



1 *Type of the Paper (Article)*

2 **Identification of candidate genes controlling red seed coat color 3 in cowpea (*Vigna unguiculata* [L.] Walp.)**

4 **Ira A. Herniter^{1,2,*}, María Muñoz-Amatriaín^{1,3}, Sassooum Lo^{1,4}, Yi-Ning Guo¹, Stefano Lonardi⁵, and Timothy J.
5 Close¹**

6 ¹ Department of Botany and Plant Sciences, University of California, Riverside, California, USA

7 ² Department of Plant Science, Rutgers University, New Brunswick, New Jersey, USA

8 ³ Departamento de Biología Molecular - Área de Genética, Universidad de León, León, Spain

9 ⁴ Department of Plant Science, University of California, Davis, California, USA

10 ⁵ Department of Computer Science and Engineering, University of California, Riverside, California, USA

11 * Correspondence: ira.herniter@rutgers.edu

12
13
14 **Abstract:** Seed coat color is an important consumer-related trait in cowpea (*Vigna unguiculata* [L.]
15 Walp.) and has been a subject of study for over a century. Utilizing newly available resources, in-
16 cluding mapping populations, a high-density genotyping platform, and several genome assem-
17 blies, red seed coat color has been mapped to two loci, *Red-1* (*R-1*) and *Red-2* (*R-2*), on Vu03 and
18 Vu07, respectively. A gene model (*Vigun03g118700*) encoding a dihydroflavonol 4-reductase, a
19 homolog of anthocyanidin reductase 1, which catalyzes the biosynthesis of epicatechin from
20 cyanidin, has been identified as a candidate gene for *R-1*. Possible causative variants have been also
21 identified for *Vigun03g118700*. A gene model on Vu07 (*Vigun07g118500*), with predicted nucleolar
22 function and high relative expression in the developing seed, has been identified as a candidate for
23 *R-2*. The observed red color is believed to be the result of a buildup of cyanidins in the seed coat.

24 **Keywords:** anthocyanidin reductase; cowpea; QTL analysis; seed coat color; SNP genotyping;
25 *Vigna unguiculata*

26

Citation: Lastname, F.; Lastname, F.;
Last-name, F. Title. *Horticulturae*

27 2021, 7, x.

<https://doi.org/10.3390/xxxxx>

28

Received: date

29

Accepted: date

30

Published: date

31

Publisher's Note: MDPI stays
neutral with regard to jurisdictional
claims in published maps and
institutional affiliations.

32

33

34
35 Copyright: © 2020 by the authors.
Submitted for possible open access
publication under the terms and
conditions of the Creative Commons
Attribution (CC BY) license
<http://creativecommons.org/licenses/by/4.0/>.

36
37
38

39

40

41

42

43

44

45

46

47

48

49

50

51

52

53

54

55

56

57

58

59

60

61

62

63

64

65

66

67

68

69

70

71

72

73

74

75

76

77

78

79

80

81

82

83

84

85

86

87

88

89

90

91

92

93

94

95

96

97

98

99

100

101

102

103

104

105

106

107

108

109

110

111

112

113

114

115

116

117

118

119

120

121

122

123

124

125

126

127

128

129

130

131

132

133

134

135

136

137

138

139

140

141

142

143

144

145

146

147

148

149

150

151

152

153

154

155

156

157

158

159

160

161

162

163

164

165

166

167

168

169

170

171

172

173

174

175

176

177

178

179

180

181

182

183

184

185

186

187

188

189

190

191

192

193

194

195

196

197

198

199

200

201

202

203

204

205

206

207

208

209

210

211

212

213

214

215

216

217

218

219

220

221

222

223

224

225

226

227

228

229

230

231

232

233

234

235

236

237

238

239

240

241

242

243

244

245

246

247

248

249

250

251

252

253

254

255

256

257

258

259

260

261

262

263

264

265

266

267

268

269

270

271

272

273

274

275

276

277

278

279

280

281

282

283

284

285

286

287

288

289

290

291

whole or ground into flour [4], [5]. Seed coat color is an important consumer-related trait in cowpea, as consumers usually make decisions about the quality and presumed taste of a product based on appearance [6], [7]. Cowpea grain displays a variety of colors, including black, brown, red, purple, and green. Each cowpea production region has preferred varieties, valuing certain color and pattern traits above others for determining quality and use. In West Africa, consumers pay a premium for seeds exhibiting certain characteristics specific to the locality, such as lack of color for use as flour or solid brown for use as whole beans [3], [8]–[10]. In the United States consumers prefer varieties with tight black eyes, commonly referred to as “black-eyed peas” [11].

Seed coat traits in cowpea have been studied since the early 20th century, when Spillman [12] and Harland [13], reviewed by Fery [14], explored the inheritance of factors controlling seed coat color and pattern. In a series of F₂ populations, Spillman [12] and Harland [13] identified genetic factors responsible for color expression, including “Red” (R), which was confirmed by Saunders [15] and Drabo et al [16]. The previous studies noted above, as well as similar studies in soybean (*Glycine max*) [17] and common bean (*Phaseolus vulgaris*) [18]–[20], have shown that there are two loci involved in red coat color, each following a simple Mendelian inheritance. One of those loci seems closely associated with the Color Factor locus, which controls overall pigmentation distribution on the seed coat [19]–[21]. This simple inheritance makes it relatively easy to develop reliable markers for use in marker-assisted breeding. Due to the high level of synteny between cowpea and other leguminous grain crops [22], markers developed in cowpea may be directly useful for breeding in those closely related species.

A genotyping array for 51,128 single nucleotide polymorphisms (SNP) been developed for cowpea [23] which, combined with relatively large populations, offers opportunities to improve the precision of genetic mapping. In particular, new populations have been developed for higher-resolution mapping including a minicore collection representing a worldwide cross-section of cultivated cowpea (“UCR Minicore”) [24] and an eight-parent Multi-parent Advanced Generation Inter-Cross (MAGIC) population [25]. In addition, a reference genome sequence of cowpea (phytozome-next.jgi.doe.gov) [22] and genome assemblies of six additional diverse accessions [26]; (<https://phytozome-next.jgi.doe.gov/cowpeapan/>) have been produced recently. Here, we make use of genetic and genomic resources available for cowpea to map the red seed coat trait, determine candidate genes, and propose sources of the observed variation.

2. Materials and Methods

73

2.1. Plant Materials

74

Three cowpea populations were used for mapping: the UCR Minicore, which includes 368 accessions [24], a MAGIC population consisting of 305 lines [25], and an F_2 population developed as part of this work from a cross between Sasaque (Red, *r-1*) and Sanzi (Not Red, *R-1*), consisting of 108 individuals. Example images of the seed coats, including these two accessions, can be found in **Figure 1**.

75

76

77

78

79

80

81

82

83

84

Candidate gene sequences were compared using the sequences of the cowpea reference genome (IT97K-499-35-1) and the six additional sequences from the pangenome (Sanzi, CB5, Suvita-2, UCR779, ZN016, and TZ30) [26].

These accessions used as parents in the mapping populations and as part of the cowpea pangenome [26] consist of at least one from each of the six identified subpopulations in global cowpea [24] (**Supplementary Table S1**).

85

2.2. SNP Genotyping and Data Curation

86

87

88

89

90

91

DNA was extracted from young leaf tissue using the Qiagen DNeasy Plant Mini Kit (Qiagen, Germany) per the manufacturer's instructions. The Cowpea iSelect Consortium Array (Illumina Inc., California, USA), which assays 51,128 SNPs [23] was used to genotype each DNA sample. Genotyping was performed at the University of Southern California Molecular Genomics Core facility (Los Angeles, California, USA). The same custom cluster file as in Muñoz-Amatriaín et al [23] was used for SNP calling.

92

93

94

95

96

97

For the UCR Minicore, SNPs with >20% missing data and a minor allele frequency (MAF) < 0.05 were eliminated, leaving a total of 42,603 SNPs used for mapping. For the MAGIC population, SNP data and a genetic map were available from Huynh et al [25], including 32,130 SNPs in 1,568 genetic bins. For the Sasaque by Sanzi F_2 population, SNPs were filtered to remove non-polymorphic loci between the parents, leaving 14,772 total polymorphic markers (**Supplementary Table S2**).

98

2.3. Seed Coat Phenotyping

99

00

01

02

03

04

05

Phenotype data for seed coat traits were collected by visual examination of the seeds. The scored phenotypic classes consisted of "Red" and "Not Red," as dictated by the presence or absence of observable red pigmentation in the mature seed coat. For mapping purposes, the "Red" lines were marked with a "1" and the "Not Red" with a "0." Accessions with black seed coat pigmentation or no pigmentation whatsoever were regarded as missing data as it is believed that black pigmentation obscures other colors from view [27] and that the lack of color is controlled by the *Color Factor* locus Herniter et

06 al [21]. Phenotype data can be found in **Supplementary Tables S3** (UCR Minicore), **S4**
07 (MAGIC), and **S5** (Sasaque by Sanzi).

08 Expected segregation ratios were determined based on the type of population and
09 the parental and, when available, F_1 phenotypes. Expected segregation ratios were tested
10 by chi-square analysis.

11 Dominance relationships were determined by examining the phenotypes of the F_1 of
12 the Sasaque by Sanzi population, the segregation of the parental phenotypes in the F_2
13 generation, and individuals from the early development of the MAGIC population. Seeds
14 from relevant individuals were visually examined for the presence of red pigmentation
15 on the seed coat.

16 2.4. *Genetic Mapping of the Red Seed Coat Trait*

17 Genetic mapping of the red seed coat trait was achieved with different methods for
18 each type of population. In the MAGIC population, the R package “mpMap” [28] was
19 used as described by Huynh et al [25]. The significance cutoff values were determined
20 through 1000 permutations, resulting in a threshold of $p = 8.10E-05$ [$-\log_{10}(p) = 4.09$]. Due
21 to the high number of markers in the genotype data, imputed markers spaced at 1 cM
22 intervals were used.

23 A Genome-Wide Association Study (GWAS) was performed in the UCR Minicore to
24 identify SNPs associated with the Red phenotype. The weighted mixed-linear model
25 (MLM) function [29] implemented in TASSEL v.5 (www.maizegenetics.net/tassel) was
26 used, with five principal components accounting for population structure in the dataset.
27 The $-\log_{10}(p)$ values were plotted against the physical coordinates of the SNPs [26]. A
28 Bonferroni correction was applied to correct for multiple testing error in GWAS, with the
29 significance cut-off set at α/n , where α is 0.05 and n is the number of tested markers, re-
30 sulting in a cut-off of $-\log_{10}(p) = 5.93$.

31 For the Sasaque by Sanzi F_2 population, the genotype calls of each bulked DNA pool
32 in the population were filtered to leave only the markers known to be polymorphic be-
33 tween the parents, and these were then sorted based on physical positions in the
34 pseudochromosomes of IT97K-499-35 [26]. The genotype data was then examined vis-
35 ually in Microsoft Excel for areas where the recessive bulk was homozygous and the
36 dominant bulk was heterozygous.

37 2.5. *Candidate Gene Identification and PCR Amplification*

38 The gene-annotated sequences of the overlapping QTL and region of interest for
39 both Vu03 and Vu07 were obtained from the v1.2 of the reference genome sequence of
40 cowpea (https://phytozome-next.jgi.doe.gov/info/Vunguiculata_v1_2) [22]. Candidate
41 genes were identified through location within the significant regions and similarity to
42 genes known to be involved in red pigmentation in soybean and common bean in the
43 literature (see Discussion). Additionally, data from the Gene Expression Atlas (VuGEA)
44 of the reference genome was utilized (<https://legumeinfo.org>) [30].

45 Primers were developed to amplify sections of the candidate gene *Vigun03g118700*
46 in Sasaque to examine the sequence for causative variations. Each primer pair, as well as
47 the annealing temperatures, are listed in **Supplementary Table S7**. PCR was performed
48 using the Thermo Scientific DreamTaq Green PCR Master Mix (Thermo Scientific, Mas-
49 sachusetts, USA) per the manufacturer's instructions. Primers were developed using
50 Primer3 v0.4.1 (bioinfo.ut.ee/primer3) and ordered from Integrated DNA Technology
51 (Coralville, Iowa, USA). PCR was run for 25-45 cycles with an annealing temperature
52 compatible with the primer pair and an extension time of 60-75 sec. Amplicons were
53 confirmed by gel electrophoresis. PCR products of the expected size were purified using
54 the QIAquick PCR Purification kit (Qiagen, Germany) and Sanger sequenced in both di-
55 rections on an Applied Biosystems 3730xl DNA Sequencer (University of California,
56 Riverside IIGB Genomics Core).

57 Additional sequences of *Vigun03g118700* and *Vigun07g118600* were obtained from
58 IT97K-499-35 (<https://phytozome.jgi.doe.gov>) and the other genomes sequenced as part
59 of the pangenome, consisting of Sanzi (Red), CB5-2 (Not Red), Suvita-2 (Not Red),
60 UCR779 (Not Red), TZ30 (Red), and ZN016 (Red)
61 (<https://phytozome-next.jgi.doe.gov/cowpeapan/>) [26]. Complete nucleotide and amino
62 acid sequences of both candidate genes were compared using MultAlin [31].

63 3. Results

64 3.1. Phenotypic Variation for Red Seed Coat Color

65 As described in Materials and Methods, phenotypic data were collected from the
66 three populations and their parents (when applicable) by visual examination of the seeds.
67 A summary of the phenotypic data, along with predicted segregation ratios, chi-square
68 values, and probability can be found in **Table 1**, with demonstrations in **Figure 2**.

69 Expected segregation ratios were determined based on the type of population and
70 the parental and, when available, F₁ phenotypes. For example, the Sasaque by Sanzi F₂

71 population was expected to segregate in a 3:1 ratio. For the MAGIC population, based on
72 how the population was constructed [25] it was assumed that each fully homozygous
73 parent had a roughly 1/8 chance to pass its genotype at a particular locus to a given RIL.
74 Three of the MAGIC parents had red seed coat (IT89KD-288, IT84S-2049, and
75 IT93K-503-1) while the other five parents did not, resulting in an expected 5:3 segregation
76 ratio.

77 The observed 3:1 segregation ratio of the Sasaque by Sanzi population together with
78 the observed Not Red seed coat of the F₁ indicate the dominance of *R*-1 (Not Red) over *r*-1
79 (Red). Similarly, a cross from the early development of the MAGIC population demon-
80 strated the dominance of *R*-2 (Not Red) over *r*-2 (Red): IT84S-2246 (*R*-2*R*-2, Not Red) was
81 crossed with IT93K-503-1 (*r*-2*r*-2, Red). The resulting plant, 2011-020-3 (*R*-2*r*-2) produced
82 seeds with the Not Red phenotype (Figure 2B).

83 3.2. Loci Controlling Red Seed Coat Color

84 Two QTL, one each on Vu03 and Vu07, were identified using the three populations.
85 The Vu03 QTL was identified in the UCR Minicore and the Sasaque x Sanzi F₂ popula-
86 tion, while the Vu07 QTL was identified in the UCR Minicore and the MAGIC population
87 (Table 2; Figure 3). Higher-resolution mapping for both genomic regions was achieved
88 using the UCR Minicore due to the higher genetic diversity and number of recombination
89 events available in this population. In particular, the significant region was 135 kbp in
90 size for Vu03, and it was reduced to just one SNP position for Vu07 (Table 2; Supple-
91 mentary Table S6). An analysis of linkage disequilibrium (*r*²) between pairs of SNPs lo-
92 cated 50 kbp upstream and downstream from the significant SNP for Vu07 (2_02375)
93 revealed low LD between the marker and nearby variants (analysis not shown). Com-
94 plete information on the mapping results can be found in Supplementary Tables S7
95 (UCR Minicore) and S8 (MAGIC).

96 3.3. Identification of Candidate Genes

97 The overlapping areas of the QTL and regions of interest on Vu03 and Vu07 were
98 examined for candidate genes in the cowpea reference genome v1.2
99 (phytozome-next.jgi.doe.gov/info/Vunguiculata_v1_2) [22]. Eleven genes were identified
00 in the minimal area (the area in which all the independent mappings identified) on Vu03
01 and one gene in the minimal area on Vu07 (Supplementary Table S9). Of the genes in the
02 minimal area on Vu03, *Vigun03g118700* is the best candidate as it encodes a
03 dihydroflavonol 4-reductase, which shows a highly elevated expression levels during

04 early seed development (**Supplementary Figure S2** [30]). The peak SNP on Vu03 is also
05 located within *Vigun03g118700*. The closest homologs to *Vigun03g118700* are
06 anthocyanidin reductases identified in various leguminous plants, including *Vigna*
07 *radiata*, *Vigna angularis*, common bean (*Phaseolus vulgaris*), and soybean (*Glycine max*),
08 among others (**Supplementary Table 10**).

09 In the UCR Minicore, GWAS only identified a single significant SNP on Vu07,
10 2_02375, contained within the larger region identified in mapping the MAGIC population
11 (**Table 2**, **Supplementary Tables S7 and S8**). This SNP is located in the third intron
12 of *Vigun07g118600*, which is orthologous to the *Arabidopsis BAH1/NLA* gene encoding
13 an E3 ubiquitin-protein ligase with roles in defense response, plant growth and develop-
14 opment [32]. The significant SNP is not in high LD with nearby variants. *Vigun07g118600*
15 does not show a high expression in developing seed tissue of the reference genome cv.
16 IT97K-499-35, which possesses black-eyed seeds (<https://legumeinfo.org>) [30].

17 To identify possible causative polymorphisms, sequence comparisons were made
18 using the published cowpea genomes [26] and the sequence of *Vigun03g118700* amplified
19 from the Sasaque genome. Variant locations in the following data are calculated from the
20 start of the coding sequence.

21 In *Vigun03g118700*, several SNPs were observed among accessions, one of which
22 could be causative (**Figure 4**, **Table 3**). In the two Red-seeded genotypes from the
23 pangenome, TZ30 and ZN016, there is a substitution at base 593, C→A, resulting in an
24 amino acid substitution, Ala→Asp. In Sasaque, there were two substitutions. One at base
25 661, T→G, resulting in an amino acid substitution, Met→Ser; the other at base 672, G→T,
26 resulting in an amino acid substitution, Val→Phe.

27 In addition, there were two substitutions in UCR779, a brown-seeded genotype from
28 the pangenome. At base 378, there is a silent mutation, C→T, still encoding Ile; and at
29 base 592, G→A, resulting in an amino acid substitution, Thr-Ala.

30 In *Vigun07g118600*, additional mutations were observed (**Figure 5**, **Table 3**). In both
31 TZ30 and ZN016 there are major differences from base 850 onwards, resulting in a dif-
32 ferent amino acid sequence. In addition, at base 761 in UCR779, there is a SNP, T→C,
33 resulting in an amino acid substitution.

35 *3.5. Figures, Tables and Schemes*

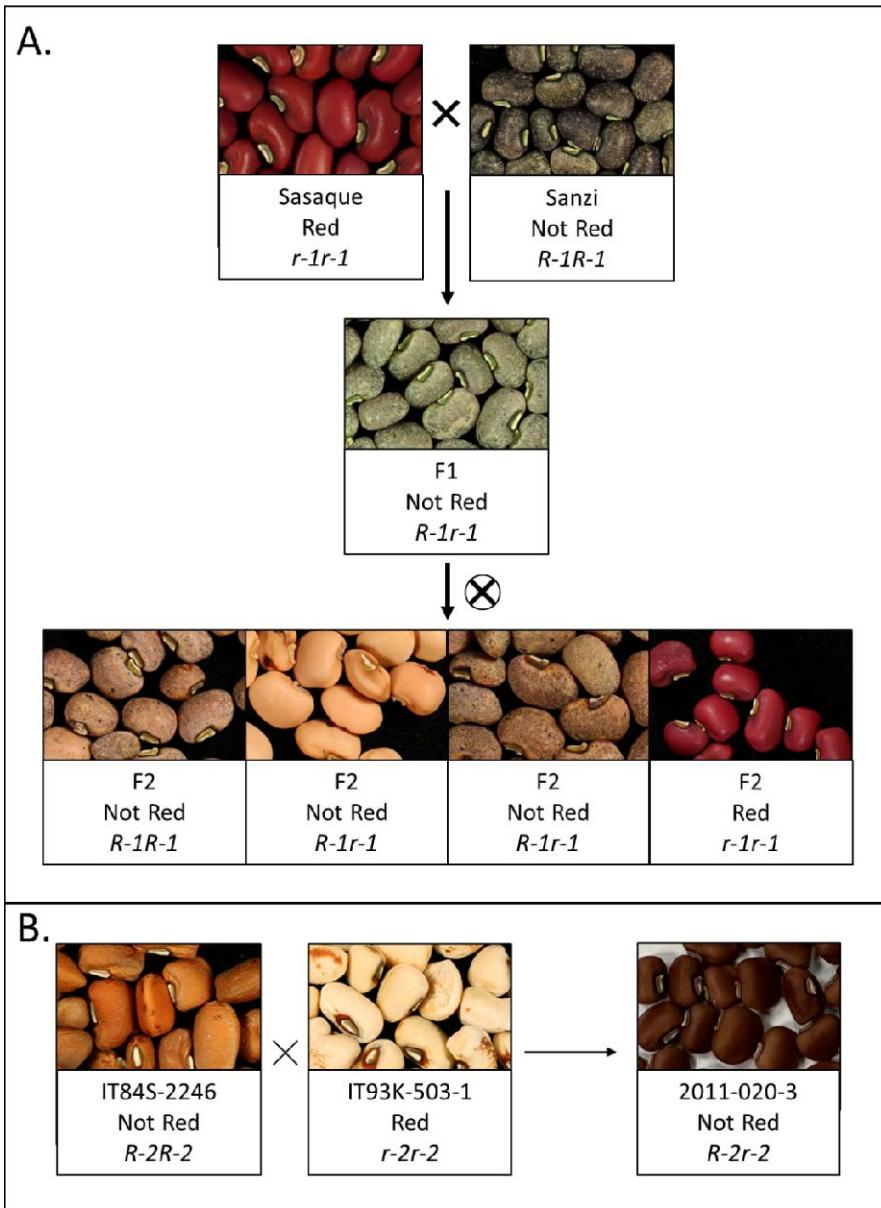


36

37

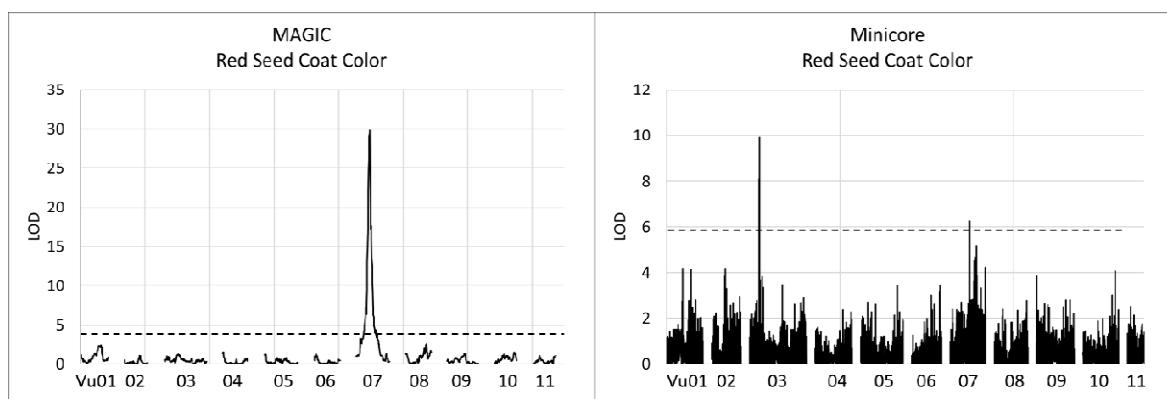
38

Figure 1. Images of accessions expressing and not expressing red seed coat color. Sasaque and IT93K-503-1 express red seed coat color; Sanzi and IT82E-18 do not.



39
40

41 **Figure 2.** Segregation at the two *Red* loci, *R-1* and *R-2*. (A) Segregation at the *R-1* locus in the
42 Sasaque by Sanzi F2 population. (B) Cross between IT84S-2246 and IT93K-503-1 from the early
43 development of the MAGIC population, resulting in Not Red phenotype in the seed coats on seeds
44 of the F1 maternal parent.



							130
IT97K-499-	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
CB5-2	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
Sanzi	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
Suvita2	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
UCR779	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
ZN016	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
TZ30	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
Consensus	MKFCKTYQRQ MQGHDHKLP VGFKKLKTM KRCCRRASTPT CPNHCOPDHCP VCDGTFPPSL LNEMSDIVGC FNQRAQKLL VHLASGFRKY FLMKGKMR NHTALIEEGK DLVIYALINA IAIRKILKKY						
	131						260
IT97K-499-	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
CB5-2	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
Sanzi	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
Suvita2	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
UCR779	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
ZN016	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
TZ30	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
Consensus	DKIHYSKQQQ LFKSKVQTMH KEILQSPWLS ELALHINLR ETKAKSREAS ALFHGCYLT KGDKPSLTCE LFDSFKIDIN LTCSCICLDTV FDPVSLACGH IFCYTCACLA ASVTIVDGK AAKFKAKCAL						
	261	318					
IT97K-499-	CRQQGVFEDA VHLEELNILL GRRYEVIHKT GYYLEVLETY IYCVILVLF MIYIFIIVD						
CB5-2	CRQQGVFEDA VHLEELNILL GRRYEVIHKT GYYLEVLETY IYCVILVLF MIYIFIIVD						
Sanzi	CRQQGVFEDA VHLEELNILL GRRYEVIHKT GYYLEVLETY IYCVILVLF MIYIFIIVD						
Suvita2	CRQQGVFEDA VHLEELNILL GRRYEVIHKT GYYLEVLETY IYCVILVLF MIYIFIIVD						
UCR779	CRQQGVFEDA VHLEELNILL GRRYEVIHKT GYYLEVLETY IYCVILVLF MIYIFIIVD						
ZN016	CRQQGVFEDA VHLEELNILL GRRKEYWEQ RLRLMERIERV KQVKEHWQTQ CREFVGI						
TZ30	CRQQGVFEDA VHLEELNILL GRRKEYWEQ RLRLMERIERV KQVKEHWQTQ CREFVGI						
Consensus	CRQQGVFEDA VHLEELNILL GRRkeyweq rlrl\$erierv kqvkehwtq crefvg!						

Figure 5. Alignment of proposed amino acid sequences of *Vigun07g118600* among genotypes in the pangenome. ZN016 and TZ30 express the red seed coat phenotype. The other genotypes do not have red seed coat coloring.

53

54
55
56

57

58

Population (# of lines)	Not Red	Red	Other	Pred. Seg. Ratio	X ²	Probability
UCR Minicore (368)	190	80	98	--	--	--
MAGIC (305)	157	92	56	5:3	0.032	0.86
Sasaque by Sanzi (108)	86	22	0	3:1	1.23	0.27

59

60
61

62

Table 1. Seed coat phenotypes and segregation ratios of the tested populations. "Other" includes seeds with black pigmentation or no pigmentation.

Population	Chr	Peak SNP	Flanking SNPs	QTL pos. (bp)	LOD/-log10 (p)	R ² (%)	Effect
Sasaque x Sanzi (F ₂)	Vu03	--	2_38494 - 2_05272	7,657,076 - 41,457,212	--	--	
UCR Minicore	Vu03	2_20787	2_03389 - 2_20787	10,933,603 - 11,068,119	9.82	16.6	-0.56 (G)
UCR Minicore	Vu07	2_02375	--	22,014,760	6.26	9.8	-0.61 (A)
MAGIC	Vu07	--	2_51274 - 2_04829	14,976,909 - 22,928,810	29.81	38.82	

63

Table 2. Mapping results in the different populations.

64

Vigun03g118700						
<u>BP</u>	<u>RefGen Base</u>	<u>Mutant Base</u>	<u>Genotypes Present In</u>	<u>AA Code</u>	<u>AA</u>	<u>Mutation Type</u>
378	C	T	UCR779	AT <u>T</u> →AT <u>C</u>	Ile→Ile	Silent
592	G	A	UCR779	<u>A</u> CT→G <u>C</u> T	Thr→Ala	Polar Uncharged → hydrophobic
593	C	A	ZN016, TZ30	G <u>C</u> T→G <u>A</u> T	Ala→Asp	hydrophobic → polar negative
661	T	G	Sas aque	A <u>T</u> G→A <u>G</u> C	Met→Ser	hydrophobic → polar uncharged
672	G	T	Sas aque	G <u>T</u> T→T <u>TT</u>	Val→Phe	hydrophobic → hydrophobic
Vigun07g118600						
<u>BP</u>	<u>RefGen Base</u>	<u>Mutant Base</u>	<u>Genotypes Present In</u>	<u>AA Code</u>	<u>AA</u>	<u>Mutation Type</u>
761	T	C	UCR779	T <u>T</u> C→T <u>C</u> C	Phe→Ser	hydrophobic → polar uncharged
849+			ZN016, TZ30			Completely different AA sequence

65 **Table 3.** Observed variations in the coding DNA sequence of the candidate genes. The base number
66 is calculated from the start of the coding sequence. “849+” indicates that including base 849 to the
67 end of the coding sequence.

68 **4. Discussion**

69 Previous studies, dating back over a hundred years, by Harland [13] and Spillman
70 [12], reviewed by Fery [14] and confirmed by Drabo et al [16], identified only the single
71 locus *R* as being responsible for the red seed coat phenotype in cowpea. The present data
72 suggests the presence of two loci each of which can be independently responsible for the
73 expression of red seed coat color. To avoid confusion, the locus on Vu03 is referred to as
74 *R-1* while the locus on Vu07 is referred to as *R-2*. While the previous studies in cowpea
75 and soybean have only identified a single locus responsible for red seed coat color, Bassett
76 et al [33] identified two loci in common bean, termed *Rk* (red kidney) and *R* (oxblood
77 red). The candidate gene on Vu07, *Vigun07g118600*, is closely linked the Color Factor
78 locus, believed to be controlled by *Vigun07g110700* [21], just 1.5 Mb [26] and 0.96 cM
79 distant according to the consensus map established by Muñoz-Amatriaín et al [23]. This
80 occurrence could explain why a second locus for Red color in cowpea was not previously
81 identified, as on seeds with only small areas of pigmentation, such as IT93K-503-1 (Fig-
82 ure 1) it can be difficult to discern Red (*r-1/r-1, r-2/r-2*) from Not Red (*R-1, R-2*). A similar
83 linkage has been well-documented in common bean (*Phaseolus vulgaris*) [33], and has
84 been recently confirmed through trait mapping [19], [20].

85 The flavonoid biosynthesis pathway has been well characterized in a number of
86 plants, including soybean [17] and *Arabidopsis* [34], among others. An intermediate

87 group of molecules in the flavonoid biosynthesis pathway are cyanidins. Cyanidins are
88 well-described pigmentation molecules, identified in a wide range of plants, including
89 red cabbage [35], clover [36], and various red berries [37], among others. Cyanidins ap-
90 pear red in low pH (<3.5) conditions [38]. It has been well-established much of the ob-
91 served variation in plant organ pigmentation arises from differential levels of pigmenta-
92 tion molecules and pH in those organs [39].

93 The present study reports the identification of the *R*-1 locus on Vu03. The most likely
94 candidate gene, *Vigun03g118700*, encodes a dihydroflavonol 4-reductase protein (D4R).
95 Based on the dominance pattern, with Not Red (*R*-1) dominant towards Red (*r*-1), it is
96 most likely that this is the previously identified *R* locus from the literature [14], [16]. A
97 BLAST search of the soybean (*Glycine max*) genome on Phytozome, showed that the
98 closest homolog of *Vigun03g118700* in soybean *Glyma.08G062000* (previously referred to
99 as *Glyma08g06630*), with 90.63% identity and a e-value of 0.0 from pairwise BLAST of the
00 coding sequence.

01 Kovinich et al [17] demonstrated that soybean mutants with reduced function of the
02 soybean *ANR1* gene (*Glyma.08G062000*) had a red seed coat phenotype. This would be
03 consistent with a complete dominance model as a single functional copy of the gene
04 would allow successful catalysis of cyanidin to epicatechin and result in no red pigmen-
05 tation, while an individual with two recessive alleles and nonfunctional D4R would not
06 catalyze the biosynthesis of epicatechin, resulting in a buildup of cyanidin, and instead
07 would express red seed coat pigmentation. The observed allele sequence in common
08 bean would match this, with the dominant *Rk* allele not expressing red pigmentation, the
09 recessive *rk^d* allele expressing red pigmentation, and the semidominant *rk* allele in the
10 middle, with some red pigmentation [18].

11 Similarly, the closest homolog of *Vigun03g118700* in common bean is
12 *Phvul.002G218700*, with 95.66% identity and an e-value of 0.0 from pairwise BLAST of the
13 coding sequences. This gene is also predicted to encode an ANR. The recent mapping of
14 color in common bean by Campa et al [19] and Sadohara et al [20] identified SNPs which
15 were significantly associated with the *a** trait, which measures greenness-redness, as de-
16 termined by the International Commission on Illumination [40]. Both found significantly
17 correlated markers located 6-15 Mb up and downstream of *Phvul.002G218700*. It is possible
18 that the mapping did not identify markers closer because they did not map Red versus
19 Not Red phenotypes directly, incorporating the green color as part of the phenotyping.

20 While the sequence comparisons of *Vigun03g118700* identified several potentially
21 causative variations in Sasaque, TZ30, and ZN016, they also identified variations present
22 in UCR779, which does not have any red pigmentation (Table 3). While the base substi-
23 tution at base 378 (C→T, Ile→Ile) in *Vigun03g118700* is a silent mutation, the substitutions
24 at base 592 (G→A, Thr→Ala) in *Vigun03g118700* and at base 761 (T→C, Phe→Ser) in
25 *Vigun07g118600* are more substantial. It is unclear why these mutations do not result in a
26 similarly expected loss of the function to the mutations found in the Red phenotyped
27 varieties. Perhaps some insight could be gained from the determination that UCR779 is
28 part of the most highly divergent subpopulation of cowpea [24], from eastern and
29 southern Africa. It may be that its sequence represents alternative Not Red alleles (*r-1* &
30 *R-2*). To better understand this, future studies should examine other diverse cowpea
31 genotypes.

32 The *R-2* locus identified in the present study mapped to Vu07. Like the dominant
33 red *R* locus in common bean, Not Red color was dominant at this locus (Figure 2B) and it
34 is closely linked to the Color Factor (*C*) locus, previously mapped by Herniter et al [21].
35 The lone significant SNP identified here falls within an intron of *Vigun07g118600*, which
36 is predicted to encode an E3 ubiquitin ligase. The available expression data for this gene
37 shows quite low expression, with only about 0.5 TPM in the early developing seed and
38 dropping off thereafter (Supplementary Figure S3). However, it should be noted that the
39 only genotype for which transcription data is available is the cowpea reference genome,
40 IT97K-499-35, which does not have red pigmentation and likely has the recessive, non-
41 functional (*r-2*) allele, which is likely not expressed. This expression differential might be
42 explained by variation in the promoter region. Future studies should seek to identify the
43 transcription factor which controls expression of *Vigun07g118600* and identify potential
44 binding sites which could vary between genotypes.

45 While the function of *Vigun03g118700* can be inferred based on previous research in
46 other species, the function of *Vigun07g118600* in seed coat color is much less clear. Some
47 possibilities include regulation of the flavonoid biosynthesis pathway, perhaps even ex-
48 erting negative control over D4R (*R-1* locus, *Vigun03g118700*). Similarly, the closest
49 homolog in soybean is *Glyma.10G018800*, with 90.26% identity and an e-value of 0.0 from
50 pairwise BLAST of the coding sequence. Interestingly, *Glyma.10G018800* shows elevated
51 expression in the root
(https://phytozome-next.jgi.doe.gov/report/transcript/Gmax_Wm82_a4_v1/Glyma.10G018800)

53 18800.1) and has been identified as a candidate gene controlling root growth [41], [42],
54 potentially resulting from divergent evolution.

55 The closest homolog of *Vigun07g118600* in common bean is *Phvul.007G143600*. In
56 their mappings, both Sadohara et al [20] and Campa et al [19] identified a SNP 6 Mb
57 downstream of *Phvul.007G143600*, as significantly associated with the *a** trait. However,
58 Sadohara et al [20] did not proffer candidate genes and Campa et al [19] did not consider
59 this gene. It may be that, similarly to the function of the homolog in soybean, the function
60 of *Phvul.007G143600* has diverged over evolutionary time.

61 62 5. Conclusions

63 This study presents the mapping of red seed coat color in cowpea, identifying two
64 loci, one each on Vu03 and Vu07, and proposing a candidate gene for each,
65 *Vigun03g118700*, encoding a dihydroflavonol 4-reductase, and *Vigun07g118600*, encoding
66 an E3 ubiquitin ligase. Sequence comparisons between accessions either expressing or
67 not expressing red pigmentation identified multiple potentially causative variations. The
68 red seed coat color is likely caused by the disruption of the flavonoid biosynthesis
69 pathway, resulting in a buildup of anthocyanidin in the seed coat.

70 **Supplementary Materials:** The following are available online at www.mdpi.com/xxx/s1, Figure S1:
71 *Vigun03g118700* expression profile, Figure S2: *Vigun07g118600* expression profile, Figure S3: Coding
72 sequence alignment of *Vigun03g118700* among the pangenome sequences, Figure S4: Coding
73 sequence alignment of *Vigun07g118600* among the pangenome sequences, Table S1: Country of
74 origin and subpopulation affinity of the population parents and pangenome accessions, Table S2:
75 Genotypes of the bulked samples from the F2 population Sasaque by Sanzi, Table S3: Phenotypes
76 of the UCR Minicore Collection, Table S4: Phenotypes of the UCR MAGIC population, Table S5:
77 Phenotypes of the F2 Sasaque by Sanzi population, Table S6: Primers used to amplify the candidate
78 gene *Vigun03g118700* for sequencing, Table S7: GWAS results from the UCR Minicore Collection,
79 Table S8: Mapping results from the UCR MAGIC population, Table S9: Candidate genes identified
80 within the QTL peaks, Table S10: BLASTP Results of the *Vigun03g118700* transcript.

81 **Author Contributions:** Conceptualization, I.A.H., T.J.C.; data curation, I.A.H., St.L.; formal analy-
82 sis, I.A.H., M.M.A.; funding acquisition, St.L., T.J.C.; investigation, I.A.H., M.M.A., Sa.L., and
83 Y.N.G.; methodology, I.A.H.; project administration, I.A.H., M.M.A., and T.J.C.; resources, I.A.H.,
84 T.J.C.; software, St.L.; supervision, T.J.C.; validation, I.A.H., M.M.A.; visualization, I.A.H., M.M.A.;
85 writing – original draft, I.A.H.; writing – review and editing, I.A.H., St.L., M.M.A., Sa.L., Y.N.G.,
86 and T.J.C. All authors have read and agreed to the published version of the manuscript.

87 **Funding:** This study was supported by the Feed the Future Innovation Lab for Climate Resilient
88 Cowpea (USAID Cooperative Agreement AID-OAA-A-13-00070), the National Science Foundation
89 BREAD project “Advancing the Cowpea Genome for Food Security” (NSF IOS-1543963) and Hatch
90 Project CA-R-BPS-5306-H.

91 **Informed Consent Statement:** Not applicable.

92 **Data Availability Statement:** The original contributions presented in the study are included in the
93 article/supplementary material; further inquiries can be directed to the corresponding author.

94
95
96
97
98
99

Acknowledgments: The authors thank Eric Castillo, Mia Rochford, Julia Valdepeña, and Sabrina Phengsy for assistance with seed photography; Steve Wanamaker for assistance in the analysis of the various genome sequences; Bao-Lam Huynh for providing the MAGIC population and attendant genotypic and pedigree information; Pei Xu and Xinyi Yu for providing images of the TZ30 and ZN016 genotypes.

Conflicts of Interest: The authors declare no conflict of interest.

00 References

01 [1] FAOSTAT, "Crops," 2023. <http://www.fao.org/faostat/en/#data/QC> (accessed Feb. 26, 2020).

02 [2] J. D. Ehlers and A. E. Hall, "Cowpea (*Vigna unguiculata* L. Walp.)," *F. Crop. Res.*, vol. 53, no. 1–3, pp. 187–204, 1997, doi: 10.1016/S0378-4290(97)00031-2.

03 [3] O. Boukar *et al.*, "Cowpea (*Vigna unguiculata*): Genetics, genomics and breeding," *Plant Breed.*, vol. 138, no. 4, pp. 415–424, Aug. 2019, doi: 10.1111/pbr.12589.

04 [4] B. B. Singh, *Cowpea: The Food Legume of the 21st Century*. 2014. Accessed: May 11, 2016. [Online]. Available: <https://dl.sciencesocieties.org/publications/books/tocs/acsesspublicati/cowpeathefoodle>

05 [5] A. R. Tijjani, R. T. Nabinta, and M. Muntaka, "Adoption of innovative cowpea production practices in a rural area of Katsina State, Nigeria," *J. Agric. Crop Res.*, vol. 3, no. June, pp. 53–58, 2015.

06 [6] S. R. Jaeger, L. Antúnez, G. Ares, M. Swaney-Stueve, D. Jin, and F. R. Harker, "Quality perceptions regarding external appearance of apples: Insights from experts and consumers in four countries," *Postharvest Biol. Technol.*, vol. 146, pp. 99–107, Dec. 2018, doi: 10.1016/J.POSTHARVBIO.2018.08.014.

07 [7] A. S. Kostyla, F. M. Clydesdale, and M. R. McDaniel, "The psychophysical relationships between color and flavor," *Food Sci. Nutr.*, vol. 10, no. 3, pp. 303–321, 1978, doi: 10.1080/10408397809527253.

08 [8] A. S. Langyintuo *et al.*, "Cowpea supply and demand in West and Central Africa," *F. Crop. Res.*, vol. 82, no. 2, pp. 215–231, 2003, doi: 10.1016/S0378-4290(03)00039-X.

09 [9] F. J. Mishili *et al.*, "Consumer preferences for quality characteristics along the cowpea value chain in Nigeria, Ghana, and Mali," *Agribusiness*, vol. 25, no. 1, pp. 16–35, Sep. 2009, doi: 10.1002/agr.20184.

10 [10] A. H. Ira, J. Zhenyu, and K. Francis, "Market preferences for cowpea (*Vigna unguiculata* [L.] Walp) dry grain in Ghana," *African J. Agric. Res.*, vol. 14, no. 22, pp. 928–934, May 2019, doi: 10.5897/AJAR2019.13997.

11 [11] R. L. Fery, "The Genetics of Cowpeas: A Review of the World Literature," in *Cowpea Research, Production and Utilization*, S. R. Singh and K. O. Rachie, Eds. New York: John Wiley & Sons, Inc., 1985, pp. 25–62.

12 [12] W. J. Spillman, "Inheritance of the 'eye' in *Vigna*," *Am. Nat.*, vol. XLV, no. 53, pp. 513–523, 1911.

13 [13] S. C. Harland, "Inheritance of certain characters in the cowpea (*Vigna sinensis*)," *J. Genet.*, vol. 8, no. 2, pp. 101–132, Apr. 1919, doi: 10.1007/BF02983490.

14 [14] R. L. Fery, "Genetics of *Vigna*," in *Horticultural Reviews*, vol. 2, Hoboken, NJ, USA: John Wiley & Sons, Inc., 1980, pp. 311–394. doi: 10.1002/9781118060759.ch7.

15 [15] A. R. Saunders, "Inheritance in the cowpea (*Vigna sinensis* Endb.). II: seed coat colour pattern; flower, plant, and pod color," *South African J. Agric. Sci.*, vol. 3, no. 2, pp. 141–162, 1960.

16 [16] I. Drabo, T. A. O. Ladeinde, J. B. Smithson, and R. Redden, "Inheritance of Eye Pattern and Seed Coat Colour in Cowpea (*Vigna unguiculata* [L.] Walp)," *Plant Breed.*, vol. 100, no. 2, pp. 119–123, Mar. 1988, doi: 10.1111/j.1439-0523.1988.tb00226.x.

17 [17] N. Kovinich, A. Saleem, J. T. Arnason, and B. Miki, "Identification of two anthocyanidin reductase genes and three red-brown soybean accessions with reduced anthocyanidin reductase 1 mRNA, activity, and seed coat proanthocyanidin amounts," *J. Agric. Food Chem.*, vol. 60, no. 2, pp. 574–584, Jan. 2012, doi: 10.1021/jf2033939.

18 [18] M. J. Bassett, "Genetics of Seed Coat Color and Pattern in Common Bean," in *Plant Breeding Reviews*, 28th ed., Hoboken, NJ, USA: Wiley, 2007, pp. 239–315. doi: 10.1002/9780470168028.ch8.

37 [19] A. Campa, R. Rodríguez Madrera, M. Jurado, C. García-Fernández, B. Suárez Valles, and J. J. Ferreira, "Genome-wide
38 association study for the extractable phenolic profile and coat color of common bean seeds (*Phaseolus vulgaris* L.)," *BMC
39 Plant Biol.*, vol. 23, no. 1, pp. 1–16, Dec. 2023, doi: 10.1186/S12870-023-04177-Z/TABLES/5.

40 [20] R. Sadohara, Y. Long, P. Izquierdo, C. A. Urrea, D. Morris, and K. Cichy, "Seed coat color genetics and genotype ×
41 environment effects in yellow beans via machine-learning and genome-wide association," *Plant Genome*, vol. 15, no. 1, p.
42 e20173, Mar. 2021, doi: 10.1002/tpg2.20173.

43 [21] I. A. Herniter *et al.*, "Seed Coat Pattern QTL and Development in Cowpea (*Vigna unguiculata* [L.] Walp.)," *Front. Plant Sci.*,
44 vol. 10, p. 514455, Oct. 2019, doi: 10.3389/fpls.2019.01346.

45 [22] S. Lonardi *et al.*, "The genome of cowpea (*Vigna unguiculata* [L.] Walp.)," *Plant J.*, vol. 98, no. 5, pp. 767–782, Jun. 2019, doi:
46 10.1111/tpj.14349.

47 [23] M. Muñoz-Amatriaín *et al.*, "Genome resources for climate-resilient cowpea, an essential crop for food security," *Plant J.*,
48 vol. 89, no. 5, pp. 1042–1054, Mar. 2017, doi: 10.1111/tpj.13404.

49 [24] M. Muñoz-Amatriaín *et al.*, "The UCR Minicore: a valuable resource for cowpea research and breeding," *Legum. Sci.*, p.
50 leg3.95, May 2021, doi: 10.1002/leg3.95.

51 [25] B.-L. Huynh *et al.*, "A multi-parent advanced generation inter-cross (MAGIC) population for genetic analysis and
52 improvement of cowpea (*Vigna unguiculata* L. Walp.)," *Plant J.*, vol. 93, no. 6, pp. 1129–1142, Mar. 2018, doi:
53 10.1111/tpj.13827.

54 [26] Q. Liang *et al.*, "A view of the pan-genome of domesticated Cowpea (*Vigna unguiculata* [L.] Walp.)," *Plant Genome*, p.
55 e20319, Mar. 2023, doi: 10.1002/TPG2.20319.

56 [27] I. A. Herniter, M. Muñoz-Amatriaín, S. Lo, Y.-N. Guo, and T. J. Close, "Identification of Candidate Genes Controlling Black
57 Seed Coat and Pod Tip Color in Cowpea (*Vigna unguiculata* [L.] Walp.)," *G3; Genes|Genomes|Genetics*, vol. 8, no. 10, pp.
58 3347–3355, Oct. 2018, doi: 10.1534/g3.118.200521.

59 [28] B. E. Huang and A. W. George, "R/mpMap: a computational platform for the genetic analysis of multiparent recombinant
60 inbred lines," *Bioinformatics*, vol. 27, no. 5, pp. 727–729, Mar. 2011, doi: 10.1093/bioinformatics/btq719.

61 [29] Z. Zhang *et al.*, "Mixed linear model approach adapted for genome-wide association studies," *Nat. Genet.*, vol. 42, no. 4, pp.
62 355–360, Apr. 2010, doi: 10.1038/ng.546.

63 [30] S. Yao *et al.*, "The *Vigna unguiculata* Gene Expression Atlas (VuGEA) from de novo assembly and quantification of
64 RNA-seq data provides insights into seed maturation mechanisms," *Plant J.*, vol. 88, no. 2, pp. 318–327, Oct. 2016, doi:
65 10.1111/tpj.13279.

66 [31] F. Corpet, "Multiple sequence alignment with hierarchical clustering," *Nucleic Acids Res.*, vol. 16, no. 22, pp. 10881–10890,
67 Nov. 1988, doi: 10.1093/NAR/16.22.10881.

68 [32] T. Yaeno and K. Iba, "BAH1/NLA, a RING-type ubiquitin E3 ligase, regulates the accumulation of salicylic acid and
69 immune responses to *Pseudomonas syringae* DC3000," *Plant Physiol.*, vol. 148, no. 2, pp. 1032–1041, 2008, doi:
70 10.1104/pp.108.124529.

71 [33] M. J. Bassett, R. Lee, C. Otto, and P. E. McClean, "Classical and Molecular Genetic Studies of the Strong Greenish Yellow
72 Seedcoat Color in 'Wagenaar' and 'Enola' Common Bean," *J. Am. Soc. Hortic. Sci.*, vol. 127, no. 1, pp. 50–55, 2002.

73 [34] K. Petroni and C. Tonelli, "Recent advances on the regulation of anthocyanin synthesis in reproductive organs," *Plant Sci.*,
74 vol. 181, no. 3, pp. 219–229, 2011, doi: 10.1016/j.plantsci.2011.05.009.

75 [35] N. Ahmadiani, R. J. Robbins, T. M. Collins, and M. M. Giusti, "Molar absorptivity (ϵ) and spectral characteristics of
76 cyanidin-based anthocyanins from red cabbage," *Food Chem.*, vol. 197, pp. 900–906, Apr. 2016, doi:
77 10.1016/j.foodchem.2015.11.032.

78 [36] H. Zhang, H. Tian, M. Chen, J. Xiong, H. Cai, and Y. Liu, "Transcriptome analysis reveals potential genes involved in

79 flower pigmentation in a red-flowered mutant of white clover (*Trifolium repens* L.),” *Genomics*, vol. 110, no. 3, pp. 191–200,
80 May 2018, doi: 10.1016/j.ygeno.2017.09.011.

81 [37] M. J. Jara-Palacios, A. Santisteban, B. Gordillo, D. Hernanz, F. J. Heredia, and M. L. Escudero-Gilete, “Comparative study of
82 red berry pomaces (blueberry, red raspberry, red currant and blackberry) as source of antioxidants and pigments,” *Eur.
83 Food Res. Technol.*, vol. 245, no. 1, pp. 1–9, Jan. 2019, doi: 10.1007/s00217-018-3135-z.

84 [38] S. Asen, K. H. Norris, and R. N. Stewart, “Absorption spectra and color of aluminium-cyanidin 3-glucoside complexes as
85 influenced by pH,” *Phytochemistry*, vol. 8, no. 3, pp. 653–659, Mar. 1969, doi: 10.1016/S0031-9422(00)85416-3.

86 [39] J. Mol, E. Grofewold, and R. Koes, “How genes paint flowers and seeds,” *Trends in Plant Science*, vol. 3, no. 6, pp. 212–217,
87 Jun. 01, 1998. doi: 10.1016/S1360-1385(98)01242-4.

88 [40] CIE, “Commission Internationale de l’Eclairage CIE 015: 2018 Colorimetry,” 2018. doi: 10.25039/TR.015.2018.

89 [41] S. Sarkar, A. Shekoofa, A. McClure, and J. D. Gillman, “Phenotyping and Quantitative Trait Locus Analysis for the Limited
90 Transpiration Trait in an Upper-Mid South Soybean Recombinant Inbred Line Population (‘Jackson’ × ‘KS4895’): High
91 Throughput Aquaporin Inhibitor Screening,” *Front. Plant Sci.*, vol. 12, p. 779834, Jan. 2022, doi:
92 10.3389/FPLS.2021.779834/BIBTEX.

93 [42] X. Li, X. Zhang, Q. Zhao, and H. Liao, “Genetic improvement of legume roots for adaption to acid soils,” *Crop J.*, vol. 11, no.
94 4, pp. 1022–1033, Aug. 2023, doi: 10.1016/J.CJ.2023.04.002.

95