**SLAF-based construction of a high-density genetic map and its**

**application in QTL mapping of carotenoids content in citrus fruit**

Xiongjie Zheng1, Yuqing Tang1, Junli Ye1, Zhiyong Pan1, Meilian Tan1a, Zongzhou

Xie1, Lijun Chai1, Qiang Xu1, Paul D. Fraser 2, Xiuxin Deng1\*

1Key Laboratory of Horticultural Plant Biology (Ministry of Education), Huazhong

Agricultural University, Wuhan, China. 2School of Biological Sciences, Royal

Holloway, University of London, Egham, Surrey, TW20 0EX, UK

\* Corresponding author:

Xiuxin Deng

Key Laboratory of Horticultural Plant Biology (Ministry of Education), Huazhong

Agricultural University, Wuhan, China

Email: xxdeng@mail.hzau.edu.cn

Tel number: +86-27-87286906, Fax number: +86-27-87281850

a M.T. current address is Key Laboratory of Biology and Genetic Improvement of Oil

Crops, Ministry of Agriculture, Oil Crops Research Institute of the Chinese Academy

of Agricultural Sciences, Wuhan, China.

Author Contributions:

X.D. conceived the project and supervised the experiments; X.D. and X.Z. designed theexperiments; X.Z. and Y.T. performed most of the experiments and analyzed the data; M.T. and Z.X. provide vital materials; X.Z. wrote the article with contributions of all the authors; Z.P., J.Y., L.C., Q.X. and P.D.F provided critical comments for

experiments and writing improvement.

**ABSTRACT**

Carotenoids are important antioxidant components in the human diet. To develop

carotenoid-rich agricultural products by genetic intervention, understanding the genetic basis of carotenoids variation is essential. In this study, we constructed a high-density integrated genetic map with 3,817 molecular markers using specific-locus amplified fragment (SLAF) sequencing from a *C. reticulata* x *P. trifoliata* F1 pseudo-testcross population. A total of 17 significant quantitative trait loci (QTLs) distributed on Chromosomes (Chr) 2, 3, 5, 6 and 9 were detected to determine the carotenoid variation in the population. In particular, three QTL co-localizations for multiple carotenoid constituents were observed on Chr 2, 3 and 9, one of which was located on Chr2: 34,654,608~ 35430715 accounted for 20.1%~ 25.4% of the variation of luteoxanthin auroxanthin, lutein, violaxanthin and the total carotenoid content. Overall, this study provides a genetic foundation for marker-assisted selection (MAS) breeding of nutritionally enhanced citrus fruit.

**KEYWORDS:** Citrus, SLAF, high-density genetic map, carotenoids, QuantitativeTrait Loci (QTL)

**INTRODUCTION**

Carotenoids are high-value antioxidant compounds widely present in most fruits and

vegetables, which are essential nutrition components of the human diet and are also

used in the cosmetic, feed and pharma sectors.1, 2 The abundance of this pigment

dictates color in fruits and vegetables, ranging from yellow and orange to red,3 which

results in the aesthetic properties of agricultural products for consumers. Citrus are one of the most economically important fruit crops in the world and is grown in more than 114 countries.4 As fresh fruits and juice products, they are an important carotenoid and Vitamin C source for the human diet. Unlike other important provitamin A carotenoid-rich fruits, such as carrot and tomato, citrus fruit predominantly accumulates non-provitamin A carotenoids (e.g., lutein, zeaxanthin, violaxanthin and β-cryptoxanthin). Although these non-provitamin A carotenoids are

not converted into vitamin A in humans, they are proven to reduce the risk of lifestyle-related diseases and age-related eye diseases.5, 6 Citrus fruit pulp shows a diverse array of carotenoid content and composition among different cultivars,7-9 which indicate that carotenoid traits are highly influenced by genetic factors.9 For example, Satsuma mandarin (*Citrus unshiu* Marc.) predominantly accumulates β-cryptoxanthin in the juice sacs of mature fruits, whereas sweet oranges (*Citrus sinensis* Osbeck) often accumulate 9-*cis*-violaxanthin and all-trans-violaxanthin. By contrast, cultivars with a white pulp color, including pummelo (*Citrus grandis* (L.) Osbeck), lemon (*Citrus limon* (L.) Burm. f.), and lime (*Citrus aurantifolia* (Cristm.) Swingle) contain low carotenoid amounts in the juice sacs.10 Citrus germplasms with a large variation of carotenoids could be used as ideal material for genetic studies of carotenoid traits and the further breeding of nutritionally enhanced foods and/or by-products. However, for the woody plant (such as citrus) the large tree size and long juvenile phase make the conventional breeding time-consuming and costly. Therefore, it is necessary to identify the genetic factors of carotenoid diversity and develop associated marker-assisted selection (MAS) approaches to facilitate the breeding processes. Quantitative trait loci (QTL) analysis is an efficient strategy for understanding the genetic mechanism of the complex quantitative traits of carotenoids. For instance, a QTL analysis showed that the carotenoid cleavage dioxygenase 4 (CCD4) and zeaxanthin epoxidase (ZEP) genes were the major loci responsible for the natural variation in the β-carotene content and β-carotene-derived xanthophylls in *Arabidopsis* seeds, respectively.11, 12 Likewise, a QTL of the carotenoids in maize kernels (*Zea mays*) demonstrated that allelic variants of BCH (β-carotene hydroxylases) and LCYe (lycopene ε-cyclase) were responsible for

the natural variation of the content and composition of carotenes and xanthophylls.13, Currently, genetic linkage maps are widely used in high-throughput QTL mapping for fruit quality traits in horticultural crops.15, 16 The density of the molecular markers of the genetic map in many cases determines the quality of the map, which in turn largely affects the QTL results. The reduced representation genome sequencing (RRGS) strategy is a rapid and cost-effective next-generation sequencing (NGS) technology for large-scale marker discovery and genotyping that can be used in the construction of a high-density genetic map,17-19 among which the Specific Locus Amplified Fragment

66 sequencing (SLAF-seq) is an efficient RRGS approach based on high-throughput pair-

Page 4 of 37

67 end sequencing.18 Indeed, the development of an SLAF-seq based high-density genetic

68 map contributes to an efficient QTL analysis of some important agronomic traits in

69 crops.20-23 Notably, a high-density genetic map has also been successfully used in the

70 QTL analysis of the carotenoid content in horticultural crops, such as carrot,

71 watermelon and broccoli.24-26 In citrus, low-density markers have been constructed for

72 a QTL analysis of carotenoid contents.27 However, its low density was obviously not

73 enough to provide complete information for the exact QTLs of the carotenoid

74 constituents in citrus fruits.

75 In this study, we constructed a SLAF-based high-density genetic map of citrus and

76 identified the QTLs and associated molecular markers underlying the carotenoid

77 variation, which would provide an effective molecular breeding strategy for future fruit

78 quality improvement.

79 **MATERIALS AND METHODS**

80 **Plant Materials.** An F1 pseudo-testcross population with 94 progenies was developed

81 by crossing *C. reticulata* (female parent) and *P. trifoliata* (male parent), with an embryo

82 rescue at 80 and 85 days after pollination, as previously described.28 This F1 population

83 was used for the QTL analysis of the carotenoid traits and was grown at the Huazhong

84 Agricultural University (Wuhan, China) under the same nutrition, field soil and

85 irrigation conditions. The mature stage of the fruit was determined as previously

86 reported.7 The HJ x ZK population was composed of 94 individuals, 79 of which fruited

87 in 2015. Fruits were harvested in middle of December, 2015. At least nine mature fruits

88 were harvested for each line with three biology replicates. The freeze-dried citrus pulp

89 was then used for the carotenoid analysis. Young leaves from the F1 plants and two

90 parents were used for the DNA extraction according to a previous description.29

91 **Measurement of the Fruit Color Index.** The fruit pulp color parameters L, a and b

92 scores were calculated using the KONICA MINOLTA CR-400 (Japan), with a

93 minimum of three biological replicates. The Citrus color index (CCI) (CCI=1000 x a/

94 (L x b)) was used to measure the fruit color as previously reported.30

95 **Carotenoid Profile.** The carotenoid extraction and analysis were performed according

96 to a previous description,31, 32 with some modifications. Briefly, the carotenoids were

97 extracted from the dried powder samples by using acetone: n-hexane: ethanol (1:2: 1

98 [v/v/v]), with 0.01% (w/v) 2,6-di-tert-butyl-4-methylphenol (BHT). The extracted

99 carotenoids were then saponified for nine hours using a 25% (w/v) KOH methanol

100 solution. The water-soluble impurities were removed from the crude carotenoid by

101 using a saturated NaCl solution. The carotenoids were then evaporated to dryness and

102 were re-dissolved in methyl-tert-butyl ether (MTBE) containing 0.01% BHT. Mobile

103 phase A (methanol with 10% (v/v) H2O) and mobile phase B (MTBE with 0.01% (w/v)

104 BHT) were used for the reverse phase high-performance liquid chromatography (HPLC)

105 analysis, which was performed in a Waters HPLC system equipped with a photodiode

106 array detection (PAD) as previously reported.31, 32 The carotenoids were eluted from a

107 YMC C30 carotenoid column (250 × 4.6 mm, 5 μm; Japan) at a flow rate of 1 mL/min

108 using the following linear gradient: 8% B to 25% B in 35 min, 25% B to 50% B in 15

109 min, 50% B to 70% B in 10 min, 70% B in 10 min, and back to 8% B for 10 min of re-

110 equilibration. The carotenoids were identified by a comparison of the characteristic

111 absorption spectra and the retention times based on the previous literature and authentic

112 carotenoid standards (CaroteNature, Lupsingen, Switzerland). The contents of

113 violaxanthin, neoxanthin, lutein zeaxanthin, β-cryptoxanthin, phytofluene, phytoene

114 and β-carotene were measured by calibration curves of authentic standards.

115 **SLAF Library Construction and High-throughput Sequencing.** The SLAF library

116 construction was performed as previously described.18, 20 Briefly, according to the in-

117 silico prediction of the appropriate restriction enzymes in the *Citrus clementina*

118 genome,33 HaeIII and RsaI (New England Biolabs, NEB, USA) were used to digest the

119 genomic DNA of the ninety-four lines of the F1 population and the parents. The Klenow

120 Fragment of DNA Polymerase I and dATP were subsequently used to add the Single-

121 nucleotide A overhang to these digested DNA fragments at 37 °C. Then, these A-tailed

122 DNA fragments were ligated with the PAGE-purified Duplex Tag-labeled Sequencing

123 adapters (Life Technologies, USA)) with T4 DNA ligase at 37 °C. Then, these diluted

124 restriction-ligation DNA samples were used to performed the PCR reaction with Q5®

125 High-Fidelity DNA Polymerase and PAGE purified primers

126 (AATGATACGGCGACCACCGA and CAAGCAGAAGACGGCATACG). The

127 purified PCR products were pooled and separated by a 2% agarose gel. Fragments

128 ranging from 264 to 364 base pairs (with indexes and adaptors) in size were excised

129 and then purified. The pair-end sequencing of the gel-purified products was loaded on

130 an Illumina HiSeq 2500 system (Illumina, Inc; San Diego, CA, USA) according to the

131 manufacturer’s recommendations.

132 **SLAF-seq Data Analysis and Genotyping.** The SLAF marker development and

Page 7 of 37

133 genotyping were performed according to previous reports.18, 20 Briefly, the low-quality

134 reads from the raw data, with a quality score <30, were filtered out, and the high-quality

135 raw reads were sorted to ninety-four individuals in the F1 population and the parents by

136 the alignment of the different barcode sequences. After trimming the barcode sequences

137 and the terminal 5-bp positions, the clean high-quality reads were mapped on the *Citrus*

138 *clementina* genome. These SLAF-seq data were deposited in NCBI short read archive

139 (SRA) database (accessions: SRR7628704-SRR7628799). Reads mapping to the same

140 genome position, with over 90% identity, were grouped into one SLAF locus.18 The

141 single nucleotide polymorphism (SNP) loci of all the polymorphic SLAF loci were

142 genotyped for consistency in the parents and ninety-four progenies. Because the two

143 parents of this F1 population were diploid citrus species, the loci with more than four

144 allele SLAF tags were incorrect. These tags were defined as repetitive SLAFs and were

145 subsequently filtered out. The other Polymorphic SLAF markers were classified into

146 eight segregation types hk × hk, ab × cd, ef × eg, lm × ll, nn × np, ab × cc and cc × ab,

147 aa × bb.

148 **Linkage Map Construction and QTL Analysis.** To ensure a high-quality genetic map

149 of a double pseudo-testcross F1 population, the high-quality polymorphic SLAF

150 markers, except the aa x bb genotype, were filtered by the following quality control for

151 the Linkage Map Construction. First, the average sequence depths should be >10-fold

152 in the parents. In addition, the SLAF markers with missing data in more than 40% of

153 the progenies were filtered. Finally, markers with a significant segregation distortion (P

154 < 0.01) were filtered out. The other high-quality SLAF markers were then used in the

155 construction of the genetic map.

156 HighMap and Joinmap 4.0 software were used for the construction of the HJ x ZK F1

157 population. According to the reference genome position of the SLAF markers, they

158 were grouped into nine LGs. A High Map Strategy was also used to order the SLAF

159 markers for each linkage group, with a single-linkage clustering algorithm, which had

160 a maximum recombination rate of 0.4 and a logarithm of odds (LOD) threshold ≥5.0.

161 The Kosambi mapping function was employed to convert the recombination

162 percentages to map distances.34 The two sex-specific (female and male) linkage maps

163 were separately developed. These two linkage maps were further integrated into a sex-

164 averaged linkage map by averaging lengths over anchored segments of two sex-specific

165 maps and by extrapolating or interpolating for molecular markers segregating in only

166 female or male parent.35 Finally, the quality of the high-density integrated genetic map

167 was evaluated by the Haplotype maps and heat maps according to the previous

168 description.18

169 MapQTL® version 6.0 software,36 with the Interval mapping (IM) algorithm, was used

170 for the QTL mapping. QTLs with LOD scores higher than the threshold with 1,000-

171 permutation test (PT) and Kruskal-Wallis (K-W) test at P<0.05 were declared

172 significant. The percentage of phenotype variance explained (PVE) for each QTL was

173 calculated in the population by using the MapQTL software. The sequence of the SLAF

174 markers associated with the tested QTL traits are listed in (Dataset S3).

175 **Statistical Analysis.** All the phenotypic data used in this study are presented as the

176 mean ± SD of at least three biological replications. The statistical analysis of the

Page 9 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

10

177 phenotypic data was done using the Microsoft Excel program (Microsoft Office, 2010)

178 and IBM SPSS 17.0 Statistics software. A principal component analysis (PCA) of the

179 carotenoids, among the tested lines in the population, was performed by using the

180 SIMCA- P 11.5 program (Version 11, Umetrics, Umea, Sweden), with a correlation

181 matrix analysis and a calculation of the eigenvalues and eigenvectors. A Pearson

182 correlation (r) analysis was performed by using ANOVA in the Microsoft Office 2010

183 and IBM SPSS 17.0 Statistics software.

184 **RESULTS**

185 **SLAF sequencing and genotyping of an F1 pseudo-testcross population.** A total of

186 94 F1 progenies from the crossing of *C. reticulata* (female parent) and *P. trifoliata* (male

187 parent) were used for DNA extraction and SLAF sequencing. The enzymes HaeIII and

188 RsaI were selected for DNA digestion based on an in-silico prediction of the appropriate

189 restriction enzymes in the *Citrus clementina* genome.33 In total, the DNA sequencing

190 of four SLAF libraries generated 41.85 Gb of raw data, with a Q30 of 91.05% and

191 guanine-cytosine (GC) content of 39.90%, consisting of 209,780,148 pair-end reads of

192 ~100 bp in length (Table 1). These reads were then used to develop SLAF-tags based

193 on the read clustering. As a result, 170,687 SLAFs were detected, of which the average

194 sequencing depths were 45.29-fold in the parent and 7.82-fold in each offspring (Table

195 1). Of these SLAFs, 69,708 (40.84% of the total SLAFs) were polymorphic SLAFs,

196 and 10,645 of the polymorphic SLAFs with segregation patterns of ab × cd, ef × eg, hk

197 × hk, lm × ll, nn × np, ab × cc, and cc × ab were defined as high-quality SLAFs, which

198 was suitable for the linkage map construction of the F1 population. To improve the

199 mapping accuracy, only 3,931 of the 10,645 high-quality SLAF markers, with more

200 than 60% integrity and more than 10-fold average sequence depths in the parents, were

201 used for the construction of a linkage map (Table 1).

202 **Construction of a high-density integrated genetic map based on the SLAF-seq**

203 **genotyping data.** A total of 3,817 out of the 3,931 high-quality polymorphic SLAF

204 markers were distributed into nine linkage groups (LGs) according to their locations in

205 the *Citrus clementina* genome. For the female map, 2,354 markers were positioned in

206 nine LGs spanning 1256.92 cM of map distance, with an average distance of 0.57 cM.

207 On the male map, 1,934 markers were assembled into nine LGs and was 1,580.71 cM

208 in length, with an average interval length of 0.99 cM (Table S1 and Figure S1). These

209 two separate parental linkage maps were further integrated into a sex-averaged linkage

210 map. The further integrated genetic map was 1,502.44 cM in length, with an average

211 distance of 0.43 cM between adjacent markers (Table S1). The length of nine LGs on

212 the integrated genetic map ranged from 124.03 to 220.48 cM, with average inter-marker

213 distances of 0.26-0.62 cM (Table S1). In addition, the number of markers varied from

214 260 to 769 for each LG, with an average of 424 markers per LG (Table S1). The

215 distribution of all the SLAF markers on nine LGs is displayed in (Figure 1). The

216 percentages of the gaps in which the inter-marker distance was smaller than five cM

217 (Gap ≤ 5 cM value) of the nine linkage groups varied from 96.90% to 100.00% (average,

218 98.69%) (Table S1). The maximum gap in this linkage map was 19.22 cM, which was

219 located in LG4 and LG5. The chi-square test demonstrated that the average frequency

220 of the segregation distortion markers (SDMs) was 8.20% (P< 0.01). These segregation

221 distortion regions may be a result of preferential selection and will not affect the

222 construction of the genetic map.20 In addition, we displayed collinearity between the

223 physical and genetic distances of all the SLAF markers in the nine LGs by mapping the

224 SLAF markers to the *Citrus clementina* genome (Figure S2). In summary, we

225 constructed a high-quality integrated genetic linkage map with high-density SNPs in

226 the F1 pseudo-testcross population of *C. reticulata* x *P. trifoliata* (HJ x ZK), which

227 should be suitable for a later genetic linkage analysis of carotenoid traits.

228 **Phenotypic Variation and Correlation Analysis.** As shown in Dataset S1, the

229 carotenoid composition and content in the citrus pulp was profiled by using the HPLC-

230 PDA analysis in the HJ x ZK F1 population. The descriptive statistics showed that the

231 content of lutein and zeaxanthin varied, with a transgressive segregation manner,

232 among the individuals in the HJ x ZK population (Table S2). The coefficient of

233 variation (CV) of the individual carotenoid in the population ranged from 0.49 ~0.70

234 (Table S2). The male parent ZK showed a lower total carotenoid content than the female

235 parent HJ. However, the content of colorless carotenoids (phytoene and phytofluene)

236 in ZK was higher than HJ. The HJ showed higher contents of almost all the carotenoid

237 components (except for zeaxanthin) compared to ZK and to all the progeny in the

238 population. All the carotenoid contents and the CCI (citrus color index) traits showed

239 continual variation (Figure 2) among the individuals in the F1 population. Lut, Zea,

240 Phytof, Total and CCI traits followed a skew normal or normal distribution

241 (Kolmogorov–Smirnov test, P-value>0.05), however, Vio, Neo, Luteo, Auro, β-cry, β-

242 car and Phy displayed non-normal (P-value<0.05) (Table S2 and Figure S3).

243 To test the correlation between the color index and the carotenoid content in the F1

244 population, we selected 32 individuals with two extreme color traits (16 orange and 16

245 yellow individuals) (Figure 3A) to perform the PCA analysis of the carotenoids.

246 Correspondingly, these 32 F1 individuals were well classified into two groups (Figure

247 3B). As shown in (Figure 3C-D), the female parent HJ and one offspring HZ-80 showed

248 higher contents of β-carotene, β-cryptoxanthin and the total carotenoid content in the

249 pulp as well as a higher CCI color index value than the male parent ZK and one

250 offspring HZ-71, which reinforced the important role of carotenoids in the color

251 formation of citrus pulp. Consistently, the Pearson correlation analysis among the

252 individuals in the F1 population also showed that the colorful carotenoids (auroxanthin,

253 lutein, zeaxanthin, etc.) showed a significant positive correlation with the CCI value,

254 whereas the colorless carotenoids (phytoene and phytofluene) showed a significant

255 negative correlation with the color index (Table 2). The total carotenoid was

256 significantly correlated with all the individual carotenoids, with correlation coefficients

257 r ranging from 0.5 to 0.74. With the exception of phytoene and phytofluene, the

258 correlation coefficients between the contents of the individual carotenoids were all

259 highly significant(P<0.05) and ranged between r=0.15 and 0.77 (Table 2).

260 **QTL Analysis of the Carotenoids Using the High-Density Genetic Map.** The QTL

261 analysis of 11 carotenoid traits (10 individual carotenoid traits and total carotenoid) and

262 the CCI color index trait was performed by using MapQTL6.0 software, with an SLAF-

263 based high density integrated linkage map, as mentioned above. Seventeen and two

264 significant QTLs were identified for these carotenoid contents and CCI values after

265 Permutation Testing (P<0.05) and K-W test (P <0.05), respectively (Table 3 and Figure

266 4). The percentage of phenotypic variance explained (PVE) by each QTL ranged from

267 17.5 to 30.4%, and the LOD scores were from 3.3 to 5.03 (Table 3). The sequences and

268 genotype information of the markers nearest to each QTL peak are listed in Dataset S3.

269 Two QTLs, located on Chr2 (with a PVE of 25.2%) and Chr5 (20.9%), were detected

270 for Auroxanthin; two QTLs on Chr2 (25.4%) and Chr9 (21.1%) were for Luteoxanthin;

271 two QTLs on Chr2 (20% and 19.5%) were for Neoxanthin; only one QTL on Chr2

272 (24.7%) and one on Chr3 (22.3%) was for lutein and β-cry, respectively; and the same

273 QTL on Chr9 was for both phytoene (17.5%) and phytofluene (18.1%). For the two

274 main carotenoids (violaxanthin and zeaxanthin) in the pulp of the progeny in the

275 population, two QTLs on Chr 3 (20.1%) and Chr6 (20.7%) were detected for

276 violaxanthin, and two QTLs on Chr3 (23.3%) and Chr6 (24.3%) were for zeaxanthin.

277 Three QTLs were identified for the total carotenoid trait, and two of which were located

278 on Chr 3 (20.1% and 21.6%) and the third one was located on Chr2 (22.0%). Both of

279 the QTLs for the color index (CCI) value were located on Chr2 (24.8% and 30.4%)

280 (Table 3). Among these QTLs, a QTL on Chr2 for total carotenoid trait was only

281 detected on the genetic map of female (HJ) with high content of total carotenoids (Table

282 S3). In addition, the segregation mode of the nearest marker closet to this QTL also

283 showed that the allele associated with high total carotenoid content was from female

284 (HJ) (Table 3), which suggested that this QTL locus in HJ might contain a positive

285 regulator for total carotenoid content.

286 As shown in Figure 4, the QTL co-localizations for multiple carotenoid traits were

Page 14 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

15

287 observed on Chr 2, 3 and 9. Five carotenoid traits (luteoxanthin, auroxanthin, lutein,

288 violaxanthin and total carotenoid trait) shared an overlapping QTL genetic region

289 (123.291~127.091 cM) at linkage group 2 between Marker1968071-Marker2096303

290 (Table S4), suggesting that these carotenoid traits were probably controlled by the same

291 genetic determinant. Among these traits, luteoxanthin showed the highest percentage

292 of phenotypic variance explained (PVE, 25.4%), with the highest LOD score of 5.03 in

293 this overlapping QTL genetic region. According to the genome positions of the SLAF

294 markers in the *Citrus clementina* genome,33 the genome interval of this overlapping

295 QTL was located at 34,654,608~35,430,715 on Chromosome 2 and contained 100

296 predicted protein encoding genes (Dataset S4). One of the genes involved in carotenoid

297 degradation encodes a nine-cis-epoxycarotenoid dioxygenase 5 (Table S4).

298 **DISCUSSION**

299 Carotenoids are a nutrition compound of the human diet and an important determinant

300 of fruit quality. Higher total and/or individual carotenoid components in citrus fruit is

301 an important objective for citrus breeding. The current work applied next-generation

302 sequencing based high-density genetic mapping in a genetic linkage analysis of

303 carotenoid compounds in citrus fruit, providing an effective molecular breeding

304 strategy for the development of nutritionally enhanced agricultural products.

305 **SLAF-seq-based construction of a high-density genetic map in citrus**

306 Developing a suitable mapping population is important for the construction of a high-

307 density genetic map and the subsequent QTL analysis. Unlike other agricultural crops,

308 in the fruit tree species, with a long juvenile phase, it is very hard to generate inbred

309 lines, such as Double haploid (DH) and Recombination Inbred Line (RIL) populations.

310 However, because of high heterozygosis, the F1 population in most tree species often

311 show substantial segregation, which is called the pseudo-testcross strategy.37, 38 The F1

312 pseudo-testcross populations have been used for linkage mapping studies that identify

313 QTLs and markers associated with agronomic traits.20, 39 We previously developed an

314 F1 pseudo-testcross population by crossing *C. reticulata* and *P. trifoliata*.28 These two

315 parents belong to two different citrus species, with highly polymorphic genotypes,33, 40

316 and thus, this cross successfully resulted in genotype segregation among progenies in

317 the F1 population. Notably, the progenies in the population showed a significant

318 variation in flesh color, with a diversity of carotenoid contents and composition (Figure

319 3), which allowed for the subsequent QTL analysis of carotenoids. Therefore, we

320 selected this F1 pseudo-testcross population to construct a high-density genetic map and

321 obtain the linkage mapping of carotenoids.

322 RFLP, RAPD, SSRs and EST-based traditional technologies were frequently used to

323 develop molecular markers in previous constructions of linkage maps in citrus.28, 41

324 However, the number of reported molecular markers (less than 300) was insufficient to

325 construct a high-density genetic map in citrus.42, 43 With the development of next-

326 generation sequencing (NGS) platform, NGS-based large-scale marker discovery and

327 genotyping had been widely applied in the construction of high-density genetic maps.17-

328 19 SLAF-seq is a cost-effective next-generation sequencing (NGS) technology that had

329 been applied for genetic map construction in multiple crops.21-23 In this study, we

330 demonstrated the application of SLAF-seq based markers on a large scale for the

Page 16 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

17

331 construction of an integrated genetic map in citrus. In total, 170,687 high-quality SLAF

332 markers were developed, and 69,708 polymorphic markers, with a high sequencing

333 depth, were identified in the F1 population, which reinforced that SLAF-seq is a rapid

334 and high-throughput strategy for genotyping and developing markers in citrus. Finally,

335 we constructed an integrated high-density genetic linkage map, with 9 linkage groups

336 and 3,817 developed SLAF markers. This genetic map covered nearly the whole citrus

337 genome and spanned 1,502.44 cM, with an average marker density of 0.43 cM/marker.

338 For the citrus genetic map constructed by more than 800 SNP markers, Ollitrault et al.42

339 established an initial medium density map spanning 1084.1 cM, with 961 markers in

340 *Citrus clementina*. Guo et al.44 published a 976.58 cM of pummelo, including 1563

341 markers developed by restriction site-associated DNA (RAD) sequencing. Recently,

342 Curtolo et al. published a *C. sunki* map with a total of 2,778 DArTseq markers and a

343 map of *P. trifoliata* with 3,084 DArTseq markers. 45 To our knowledge, the integrated

344 genetic map constructed in our study contained a substantially higher marker density

345 than the previously published genetic maps of *C. reticulata*. Consequently, our results

346 demonstrated that the SLAF-seq is indeed an efficient NGS strategy for high-density

347 genetic map construction in citrus.

348 **QTL mapping of carotenoids by using high-density genetic maps in citrus**

349 A wide range of variation in the carotenoid content of the citrus fruit pulp was observed

350 among the progenies in the F1 population. A frequency distribution analysis showed

351 that lutein, zeaxanthin, Phytofluene and total carotenoids traits tested in this study were

352 adjusted to a normal distribution in the F1 population, suggesting that these carotenoid

Page 17 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

18

353 traits are quantitatively inherited and may be controlled by multiple genes. The normal

354 distribution of lutein and total carotenoids traits was also observed in the floret of a

355 broccoli population,25 which revealed a common genetic inheritance of carotenoids in

356 different plant species. In our study, a total of 17 significant QTLs were detected for

357 the individual and total carotenoids traits, with LOD scores from 3.3 to 5.03 (Table 3).

358 To identify the significant QTLs, a logarithm of odds (LOD) threshold of 3.2 was used

359 with 1,000-permutation test (PT) at P<0.05. Therefore, we eliminated all the QTLs with

360 LOD values <3.2. Although a previous study identified 21 QTLs for 10 carotenoid

361 components in citrus pulp,27 only one QTL showed an LOD >3.2, which was possibly

362 a result of the limited number of progenies and molecular markers. Brown et al.25

363 reported that the percentage of variation explained (PVE) by each QTL detected for the

364 carotenoids varied from 6.2% to 24.1% in broccoli florets. A small PVE, ranging from

365 5.29% and 17.31%, was also observed in the QTL mapping of the sugars and acids in

366 grape berries.46 The PVE from each detected QTL in our study ranged from

367 17.5%~30.4%. These results showed that the fruit quality was generally determined by

368 numerous QTLs, and each might have a small genetic influence due to the quantitative

369 characteristics of the fruit traits.

370 Carotenoids not only confer nutritional properties but also determine the consumer’s

371 perception of quality, with regard to the aesthetic fruit color. As expected, the PCA of

372 the carotenoids among the representative F1 progenies with a distinct flesh color showed

373 that the carotenoids were highly associated with the color of the fruit pulp. Moreover,

374 the correlation analysis displayed that auroxanthin, lutein, zeaxanthin and the total

Page 18 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

19

375 carotenoid showed a positive correlation with the fruit color index CCI value, while the

376 colorless carotenoids, such as phytoene and phytofluene, were negatively correlated

377 with the color index (Table 2). Although no QTLs co-localizations were detected

378 between the CCI and carotenoid traits, two QTLs for CCI and six QTLs for carotenoid

379 traits (auroxanthin, luteoxanthin, neoxanthin, lutein, violaxanthin and total carotenoids

380 contents) were located on the same chromosome 2, with a distance of approximately 60

381 cM. These results may partly explain the aforementioned correlation between the flesh

382 color and carotenoid content traits in genetic insight in the F1 population.

383 Three QTLs co-localization were observed among the 17 QTLs detected for the

384 carotenoids, one of which accounts for the variation of luteoxanthin, auroxanthin, lutein

385 and violaxanthin and the total carotenoid content, which suggested that there existed

386 certain pleiotropic genes that regulated the network of carotenoid metabolism.

387 Consistently, a significant correlation (P≤0.01) among these five individual carotenoids

388 contents was observed in the F1 progenies (Table 2), suggesting the feasibility of

389 improving these carotenoids simultaneously in citrus pulp in future genetic strategies.

390 The overlapping genome interval for these carotenoid traits with the QTL co-

391 localization was located at Chr2: 34,654,608~35,430,715 in the Clementina genome.

392 Previous linkage mapping analyses show that most proven genetic intervals associated

393 with carotenoid variation often contain carotenoid biosynthetic genes and/or carotenoid

394 cleavage dioxygenase genes,11, 12, 25 which allowed us further narrow down the selection

395 of plausible causal genes in the QTL interval. The large scale SLAF markers developed

396 in this study can be used to rapidly identify the genome interval by a sequence

Page 19 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

20

397 alignment with the Clementina genome.

398 According to the gene annotation, the overlapping QTL interval, detected for five

399 carotenoid traits, contained 100 predicted genes (Dataset S4), including a carotenoid

400 degradation-related gene encoding a Nine-cis-epoxycarotenoid dioxygenase 5

401 (NCED5). Zhu et al. reported that NCED5 played an important role in the carotenoid

402 metabolism of flavedo in *C. reticulate*.47 NCED5 is responsible for the biosynthesis of

403 ABA,48 indicating it possibly indirectly affects the carotenoid content by modulating

404 ABA signaling in fruits.49 The possible function of this candidate gene on carotenoid

405 metabolism in citrus pulp remains to be elucidated by a further detailed molecular and

406 functional analysis.

407 In summary, the present study demonstrated that the NGS-based SLAF-seq is an

408 efficient strategy for the construction of a high-density genetic map in citrus, which can

409 be efficiently applied in a QTL analysis of fruit quality traits (carotenoids) in citrus.

410 The identification of the QTLs for carotenoids and the developed the SLAF markers

411 linked to the QTLs is helpful for the development of nutritionally enhanced citrus

412 varieties through MAS strategies

413 **ACKNOWLEDGMENTS**

414 We greatly appreciate Yue Huang (Huazhong Agricultural University) for assistance in

415 the analysis of the SLAF-seq data; Qingjiang Wu (Huazhong Agricultural University)

416 for assistance in collecting samples.

417 **SUPPORTING INFORMATION**

418 Distribution of the SLAF markers on nine linkage groups of the female and male genetic

419 map of the HJ x ZK population (Figure S1), collinearity analysis of the nine linkage

420 groups and *Citrus clementina* genome chromosomes (Figure S2), normal Q-Q plot of

421 carotenoid contents and color index of fruit pulp in HJ x ZK F1 population (Figure S3),

422 basic characteristics of the 9 linkage groups of the genetic maps (Integrated, female and

423 male) of the HJ x ZK population (Table S1), phenotypic performance of the carotenoid

424 content in the HJ x ZK F1 population (Table S2), QTLs detected for the color index

425 and the total and individual carotenoids by using female and male genetic map (Table

426 S3), candidate gene in the overlapping QTL region that may participate in carotenoid

427 metabolism (Table S4) (Support Information\_1.pdf).

428 Data of the carotenoid compositions and contents of the fruit pulp in the F1 population

429 (Dataset S1), data of the citrus color index (CCI value) of the fruit pulp in the F1

430 population (Dataset S2), the sequence genotype information of the SLAF markers

431 (Dataset S3), candidate genes in the overlapping QTL region for the five carotenoid

432 traits (Dataset S4) (Support Information\_2.xlsx).

433 **Funding**

434 This research was funded by the National Key Research and Development Program of

435 China (2018YFD1000200), National Natural Science Foundation of China (no.

436 31330066, 31521092 and 31630065), China Agriculture Research System (CARS-27),

437 the 111 project (B13034) and a China council scholarship (to X.Z.).

438 **Notes**

439 The authors declare no competing financial interest.

440 **REFERENCES**

441 (1) Fraser, P. D.; Bramley, P. M. The biosynthesis and nutritional uses of carotenoids.

442 *Prog Lipid Res*. **2004**, 43, 228-65.

443 (2) Krinsky, N. I.; Johnson, E. J. Carotenoid actions and their relation to health and

444 disease. *Mol Aspects Med*. **2005**, 26, 459-516.

445 (3) Yuan, H.; Zhang, J.; Nageswaran, D.; Li, L. Carotenoid metabolism and regulation

446 in horticultural crops. *Hortic Res*. **2015**, 2, 15036.

447 (4) Talon, M.; Gmitter, F. G. Citrus genomics. *Int J Plant Genomics*. **2008**, 2008.

448 (5) Abdel-Aal, E.-S. M.; Akhtar, H.; Zaheer, K.; Ali, R. Dietary sources of lutein and

449 zeaxanthin carotenoids and their role in eye health. *Nutrients*. **2013**, 5, 1169-1185.

450 (6) Sugiura, M.; Nakamura, M.; Ogawa, K.; Ikoma, Y.; Ando, F.; Shimokata, H.; Yano,

451 M. Dietary patterns of antioxidant vitamin and carotenoid intake associated with bone

452 mineral density: findings from post-menopausal Japanese female subjects. *Osteoporos*

453 *Int*. **2011**, 22, 143-152.

454 (7) Rodrigo, M. J.; Alquezar, B.; Alos, E.; Medina, V.; Carmona, L.; Bruno, M.; Al-

455 Babili, S.; Zacarias, L. A novel carotenoid cleavage activity involved in the biosynthesis

456 of Citrus fruit-specific apocarotenoid pigments. *J Exp Bot*. **2013**, 64, 4461-78.

457 (8) Ikoma, Y.; Matsumoto, H.; Kato, M. Diversity in the carotenoid profiles and the

458 expression of genes related to carotenoid accumulation among citrus genotypes. *Breed*

459 *Sci*. **2016**, 66, 139-47.

460 (9) Fanciullino, A. L.; Dhuique-Mayer, C.; Luro, F.; Casanova, J.; Morillon, R.;

461 Ollitrault, P. Carotenoid diversity in cultivated citrus is highly influenced by genetic

462 factors. *J. Agric. Food Chem*. **2006**, 54, 4397-4406.

463 (10) Kato, M.; Ikoma, Y.; Matsumoto, H.; Sugiura, M.; Hyodo, H.; Yano, M.

464 Accumulation of carotenoids and expression of carotenoid biosynthetic genes during

465 maturation in citrus fruit. *Plant Physiol*. **2004**, 134, 824-837.

466 (11) Gonzalez-Jorge, S.; Ha, S. H.; Magallanes-Lundback, M.; Gilliland, L. U.; Zhou,

467 A.; Lipka, A. E., et al. Carotenoid cleavage dioxygenase 4 is a negative regulator of

468 beta-carotene content in *Arabidopsis* seeds. *Plant Cell*. **2013**, 25, 4812-26.

469 (12) Gonzalez-Jorge, S.; Mehrshahi, P.; Magallanes-Lundback, M.; Lipka, A. E.;

470 Angelovici, R.; Gore, M. A.; DellaPenna, D. ZEAXANTHIN EPOXIDASE activity

471 potentiates carotenoid degradation in maturing *Arabidopsis* seed. *Plant Physiol*. **2016**,

472 pp. 00604.2016.

473 (13) Harjes, C. E.; Rocheford, T. R.; Bai, L.; Brutnell, T. P.; Kandianis, C. B.; Sowinski,

474 S. G.; Stapleton, A. E., et al. Natural genetic variation in lycopene epsilon cyclase

475 tapped for maize biofortification. *Science*. **2008**, 319, 330-333.

476 (14) Yan, J.; Kandianis, C. B.; Harjes, C. E.; Bai, L.; Kim, E.-H.; Yang, X.; Skinner,

477 D. J.; Fu, Z.; Mitchell, S., et al. Rare genetic variation at Zea mays crtRB1 increases β-

478 carotene in maize grain. *Nature. Genet*. **2010**, 42, 322.

479 (15) Galpaz, N.; Gonda, I.; Shem‐Tov, D.; Barad, O.; Tzuri, G.; Lev, S.; Fei, Z.; Xu,

480 Y.; Mao, L.; Jiao, C., et al. Deciphering genetic factors that determine melon

481 fruit‐quality traits using RNA‐Seq‐based high‐resolution QTL and eQTL mapping.

482 *Plant J*. **2018**, 94, 169-191.

483 (16) Zhong, Y.-J.; Zhou, Y.-Y.; Li, J.-X.; Yu, T.; Wu, T.-Q.; Luo, J.-N.; Luo, S.-B.;

484 Huang, H.-X. A high-density linkage map and QTL mapping of fruit-related traits in

Page 23 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

24

485 pumpkin (*Cucurbita moschata* Duch.). *Sci Rep*. **2017**, 7, 12785.

486 (17) Miller, M. R.; Dunham, J. P.; Amores, A.; Cresko, W. A.; Johnson, E. A. Rapid

487 and cost-effective polymorphism identification and genotyping using restriction site

488 associated DNA (RAD) markers. *Genome Res*. **2006**, 17, 000-000.

489 (18) Sun, X.; Liu, D.; Zhang, X.; Li, W.; Liu, H.; Hong, W.; Jiang, C.; Guan, N.; Ma,

490 C.; Zeng, H. SLAF-seq: an efficient method of large-scale de novo SNP discovery and

491 genotyping using high-throughput sequencing. *PloS one*. **2013**, 8, e58700.

492 (19) Glaubitz, J. C.; Casstevens, T. M.; Lu, F.; Harriman, J.; Elshire, R. J.; Sun, Q.;

493 Buckler, E. S. TASSEL-GBS: a high capacity genotyping by sequencing analysis

494 pipeline. *PloS one*. **2014**, 9, e90346.

495 (20) Zhang, J.; Zhang, Q.; Cheng, T.; Yang, W.; Pan, H.; Zhong, J.; Huang, L.; Liu, E.

496 High-density genetic map construction and identification of a locus controlling weeping

497 trait in an ornamental woody plant (*Prunus mume* Sieb. et Zucc). *DNA Res*. **2015**, 22,

498 183-91.

499 (21) Li, B.; Tian, L.; Zhang, J.; Huang, L.; Han, F.; Yan, S.; Wang, L.; Zheng, H.; Sun,

500 J. Construction of a high-density genetic map based on large-scale markers developed

501 by specific length amplified fragment sequencing (SLAF-seq) and its application to

502 QTL analysis for isoflavone content in *Glycine max*. *BMC genomics*. **2014**, 15, 1086.

503 (22) Zhang, Z.; Shang, H.; Shi, Y.; Huang, L.; Li, J.; Ge, Q.; Gong, J.; Liu, A.; Chen,

504 T.; Wang, D. Construction of a high-density genetic map by specific locus amplified

505 fragment sequencing (SLAF-seq) and its application to Quantitative Trait Loci (QTL)

506 analysis for boll weight in upland cotton (*Gossypium hirsutum*.). *BMC Plant Biol*. **2016**, 507 16, 79.

508 (23) Zhu, W.-Y.; Huang, L.; Chen, L.; Yang, J.-T.; Wu, J.-N.; Qu, M.-L.; Yao, D.-Q.;

509 Guo, C.-L., et al. A high-density genetic linkage map for cucumber (*Cucumis sativus*

510 L.): based on specific length amplified fragment (SLAF) sequencing and QTL analysis

511 of fruit traits in cucumber. *Front Plant Sci*. **2016**, 7, 437.

512 (24) Branham, S.; Vexler, L.; Meir, A.; Tzuri, G.; Frieman, Z.; Levi, A.; Wechter, W.

513 P.; Tadmor, Y.; Gur, A. Genetic mapping of a major codominant QTL associated with

514 β-carotene accumulation in watermelon. *Mol. Breed*. **2017**, 37.

515 (25) Brown, A. F.; Yousef, G. G.; Chebrolu, K. K.; Byrd, R. W.; Everhart, K. W.;

516 Thomas, A.; Reid, R. W., et al. High-density single nucleotide polymorphism (SNP)

517 array mapping in *Brassica oleracea*: identification of QTL associated with carotenoid

518 variation in broccoli florets. *Theor. Appl. Genet*. **2014**, 127, 2051-64.

519 (26) Iorizzo, M.; Ellison, S.; Senalik, D.; Zeng, P.; Satapoomin, P.; Huang, J.; Bowman,

520 M.; Iovene, M., et al. A high-quality carrot genome assembly provides new insights

521 into carotenoid accumulation and asterid genome evolution. *Nat. Genet*. **2016**, 48, 657-

522 66.

523 (27) Sugiyama, A.; Omura, M.; Matsumoto, H.; Shimada, T.; Fujii, H.; Endo, T.;

524 Shimizu, T.; Nesumi, H.; Ikoma, Y. Quantitative trait loci (QTL) analysis of carotenoid

525 content in Citrus fruit. *J. Japan. Soc. Hort. Sci.* **2011**, 80, 136144Taylor.

526 (28) Tan, M.; Song, J.; Deng, X. Production of two mandarin× trifoliate orange hybrid

527 populations via embryo rescue with verification by SSR analysis. *Euphytica.* **2007**, 157,

528 155-160.

529 (29) Cheng, Y.; Guo, W.; Yi, H.; Pang, X.; Deng, X. An efficient protocol for genomic

530 DNA extraction from Citrus species. *Plant Mol. Biol Report*. **2003**, 21, 177-178.

531 (30) Luo, T.; Xu, K.; Luo, Y.; Chen, J.; Sheng, L.; Wang, J.; Han, J.; Zeng, Y.; Xu, J.;

532 Chen, J.; Wu, Q.; Cheng, Y.; Deng, X. Distinct Carotenoid and Flavonoid

533 Accumulation in a Spontaneous Mutant of Ponkan (*Citrus reticulata* Blanco) Results

534 in Yellowish Fruit and Enhanced Postharvest Resistance. *J. Agric. Food Chem*. **2015**,

535 63, 8601-14.

536 (31) Cao, H.; Zhang, J.; Xu, J.; Ye, J.; Yun, Z.; Xu, Q.; Xu, J.; Deng, X. Comprehending

537 crystalline β-carotene accumulation by comparing engineered cell models and the

538 natural carotenoid-rich system of citrus. *J Exp Bot*. **2012**, 63, 4403-4417.

539 (32) Zheng, X.; Xie, Z.; Zhu, K.; Xu, Q.; Deng, X.; Pan, Z. Isolation and

540 characterization of carotenoid cleavage dioxygenase 4 genes from different citrus

541 species. *Mol Genet Genomics*. **2015**, 290, 1589-603.

542 (33) Wu, G. A.; Prochnik, S.; Jenkins, J.; Salse, J.; Hellsten, U.; Murat, F.; Perrier, X.;

543 Ruiz, M.; Scalabrin, S., et al. Sequencing of diverse mandarin, pummelo and orange

544 genomes reveals complex history of admixture during citrus domestication. *Nat.*

545 *Biotechnol*. **2014**, 32, 656-62.

546 (34) Van Ooijen, J., JoinMap 4. Software for the calculation of genetic linkage maps in

547 experimental populations. Kyazma BV, Wageningen, Netherlands **2006**, 33.

548 (35) Zhang, L.; Li, H.; Wang, J.; Linkage analysis and map construction in genetic

549 populations of clonal F1 and double cross. *G3 (Bethesda).* **2015**, 114.

550 (36) Ooijen, J.; Kyazma, B., MapQTL 6.0, software for the mapping of quantitative

551 trait loci in experimental populations of dihaploid species. Netherlands: Wageningen

552 **2009**.

553 (37) Grattapaglia, D.; Sederoff, R. Genetic linkage maps of *Eucalyptus grandis* and

554 *Eucalyptus urophylla* using a pseudo-testcross: mapping strategy and RAPD markers.

555 *Genetics*. **1994**, 137 (4), 1121-1137.

556 (38) Grattapaglia, D.; Bertolucci, F. L.; Sederoff, R. Genetic mapping of QTLs

557 controlling vegetative propagation in *Eucalyptus grandis* and *E. urophylla* using a

558 pseudo-testcross strategy and RAPD markers. *Theor Appl Genet*. **1995**, 90 (7-8), 933-

559 947.

560 (39) Pfender, W. F., Saha, M. C., Johnson, E. A., & Slabaugh, M. B. Mapping with

561 RAD (restriction-site associated DNA) markers to rapidly identify QTL for stem rust

562 resistance in *Lolium perenne*. *Theor Appl Genet* 2011, 122, 1467-1480.

563 (40) Wang, X.; Xu, Y.; Zhang, S.; Cao, L.; Huang, Y.; Cheng, J.; Wu, G.; Tian, S.;

564 Chen, C.; Liu, Y.; Yu, H., et al. Genomic analyses of primitive, wild and cultivated

565 citrus provide insights into asexual reproduction. *Nat. Genet*. **2017**, 49, 765-772.

566 (41) Chen, C.; Bowman, K. D.; Choi, Y. A.; Dang, P. M.; Rao, M. N.; Huang, S.; Soneji,

567 J. R.; McCollum, T. G.; Gmitter, F. G. EST-SSR genetic maps for *Citrus sinensis* and

568 *Poncirus trifoliata*. *Tree Genet Genomes*. **2008**, 4, 1-10.

569 (42) Ollitrault, P.; Terol, J.; Chen, C.; Federici, C. T.; Lotfy, S.; Hippolyte, I.; Ollitrault,

570 F.; Bérard, A., et al. A reference genetic map of *C. clementina* hort. ex Tan.; citrus

571 evolution inferences from comparative mapping. *BMC genomics*. **2012**, 13, 593.

572 (43) Gulsen, O.; Uzun, A.; Canan, I.; Seday, U.; Canihos, E. A new citrus linkage map

573 based on SRAP, SSR, ISSR, POGP, RGA and RAPD markers. *Euphytica*. **2010**, 173,

574 265-277.

575 (44) Guo, F.; Yu, H.; Tang, Z.; Jiang, X.; Wang, L.; Wang, X.; Xu, Q.; Deng, X.

576 Construction of a SNP-based high-density genetic map for pummelo using RAD

577 sequencing. *Tree Genet Genomes.* **2015**, 11, 2.

578 (45) Curtolo, M; Soratto, T. A. T.; Gazaffi, R.; Takita, M. A.; Machado, M. A.;

579 Cristofani-Yaly, M. High-density linkage maps for *Citrus sunki* and *Poncirus trifoliata*

580 using DArTseq markers. *Tree Genet. Genomes*. **2018**, 14(1), 5.

581 (46) Chen, J.; Wang, N.; Fang, L.; Liang, Z.; Li, S.; Wu, B. Construction of a high-

582 density genetic map and QTLs mapping for sugars and acids in grape berries. *BMC*

583 *Plant Biol*. **2015**, 15, 28.

584 (47) Zhu, F.; Luo, T.; Liu, C.; Wang, Y.; Yang, H.; Yang, W.; Zheng, L.; Xiao, X.;

585 Zhang, M.; Xu, R., et al. An R2R3‐MYB transcription factor represses the

586 transformation of α‐and β‐branch carotenoids by negatively regulating expression of

587 CrBCH2 and CrNCED5 in flavedo of *Citrus reticulate*. *New Phytol*. **2017**, 216, 178-

588 192.

589 (48) Frey, A.; Effroy, D.; Lefebvre, V.; Seo, M.; Perreau, F.; Berger, A.; Sechet, J.; To,

590 A.; North, H. M.; Marion-Poll, A. Epoxycarotenoid cleavage by NCED5 fine-tunes

591 ABA accumulation and affects seed dormancy and drought tolerance with other NCED

592 family members. *Plant J*. **2012**, 70, 501-12.

593 (49) Wu, Q.; Bai, J.; Tao, X.; Mou, W.; Luo, Z.; Mao, L.; Ban, Z.; Ying, T.; Li, L.

594 Synergistic effect of abscisic acid and ethylene on color development in tomato

595 (*Solanum lycopersicum* L.) fruit. *Sci Hort*. **2018**, 235, 169-180**Figure Captions**

Figure 1. Distribution of the SLAF markers on nine linkage groups of the integrated

map of the HJ x ZK population. The x-axis represents nine linkage groups, and the yaxis

indicates the genetic location (centiMorgan as unit). Every black bar indicates an

SLAF marker.

Figure 2. The phenotypic distribution of the CCI (citrus color index) value and the

individual and total carotenoids of the fruit pulp in the HJ x ZK population.

Figure 3. The correlation between the carotenoids and flesh color in citrus fruit. (A)

The segregation of fruit color in the F1 population. The fruits of the parents and

representative offspring with a distinct color are displayed. (B) A PCA analysis of the

carotenoids in the citrus pulp among these progenies described in (A). (C) and (D) The

CCI and carotenoid content in the two parents and two representative individuals in the

F1 population.

Figure 4. Distribution of the QTLs detected in the mapping population HJ × ZK. The

markers are shown on the right of the linkage groups. Only the linkage groups found to

contain a significant QTL (estimated by permutation tests and Kruskal-Wallis test) are

shown. The QTLs are represented by the bars of different colors and are shown on the

right of the chromosomes.

**Tables**

Table 1. SLAF-seq data summary for the HJ x ZK population

Data Name Value

Total reads

No. of reads

Reads in high-quality SLAFs

High-quality SLAFs

No. of SLAFs

Average depth in parents

Average depth in individuals

No. of polymorphic SLAFs

No. of high-quality SLAF markers

SLAF markers used for map construction

No. of total markers

No. of “ef x eg” type markers

No. of “ab x cd” type markers

No. of “lm x ll” type markers

No. of “nn x np” type markers

209,780,148

93,883,016

170,687

45.29

7.82

69,708

10,645

3,931

43 (1.09%)

449 (11.42%)

1503 (38.23%)

1936 (49.25%)

Note: No, number; SLAF, specific length amplified fragment.

32

Table 2. Pearson's correlation coefficients (r) between the citrus color index (CCI) value

and the levels of the individual and total carotenoids of the fruit pulp in the HJ x ZK

population.

Note: Vio, 9-cis-Violaxanthin; Luteo, Luteoxanthin; Lut, Lutein; Zea, Zeaxanthin; β-cry, β-

cryptoxanthin; β-car, β-carotene; Phy, Phytoene; Phytof, phytofluene; Neo, Neoxanthin; Auro,

Auroxanthin, \*P<0.05; \*\*P<0.01 others are not significant at P>0.05.

Trait Neo Luteo Auro Vio Lut Zea β-cry Phy Phytof β-car Total

Neo

Luteo 0.77\*\*

Auro 0.22\*\* 0.27\*\*

Vio 0.72\*\* 0.84\*\* 0.17\*\*

Lut 0.51\*\* 0.45\*\* 0.41\*\* 0.30\*\*

Zea 0.22\*\* 0.18\*\* 0.29\*\* 0.15\* 0.59\*\*

β-cry 0.26\*\* 0.35\*\* 0.38\*\* 0.25\*\* 0.56\*\* 0.47\*\*

Phy 0.13\* 0.21\*\* 0.08 0.23\*\* 0.08 0.01 0.17\*\*

Phytof 0.10 0.12 0.25\*\* 0.16\* 0.16\* 0.16\* 0.12 0.79\*\*

β-car 0.54\*\* 0.44\*\* 0.31\*\* 0.40\*\* 0.51\*\* 0.41\*\* 0.54\*\* 0.29\*\* 0.20\*\*

Total 0.60\*\* 0.64\*\* 0.50\*\* 0.59\*\* 0.74\*\* 0.58\*\* 0.60\*\* 0.56\*\* 0.63\*\* 0.61\*\*

CCI -0.09 -0.10 0.22\*\* -0.09 0.24\*\* 0.33\*\* 0.12 -0.37\*\* -0.24\*\* -0.10 0.32\*\*

Page 32 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

33

Table 3. Table 3. QTLs detected for the color index and the total and individual

carotenoids by using integrated genetic map.

Trait a

Linkage

group

Genetic

interval (cM) b

Nearest

marker c

LOD

max

K-W

test d

Exp

(%) e

Allele

configuration f

Allelic

source g

Auro 2 107.556~135.178 Marker2132416 4.99 14.281\*\*\* 25.2 ab x cd Male/Female

5 48.85~54.051 Marker1119700 4.02 12.217\*\*\*\*\* 20.9 nn x np Female (+)

Luteo 2 105.656~127.091 Marker2118551 5.03 15.262\*\*\*\*\* 25.4 lm x ll Male/Female

9 23.761~30.601 Marker1251818 4.07 6.260\* 21.1 lm x ll Male (-)

Neo 2 113.542~115.693 Marker2118551 3.84 9.496\*\*\* 20.0 lm x ll Male/Female

2 131.178~141.582 Marker2206293 3.72 7.114\*\* 19.5 ab x cd Male/Female

Lut 2 119.504~136.178 Marker2042721 4.87 7.865\*\* 24.7 nn x np Female (+)

Vio 2 123.291~141.034 Marker1968553 3.84 6.803\*\* 20.1 nn x np Male/Female

3 150.634~153.439 Marker500967 3.97 12.980\*\*\*\*\* 20.7 nn x np Female (+)

Zea 3 14.938~27.442 Marker444648 4.55 13.646\*\*\*\*\* 23.3 nn x np Female (-)

6 120.691~126.661 Marker1844582 4.79 15.174\*\*\*\*\* 24.3 nn x np Female (-)

β-cry 3 194.741~207.8 Marker580865 4.33 17.896\*\*\*\*\* 22.3 ab x cd Male/Female

Phy 9 72.974~74.324 Marker1289904 3.30 9.737\*\*\* 17.5 lm x ll Male (-)

Phytof 9 72.974~74.324 Marker1289904 3.43 8.262\*\*\* 18.1 lm x ll Male (-)

Total 2 123.291~127.091 Marker2007011 4.27 8.596\*\*\* 22.0 nn x np Female (+)

3 193.969~196.366 Marker470559 3.85 15.227\*\*\* 20.1 ab x cd Male (+)

3 204.538~210.104 Marker477926 4.17 12.488\*\*\*\*\* 21.6 lm x ll Male (+)

Page 33 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

34

CCI 2 50.698~56.074 Marker2163334 3.53 12.311\*\*\*\*\* 24.8 nn x np Male/Female

2 61.399~74.472 Marker1971126 4.48 18.967\*\*\*\*\* 30.4 ab x cd Female (-)

Note: a The details of the carotenoids names are shown in the note in Table 2; b The 95% confidence

genetic interval for the detected QTL identified by the genome-wide threshold of the PT; c The

marker located closest to the position of the peak of LOD scores for the QTL; d Significance level

of Kruskal-Wallis (K-W) test of nearest markers. K-M test was performed to identify the

significance of QTLs for carotenoid traits with non-normal distribution. \* P<0.05, \*\* P<0.01, \*\*\*

P<0.005, \*\*\*\* P<0.001, \*\*\*\*\* P<0.0005. e The phenotypic variance explained by an individual

QTL; f e The allele configuration (Male alleles x Female alleles) of nearest marker as coded in

JoinMap. g The parental source of the allele associated with extreme pools with highest (+) or lowest

(-) average carotenoids content and CCI value.

**Figure Graphics**

**Figure 1**

Page 34 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

35

**Figure 2**

**Figure 3**

Page 35 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

36

**Figure 4**

Page 36 of 37

ACS Paragon Plus Environment

Journal of Agricultural and Food Chemistry

37

**Table of Contents Graphics**

Page 37 of 37

ACS Paragon Plus Environment