038/s41438-020-00455-
265 1.

266 **FIGURES**



267

268 **Figure 1. Goldenberry (*Physalis peruviana*): an emerging commercial berry crop and a**
269 **promising candidate crop to help globalize using genome editing. A)** Goldenberry fruits
270 covered by a papery husk, typical of the inflated calyx syndrome (ICS) in the genus *Physalis* and
271 other Solanaceae species. A fruit with its husk removed is also shown. **B)** Commercial

goldenberries are assorted in two sizes: “small” and “large” types (Supplementary Figure S3). **C)** Historical origin, distribution, and producers of goldenberry. **D)** Commercial goldenberry is comparable and complementary to other berries in terms of fruit size and consumer interest. **E)** Oxygen Radical Absorbance Capacity (ORAC) measurements of antioxidants of the four major berries compared to goldenberry. Arrows represent the comparative contribution to antioxidant capacity of the different compounds in the different groups of berries. **F)** Growing worldwide production of berries (strawberries, raspberries, blueberries) from 2010 to 2023. **G)** Several plant and fruit traits which could be improved in goldenberry. Orange arrows point to goldenberry fruits.

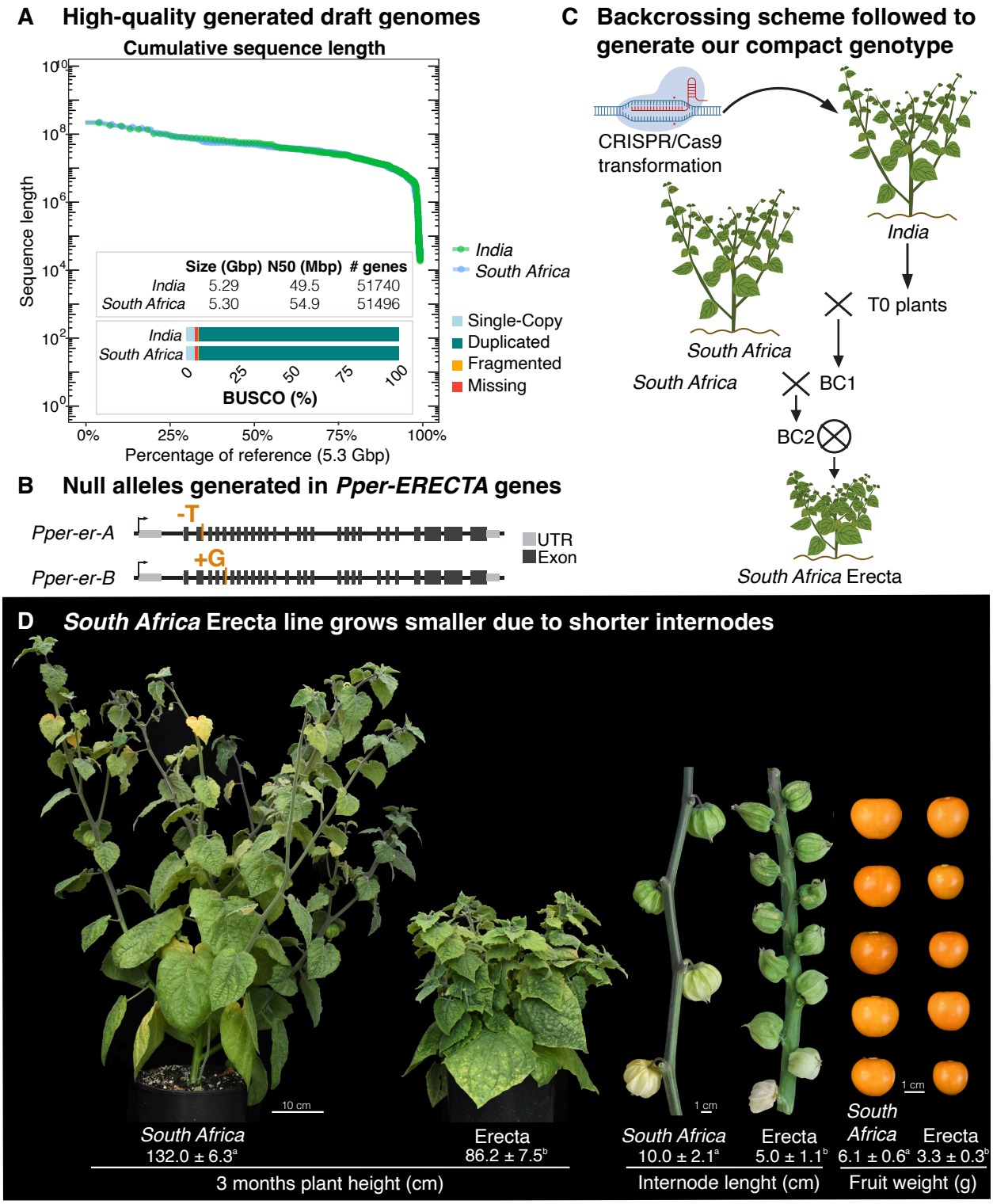


Figure 2. Establishment of reference genomes, genome editing, and a compact "Erecta" line to improve goldenberry cultivation. **A)** High-quality genome assemblies of the tetraploid parental lines *South African* and *India*, featuring maximum contig lengths approaching

285 chromosome arm lengths and high BUSCO completeness, indicating assembly integrity and
286 comprehensive gene content. **B)** Crossing scheme followed to introduce the CRISPR/Cas9-
287 engineered *erecta* mutant alleles into the commercial *South Africa* ecotype to develop two Erecta
288 varieties. **C)** Generated null mutant alleles in both *Pper-ERECTA* genes. **D)** Phenotypes of the
289 *South Africa* parental line compared to its corresponding Erecta line, showing a drastic reduction
290 in plant size and internode length. Fruits also become smaller, but are within the range of current
291 commercial types. Plant height: n = 20. Internode length: n = 50. Fruit weight: n = 50. Small letters
292 designate significantly different groups. Significant groups were defined as p-value < 0.05.

293

Figure 3. Demonstration of field-based row production of the new *South Africa* Erecta line (two middle rows) compared to the much larger and bushier *South Africa* (right row) and *India* (left row) parental lines. Note the shorter, prostrate growth habit of the *South Africa* Erecta line, facilitating high-density planting, plant maintenance, and harvestability.