

The Genetic Architecture of Strawberry Yield and Fruit Quality Traits

Short title:

Genetic Architecture of Strawberry Traits

Helen M. Cockerton^{1*}, Amanda Karlström¹, Abigail W. Johnson¹, Bo Li^{1,2}, Eleftheria Stavridou¹,
Katie J. Hopson¹, Adam B. Whitehouse¹, Richard J. Harrison¹

Helen Maria Cockerton^{1*}, Helen.Cockerton@niab.com

Amanda Karlström¹, Amanda.Karlstrom@niab.com

Abigail Johnson¹, Abi.Johnson@niab.com

Bo Li^{1,2}, Bo2.Li@uwe.ac.uk

Eleftheria Stavridou¹, Eleftheria.Stavridou@niab.com

Katie Hopson¹, Katie.Hopson@niab.com

Adam Whitehouse¹, Adam.Whitehouse@niab.com

Richard Jonathan Harrison^{1,3}, Richard.Harrison@niab.com

*Corresponding author

Helen.Cockerton@niab.com; +44 1732 523750

¹NIAB EMR, East Malling, United Kingdom

²University of the West of England, Bristol, United Kingdom

³NIAB, 93 Lawrence Weaver Road, Cambridge, CB3 0LE, United Kingdom

Helen Cockerton 0000-0002-7375-1804

Bo Li 0000-0003-0077-4383

Richard Harrison 0000-0002-3307-3519

31 Abstract

32 Over the last two centuries breeders have drastically modified the fruit quality of strawberries
33 through artificial selection. However, there remains significant variation in quality across
34 germplasm with scope for further improvements to be made. We report extensive phenotyping of
35 fruit quality and yield traits in a multi-parental strawberry population to allow genomic prediction
36 and QTL identification, thereby enabling the description of genetic architecture to inform the
37 efficacy of implementing advanced breeding strategies.

38

39 A trade-off was observed between two essential traits: sugar content and class one yield. This
40 result highlights an established dilemma for strawberry breeders and a need to uncouple the
41 relationship, particularly under June-bearing, protected production systems comparable to this
42 study. A large effect QTL was associated with perceived acidity and pH whereas multiple loci
43 were associated with firmness, we therefore recommend the implementation of both MAS and
44 genomic prediction to capture the observed variation respectively.

45

46 Ultimately, our results suggest that the best method to improve strawberry yield is through
47 selecting parental lines based upon the number of marketable fruit produced per plant. Strawberry
48 number metrics were less influenced by environmental fluctuations and had a larger additive
49 genetic component when compared to mass traits. As such, selecting using “number” traits should
50 lead to faster genetic gain. Finally, we identify a large effect locus associated with an increase in
51 class one fruit.

52

53 **Key Words:** Organoleptic, Flavour, Acidity, Achene, QTL mapping, Breeding, Yield, Genomic
54 prediction

55

56

57

58

59 Background

60 Wild strawberry fruits have evolved to attract frugivorous animals. The sweet flesh provides
61 nutrition in return for endozoochory or the dispersal of seeds (1). Achenes - the true fruits, are
62 distributed around the pseudo fruit or receptacle of a strawberry, thus ensuring that partial eating
63 of a berry is likely to result in the ingestion of seeds. In fact, digestion of seeds is required for the
64 “activation” of germination potential and therefore completion of the natural strawberry life cycle
65 (2-4). The mutualism between birds or mammals and strawberries has led to natural selection for
66 seed-disperser “desired” fruit quality traits; indeed the change in colour that develops upon
67 ripening can act as a visual signal that ripe fruit contain seeds ready for dispersal (5) and some
68 volatile organic compounds have been implicated as attractants (6–8). Thus, wild strawberries
69 have been naturally selected to attract dispersers. By contrast, breeders aim to artificially select
70 strawberries to possess “human-desirable” fruit quality traits with the ultimate aim of increasing
71 consumer consumption.

72
73 In 1766, the French botanist Duchesne was the first person to characterise *Fragaria* × *ananassa*
74 strawberry plants resulting from a hybridisation event between two octoploid species (9). *F.* ×
75 *ananassa*, named after its pineapple aroma (ananas), soon became the dominant cultivated
76 strawberry species and systematic breeding was subsequently implemented to improve fruit size
77 and vigour of strawberry plants (9). In more recent history, strawberry breeders have succeeded in
78 improving strawberry marketable yield and to a lesser extent fruit quality (10,11). Indeed, fruit
79 quality is a complex trait that is made up of multiple visual (uniformity, colour), organoleptic
80 (flavour, texture) and sensory (firmness) factors (12). Nonetheless, poor fruit quality can lead to
81 the rejection of high yielding cultivars, by grower consortia and consumers (13) and thus
82 improving strawberry fruit quality is a complex undertaking. Flavour is a key component of fruit
83 quality, which requires a balance of sugar and acid; with a high total soluble sugars: titratable acid
84 ratio believed to represent a better tasting fruit for the UK market (7,14,15). However, multiple
85 other factors have been found to significantly impact flavour (16), including the secondary
86 metabolites associated with a peach flavour (γ -decalactone)(17) and burnt caramel flavour
87 (mesifuran) (18).

88

89 Despite extensive strawberry improvement over the centuries, there remains large variation in
90 strawberry fruit quality and consistency, both within and between cultivars due to influences of
91 environmental factors (16,19). Robust phenotyping protocols will allow accurate selection to
92 capture this variation, maximise genetic gain and improve desirable traits. Organoleptic traits are
93 complex and are predominantly assessed through subjective means, nonetheless robust protocols
94 have been established (20). Scientific sensorial evaluation can be undertaken by tasting panels who
95 are trained to detect the presence and magnitude of aromas, textures and flavours (20). However,
96 the costs associated with such an organoleptic analysis are prohibitive for pre-breeding and early-
97 stage selection purposes (21). Furthermore, such tests have limited application in breeding as they
98 do not indicate whether a trait is desirable; for which, the preference of a trait must be assessed by
99 a consumer panel.

100

101 The ultimate aim of breeding is to produce varieties yielding fruit that achieve an enjoyable multi-
102 sensorial eating experience leading to repeated consumer purchasing. Initial purchases have been
103 shown to be based on appearance, however flavour and quality were indicative of repeat
104 purchasing (22). Indeed, the most influential factors on USA consumer purchases have been rated
105 as taste and produce freshness (23) with strawberry sweetness and complex flavours as the most
106 highly prized attributes, whereas nutritional content was not valued (24). These complexities make
107 fruit quality hard to dissect and leads breeding to be classified as more of an art than a science.
108 Nonetheless, here we ask 1) to what extent can we parameterize and standardise sensory fruit
109 quality assessment, 2) can robust measures truly act as a surrogate for a human scoring system and
110 3) can we implement advanced breeding strategies using subjective data sets in a fashion able to
111 assist breeding for fruit quality? Here we discuss our approach and findings whilst acknowledging
112 the subjectivity of some measures and discuss the potential applications for breeding.

113 Molecular breeding is considered to be an effective strategy to select for traits that are expensive
114 or difficult to phenotype. Marker Assisted Selection (MAS) can improve traits that are controlled
115 by a small number of major effect genes (25). By contrast, genomic prediction can abbreviate the
116 period associated with fixing polygenic traits of complex inheritance. Genomic prediction requires
117 two phases - first the training phase and secondly the validation/ selection phase (26). Genomic
118 prediction results in the generation of genomic estimated breeding values which assist the early
119 identification of good parental lines and progeny lines allowing rapid generation cycling, and a

120 reduction of the breeding cycle time. A reduced breeding cycle time results in faster genetic gain
121 thus creating a competitive advantage for breeding companies. Genomic selection approaches have
122 revolutionised animal breeding, to great success (27–30). The efficacy of genomic selection in
123 strawberries has already been established, with a selection efficiency of 74% observed in
124 increasing average fruit weight (31). Balancing the costs of genotyping with the potential benefits
125 of rapid genetic gain is a critical balance for plant breeders. The work outlined here illustrates the
126 benefits that may result from adopting genetic breeding strategies.

127 Here we study a multi-parental population of strawberry to assess the phenotypic relationships
128 between fruit traits, we assess the potential to improve each trait and the level of variation present
129 within the population and finally we report the presence of QTL associated with traits and
130 determine the potential efficacy of genomic selection breeding approaches. We present a
131 comprehensive analysis of the genetic components influencing fruit quality and yield traits in
132 strawberries and discuss how our findings may help to optimise strawberry breeding through the
133 implementation of genomic approaches.

134 **Materials and Methods**

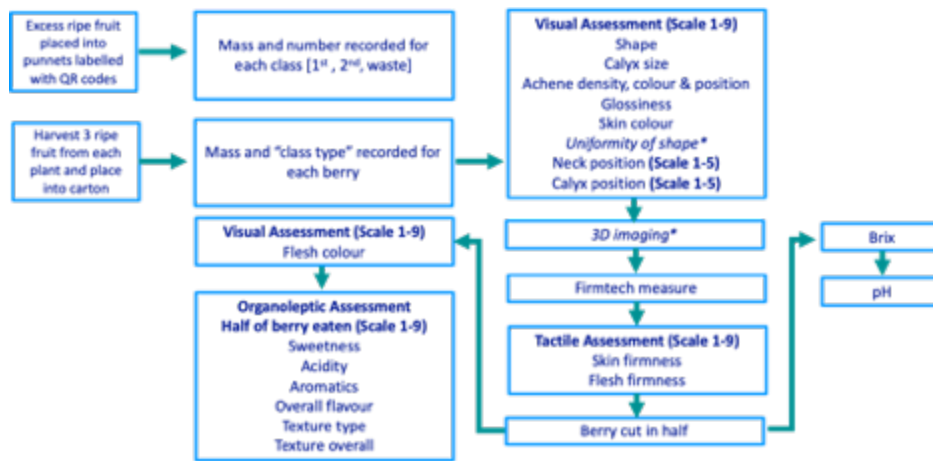
135 **Plant material and experimental set-up**

136 The multi-parental strawberry population used in this study was designed to segregate for multiple
137 fruit quality traits. Interrelated crosses between 26 parental lines were made to produce 26 families
138 of up to 16 individuals. Parental and grandparental lines were included in the population where
139 possible. A total of 270 genotypes and 28 progenitors were assessed in this study. Plants were
140 raised and allowed to go dormant over the autumn and early winter before being placed in a -2 °C
141 cold store. After five months, one cold-stored strawberry plant per genotype was potted up into
142 coir and grown under ambient polytunnel conditions. Subsequent replicate plants of each genotype
143 were removed from cold store at three-week intervals, with each cohort of plants forming a
144 replicate block. Five replicate blocks of plants were set up along table-top gutters within covered
145 polytunnels. The experiment was situated at NIAB EMR, Kent, UK (51° 17' 24.202'' N 0°
146 26'50.918'' E) along two 150 m long polytunnels covered in 150-micron plastic covers. Even

147 pollination was assisted through the addition of a Natupol Koppert bumble beehive into each
148 tunnel. Plants were grown in coir in 2 L pots, and fertigation was supplied at 1kg Vitex Vitafeed
149 (N:P:K, 176:36:255) L⁻¹ (10 s⁻¹ 45 m). Replicated blocks represented both planting date and tunnel
150 position, picking date varied for each berry as strawberries were picked when ripe between 11th
151 July and 8th November 2018, fruit were picked every weekday and assessed on the day of picking.
152 Fruit quality traits were assessed using three berries where possible for each replicate plant across
153 the 5 blocks. Yield metrics were assessed on every pick and later summed to provide a total end
154 of season value for assessment.

155 **Phenotyping**

156 The phenotyping process is detailed in Figure 1. Ripe fruits were harvested into individual punnets
157 for each genotype, and berries were then classified based on size and quality (class 1; 28- 45 mm
158 diameter, class 2; <28 mm diameter and waste; either misshapen/ physiological/ pathological
159 damage) and the number and mass of berries per plant and per class were recorded. Primary and
160 secondary ripe strawberries (as defined by Savini et al, 2005 (32)) were hand selected into
161 segmented cartons before measurement. Punnets and cartons were labelled with QR codes to allow
162 data entry using the Field Book app (33). Visual, tactile and organoleptic strawberry traits were
163 scored on a nine- or five-point scale (Figure 1), with score standardisation training provided for all
164 assessors. Trait assessment descriptors, alongside the nine discrete categorical shape and texture
165 categories, can be found in Suppl. Table 1. Traits were rated for importance in breeding on a scale
166 from 1 (not important) to 9 (highly important) as defined by breeders at NIAB EMR. 3D imaging
167 was conducted as outlined in Li et al., (2020) (34), the height to width ratio (H/W) was calculated
168 using 3D berry images and used to represent strawberry shape. Firmness measures were taken
169 using a FirmTech FT7 machine (UP GmbH, Ibbenbüren, Germany). Berries were cut
170 longitudinally to allow half of the berry to be assessed for organoleptic properties by one of four
171 assessors. Total soluble sugars and pH were measured from juice squeezed from the remaining
172 half of the berry using a refractometer meter (Atago PAL 1) and pH meter (LAQUA twin B-712),
173 respectively. Halved strawberry samples did not provide sufficient juice to measure titratable
174 acidity.



175

176

177 **Figure 1** The strawberry phenotyping process from the picking of strawberries through to
178 destructive assessments. Each box represents a discrete phenotyping station *Uniformity of shape
179 and 3D imaging have been reported by Li et al. (2020) (34).

180 Genotyping and Linkage map

181 DNA was extracted from the population using the Qiagen DNeasy plant mini extraction kit. The
182 Axiom® IStraw35 384HT array (i35k) was used for genotyping (35) and the NIAB EMR
183 strawberry consensus map was used to define marker positions (36). *Fragaria* × *ananassa*
184 chromosome number is denoted by 1-7 and the sub-genome number is represented by A-D as
185 specified in van Dijk et al. (2014) (37) and Sargent et al. (2015) (38). A total of 18,790 markers
186 segregated in the population.

187 Statistical Analysis

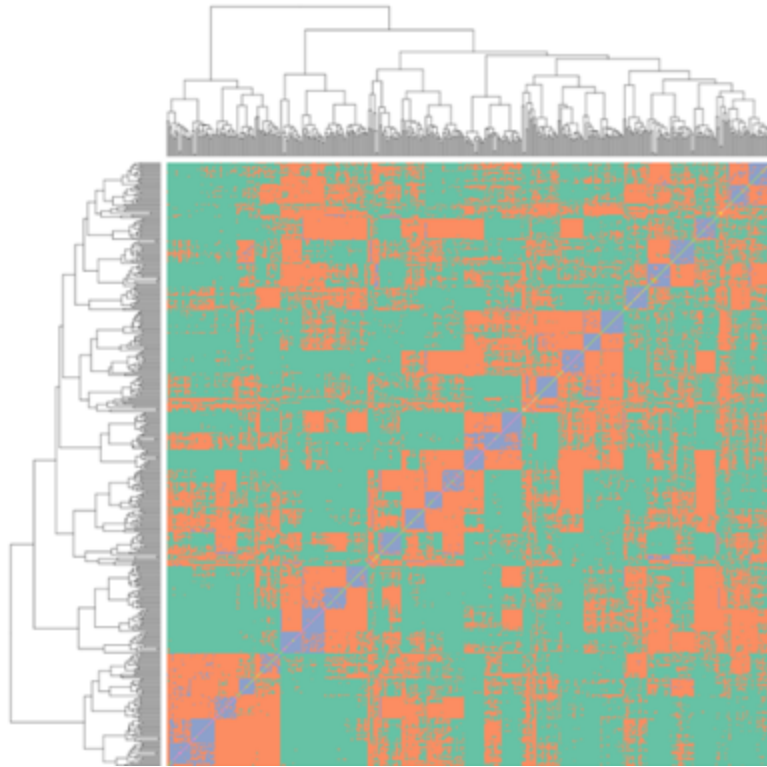
188 The best linear unbiased estimates (BLUE) were calculated for each genotype and trait using a
189 linear mixed effect model that included the cofactors of assessor, individual, picking date and
190 block. The model type fitted was specified individually for each trait as detailed in Suppl. Table
191 1. Significant co-variates were identified through comparison of a mixed model (phenotype ~
192 genotype + block + individual + date + assessor) to a model omitting the trait of interest,
193 comparisons were made using a likelihood ratio test. Significant genotype x environment (GxE)
194 interactions were assessed as specified for co-factors above but with the inclusion of the date of

195 picking x genotype interaction variable. Heritability values were calculated using the r package
196 “heritability” (39) where $H^2 = \sigma G^2 / (\sigma G^2 + \sigma E^2 / r)$ was calculated based on analysis of variance
197 statistics where r is replicate number, G represents genotypic variance and E represents residual
198 error. Narrow sense heritability was calculated by $h^2 = \sigma A^2 / (\sigma A^2 + \sigma E^2)$ where A represents
199 additive genetic variance, where the relationship matrix was calculated using the R package
200 “snpReady” (40). Phenotypic correlations were calculated using the R package “psych” (41) and
201 plotted using the R package “corrplot” (42), *p* values were adjusted for multiple testing.

202 **Genomic Analysis**

203 The R package “snpReady” was used to generate a genetic relationship matrix (Figure 2) and the
204 R package “rrBLUP” was used to conduct GWAS analysis (43). The rrBLUP model was $y = Zg$
205 $+ S\tau + \varepsilon$, where y is phenotypic observations, Z and S are matrices of 0s and 1s representing the
206 fixed effects of; β the population structure, g the genetic background and τ the additive SNPs (44).
207 GWAS was conducted with the genetic relationship covariance matrix added as a random effect
208 and a minor allele frequency set to 5%. A Bonferroni corrected *p* value of 0.001 was used to
209 identify significant QTL. R^2 of QTL effect size was calculated using a linear model comparing
210 BLUE calculated values versus predicted values assuming an additive relationship between focal
211 SNPs. A genomic best linear unbiased prediction (GBLUP) was calculated using the software
212 ASReml-R. A fivefold random subdivision of the population into the ‘training’ (80%) and ‘test’
213 (20%) was used as suggested by Erbe et al. (2010) (45). The genomic selection GBLUP linear
214 mixed model specified a variance structure which combined genotype and the inverse genetic
215 relationship matrix as random variables. Predictive ability was defined by the correlation between
216 the predicted and BLUE score for the test population over 100 permutations with random selection
217 of the genotypes forming the ‘test’ and ‘training’ population, thus allowing us to determine the
218 predictive ability of the model. Prediction accuracy was calculated as detailed in Gezan et al.
219 (2017)(31).

220



221

222

223 **Figure 2** Genetic relationship matrix for the strawberry multi-parental population, blue colouring
224 represents the full sibling relationships, orange represents half-sibling relationships between
225 individuals, green represents less than half-sibling relationships. The relationship within the 26
226 families can be observed in the blue squares along the diagonal.

227 **Results**

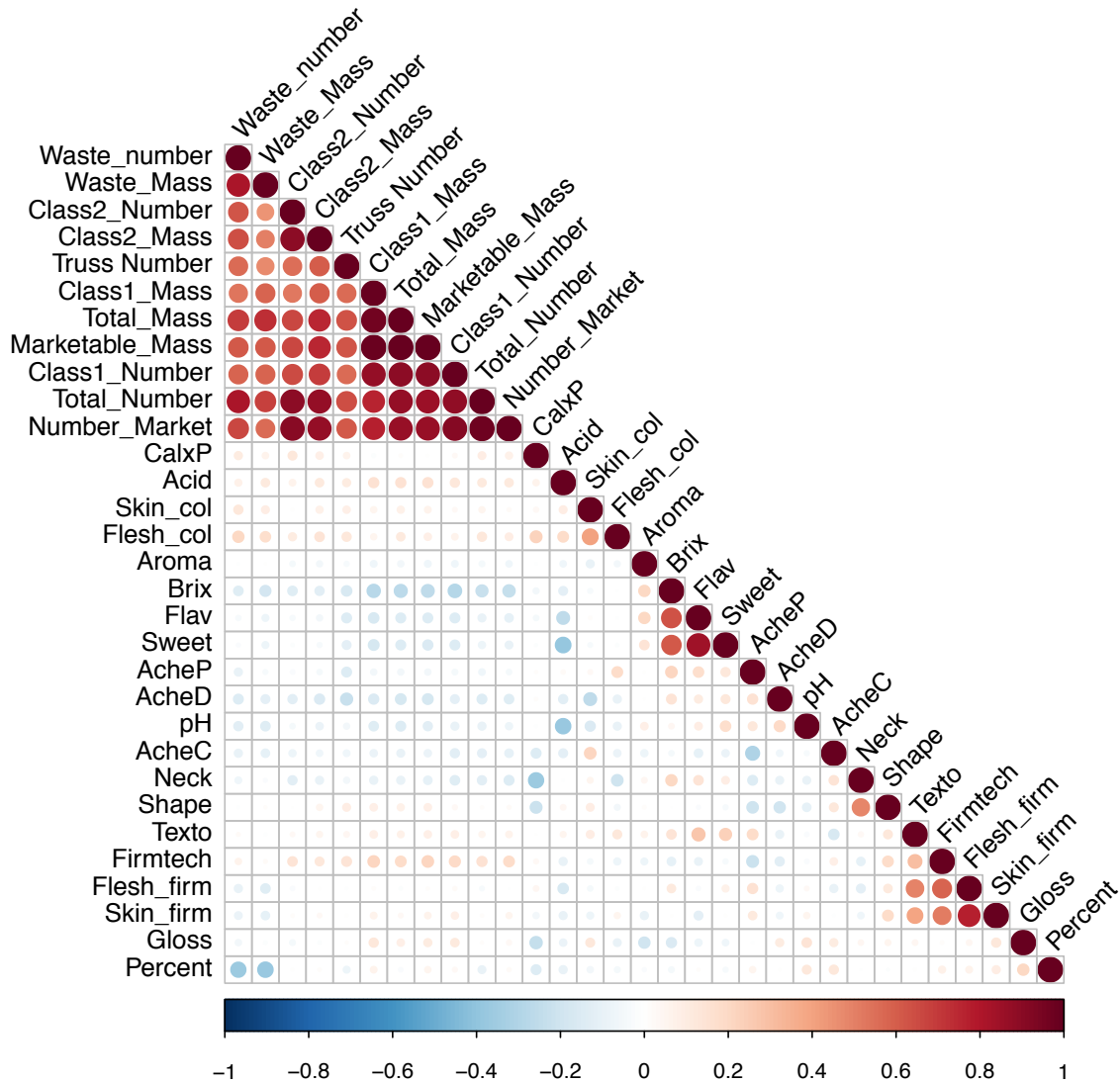
228 **Covariates**

229 A total of 21 strawberry fruit quality and 11 yield traits were measured as part of the fruit
230 phenotyping platform (Suppl. Table 2 & 3). Strawberries fruit from 270 genotypes were assessed
231 in five separate plantings replicated across the season. All measured traits were found to have
232 significant genetic and environmental components. Date of picking and block significantly
233 influenced all traits. However, variation in block was superseded by variation in picking date for
234 the following traits: flesh colour, acidity perception, sweetness perception, pH and flavour
235 perception. When assigned as a factor, the assessor was found to influence the scores for multiple

236 traits, however, interestingly the assessor did not significantly influence the scores of skin colour,
237 acidity perception, achene density, achene colour and flesh firmness (Suppl. Table 2). Significant
238 GxE terms indicate that different genotypes do not produce a consistent response across
239 environments.

240 **Phenotypic correlations between fruit quality and yield traits**

241 Flavour, sweetness perception and total soluble sugars were all shown to be positively correlated
242 ($p < 0.00001$; $r > 0.6$; Figure 3). Skin firmness, flesh firmness, automated firmness and texture
243 ratings were positively correlated ($p < 0.00001$; $r > 0.29$). Both sweetness perception ($p < 0.00001$;
244 $r = -0.38$), and to a lesser extent flavour ($p < 0.001$; $r = -0.28$), were correlated with acidity
245 perception, indicating acidity may be required for a good flavour. Negative relationships between
246 total soluble sugars and class 1 yield metrics indicate that high yielding June-bearing varieties were
247 associated with a potential trade-off ($p < 0.05$, $r = -0.22$).

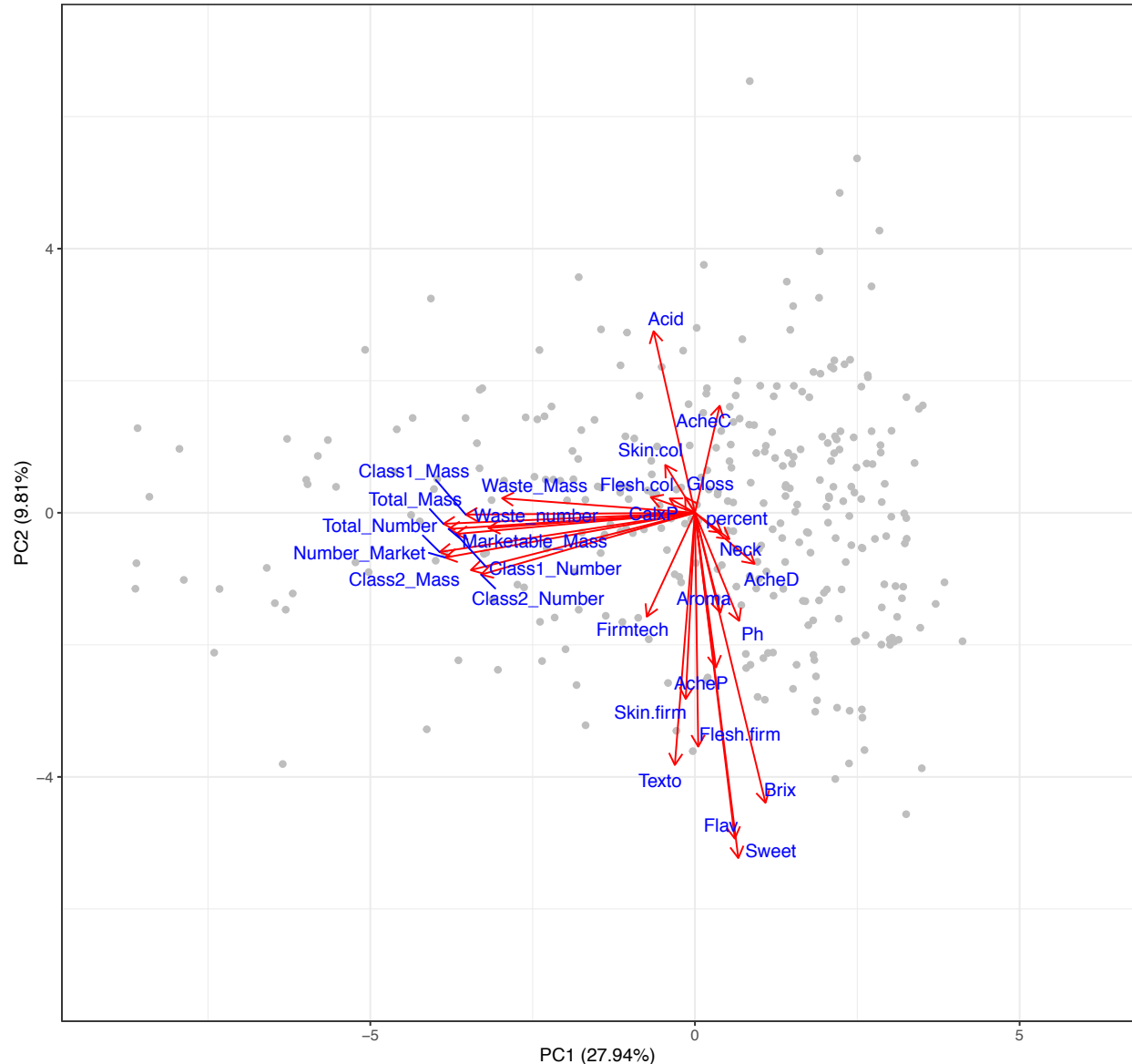


248

249 **Figure 3** Correlation matrix between the fruit quality and yield traits within the multi-parental
 250 strawberry population. Strength of colour denotes the magnitude and direction of the correlation
 251 coefficient. Size of the circle denotes significance value. CalxP - calyx position, Skin_col - skin
 252 colour, Flesh_col - flesh colour. Acid - acid perception, AcheC - achene colour, Neck - neck
 253 position, Shape- height:width, Texto - texture rating overall, Firmtech – Firmness - instrument,
 254 Flesh_firm - flesh firmness manual, Skin.firm - skin firmness, Gloss - glossiness, Percent -
 255 percentage of marketable fruit, AcheP - achene position, AcheD - achene density, Aroma -
 256 aromatics, Brix - total soluble sugars, Flav – flavour perception, Sweet - sweetness perception.

257 **Trait Variation**

258 The power to alter traits, in general, depends upon the presence of the variation within the breeding
259 germplasm. Therefore, visualisation of variation is required to define the boundaries within which
260 traits may be improved. The variation present within the multi-parental population is depicted in a
261 biplot (Figure 4). PC1 accounted for 27.9% of the variation and was largely correlated with fruit
262 number and mass, whereas PC2 represented 9.81% of the variation and was correlated with
263 organoleptic traits. Broad-sense heritability values show that between 3 and 90 % of the variation
264 observed in traits was controlled by genetic factors, whereas narrow-sense heritability scores show
265 that between 0 and 45 % of the variation was due to additive genetic effects (Suppl. Table 2).



266

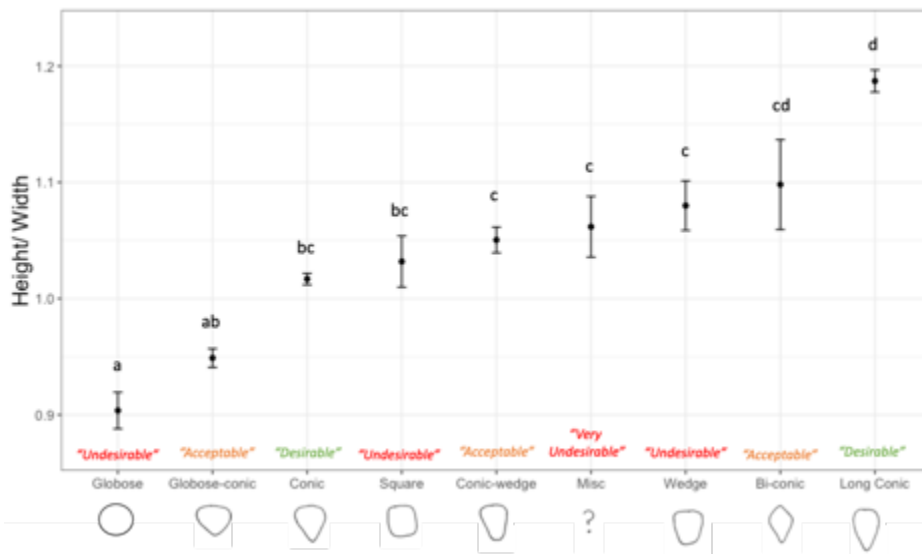
267

268 **Figure 4** Biplot representing variation in fruit quality and yield traits within the multi-parental
269 strawberry population. Numbers in brackets represent the proportion of variation explained by
270 principal components (PC). Red arrows indicate the relative influence a trait has on the PC each
271 associated with the trait denoted by a blue label. Grey points represent genotypes. CalxP - calyx
272 position, Skin.col - skin colour, Flesh.col - Flesh colour. Acid - Acidity Perception, AcheC -
273 Achene colour, Neck - Neck position, Shape- height:width, Texto - Texture rating overall, Fimtech
274 - Automated Firmness, Flesh. Firm - Flesh firmness manual, Skin.firm - Skin firmness, Gloss -
275 Glossiness, percent - percentage of marketable fruit, AcheP - Achene position, AcheD - Achene

276 density, Ph - pH, Aroma - Aromatic strength perception, Brix - Total Soluble Sugars, Flav -
277 Flavour, Sweet - Sweetness perception.

278 Objective Measure of Shape

279 As shape is an ordinal trait, a quantitative measure of strawberry shape was adopted; the height to
280 width ratio (H/W) of each berry. H/W is a continuous trait which allows data from across the
281 population to be used in genetic analysis. No QTL were associated with H/W however the
282 prediction accuracy (0.4) of this trait indicated a genomic selection approach could be effective.
283 Nonetheless H/W could not distinguish between “desirable” and “undesirable” strawberry shapes
284 (Suppl. Figure 1). The lack of relationship represents a discord between the desirability of a given
285 shape (as detailed in Li et al. 2020 (34)) and the biologically measurable trait H/W. However, H/W
286 or a similar metric, is needed to study the underlying genetic components associated with the trait
287 and thus allow the modification of shape through genome informed breeding.



288

289

290 **Supplementary Figure 1** Average height to width ratio for each manually classified strawberry
291 shape category. Desirability in coloured text terms denote the breeding goals for strawberry shape
292 within the UK. Misc – Miscellaneous undulating misshapen fruit without a clear shape.

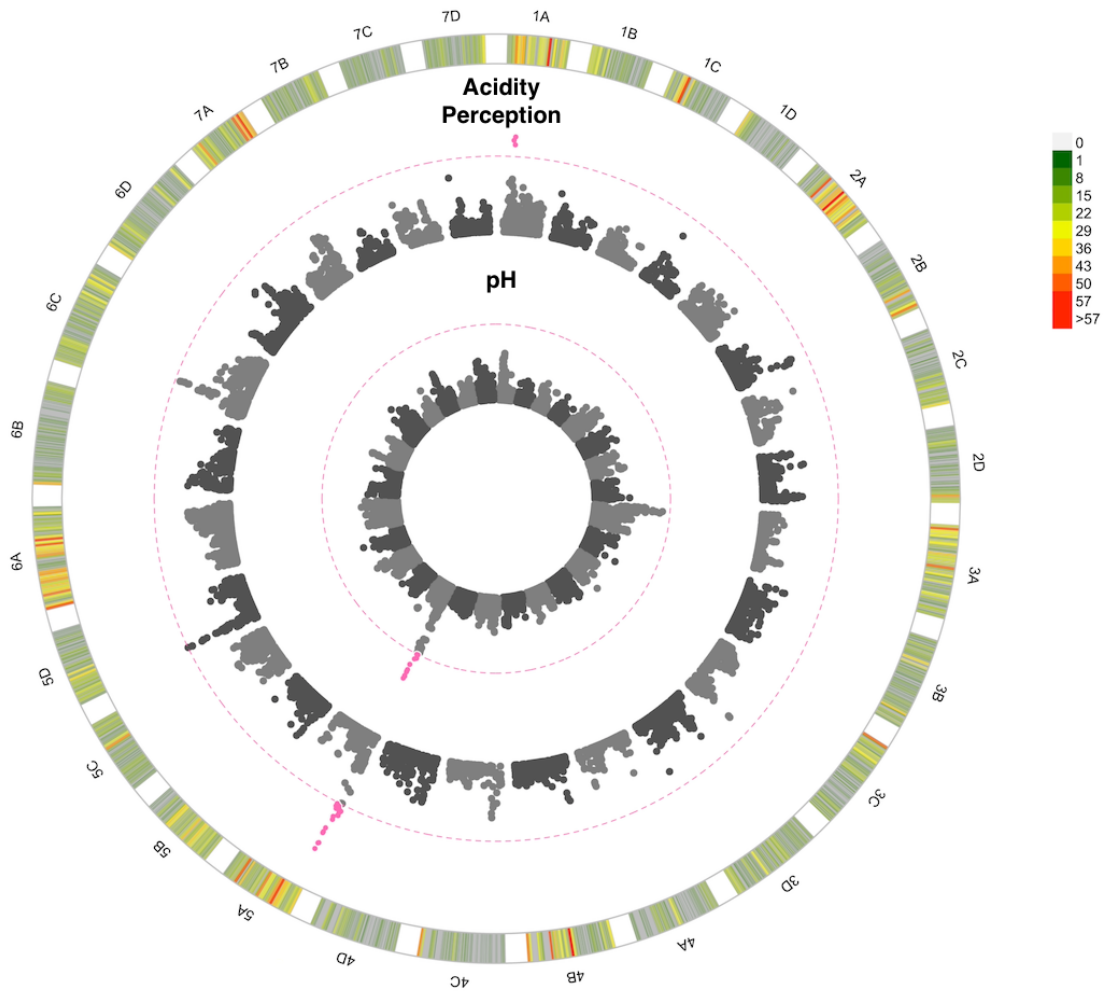
293

294 QTL identification

295 A total of 141 QTL were detected across 10 of the 19 fruit quality and 7 of the 12 yield traits
296 measured (Suppl. Table 3). A wealth of results have been generated due to the large number of
297 phenotypes assessed, here we seek to highlight the notable results relating to the traits rated as the
298 most important for breeders.

299 **Acidity & pH**

300 A highly significant QTL was detected on chromosome 5A for acidity perception and pH
301 measurements (Figure 5). This QTL was represented by the same focal SNP (Suppl. Table 3).
302 Detection of the QTL was greater for the subjective trait of acidity perception, furthermore, there
303 was no significant effect of assessor. These results indicate that acidity was perceived consistently
304 between individuals and thus human perception may act as a robust descriptor for strawberry
305 acidity (Suppl. Table 2).



306

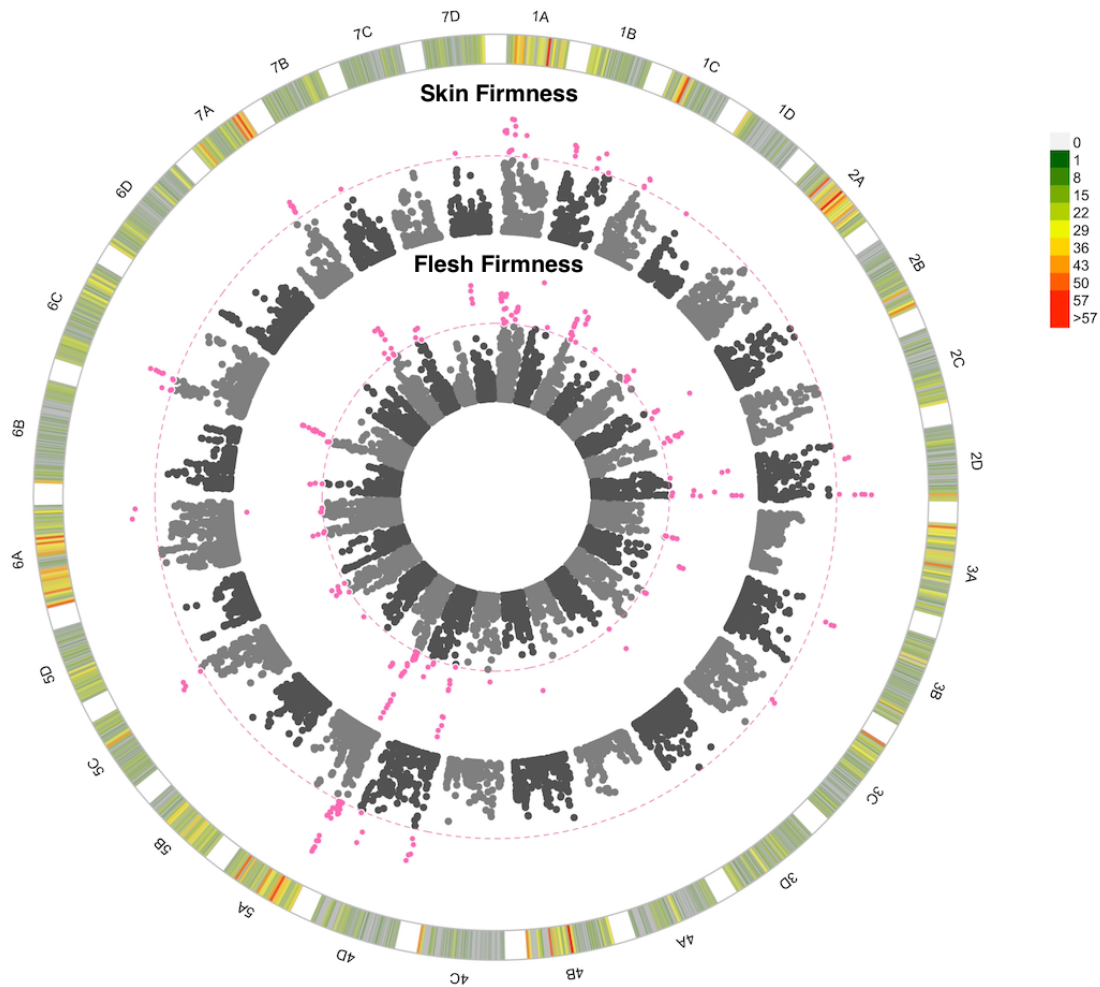
307

308 **Figure 5** Manhattan plot of GWAS looking at the association between SNPs and strawberry
309 acidity. 1A to 7D represent the 28 chromosomes of the strawberry genome. The inner Manhattan
310 plot represents acidity perception, the outer plot represents pH. The pink dotted line represents
311 Bonferroni correction at $-\log_{10} p = 7.14$ pink points are those which pass the significance
312 threshold. Marker positions are scaled to the *Fragaria vesca* genome v.4 (46). The colour coded
313 key in the outermost circle represents the number of SNPs segregating at each point across the
314 chromosome.

315

316 **Fruit firmness**

317 A total of 24 and 15 QTL were found to represent flesh and skin firmness, respectively. These
318 QTL are particularly notable - as both firmness traits are rated as 8 out of 9 for importance. Many
319 of the skin and flesh firmness QTL co-localise, with 4 of shared QTL improving both traits
320 simultaneously whereas 2 QTL impact upon the traits antagonistically (Figure 6). Flesh firmness
321 has a predictive accuracy of 0.54 and skin firmness has a predictive accuracy of 0.46 indicating
322 that a genomic prediction approach would be beneficial for improving fruit firmness in this
323 population (Suppl. Table 2). The R^2 illustrates the proportion of variation explained by the
324 identified QTL; the R^2 values for firmness traits were both greater than 40%, indicating a large
325 proportion of variation can be explained by the identified QTL (Suppl. Table 2). By contrast,
326 automated firmness measures (although positively correlated with other firmness measures) did
327 not reveal any QTL.



328

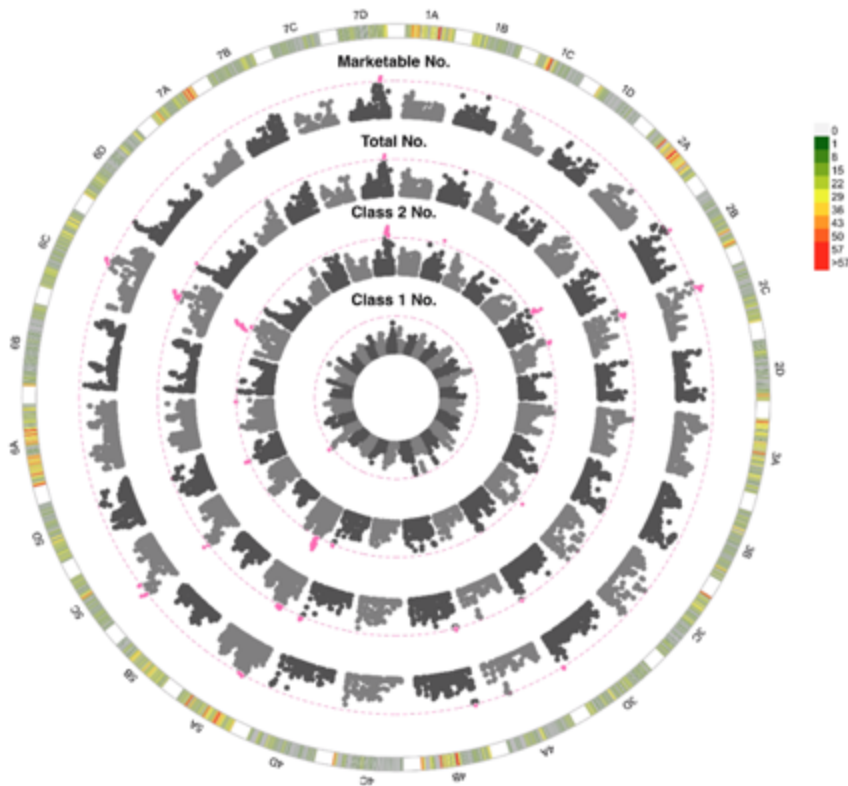
329

330 **Figure 6** Manhattan plot of GWAS looking at the association between SNPs and strawberry fruit
331 firmness. 1A to 7D represent the 28 chromosomes of the strawberry genome. The inner Manhattan
332 plot represents flesh firmness, the outer plot represents skin firmness. The pink dotted line
333 represents Bonferroni correction at $-\log_{10} p = 7.14$, pink points are those which pass the
334 significance threshold. Marker positions are scaled to the *Fragaria vesca* genome v.4 (46). The
335 colour coded key in the outermost circle represents the number of SNPs segregating at each point
336 across the chromosome.

337 **Yield and Class**

338 Several QTL were associated with variation in the number of fruits (Supplementary Figure 2).
339 Notably one QTL, represented by a single significant focal SNP, on chromosome 5C was found to
340 be associated with an 11% increase in the number of class one fruits, indicating an associated
341 improvement in fruit size and/or quality. This class one specific QTL was also associated with an
342 increase in marketable fruit and overall fruit number however it was not associated with an increase
343 in class 2 fruit. Two copies of the focal SNP were found in 17 of the progenitors, with a single
344 copy in the remaining progenitors, illustrating the SNP is abundant in the germplasm studied and
345 could be targeted through MAS to improve the quantity of high-class fruit. Furthermore, when
346 comparing yield traits, the number of marketable fruit was shown to have the greatest importance,
347 as measured by breeding priorities, and also the greatest genetic component as measured by
348 prediction accuracy, heritability and QTL number (Figure 7). These results indicate that the
349 number of marketable fruit would be the best trait to pursue and select upon if using a genomic
350 selection approach. By contrast, mass traits were associated with fewer QTL with the exception of
351 class 2 mass (Suppl. Table 3). The lack of total strawberry mass QTL may be explained by the
352 large influence of environmental factors upon the mass of berries.

353



354

355

356 **Supplementary Figure 2** Manhattan plot of GWAS looking at the association between SNPs and
357 number of strawberries. 1A to 7D represent the 28 chromosomes of the strawberry genome. The
358 inner Manhattan plot represents class one number, followed by class 2 number and total number
359 with the outermost plot representing marketable number. The pink dotted line represents
360 Bonferroni correction at $-\log_{10} p = 7.14$, pink points are those which pass the significance
361 threshold. Marker positions are scaled to the *Fragaria vesca* genome v.4 (46). The colour coded
362 key in the outermost circle represents the number of SNPs segregating at each point across the
363 chromosome.

364

365 **Traits without associated QTL**

366 No QTL were found for many of the subjective traits: aroma, sweetness perception, overall rating
367 of texture, skin colour, flavour and glossiness. Similarly, no QTL were found for several objective
368 traits: brix, objective firmness, truss number, shape (height: width ratio). The correction threshold

369 was very stringent, thus eliminating the possibility of false positive QTL. Truss number has a high
 370 broad sense heritability (90 %) indicating a highly heritable trait and yet a lower narrow sense
 371 heritability (26%) with no QTL detected, indicating that the trait may have a highly polygenic
 372 nature or potentially involves complex epigenetic interactions. The prediction ability values
 373 (Suppl. Table 2) indicate a genomic prediction approach may be used to enhance some of these
 374 traits.
 375

Visual assessments	1	2	3	4	5	6	7	8	9
Shape	Misc	Bi-conic	Globose	Globose-Conic	Conic	Long conic	Conic-Wedge	Wedge	Square/Oblong
Achene Density	V. seedy	V.seedy-seedy	Seedy	Seedy-Medium	Medium	Medium-Sparse	Sparse	Sparse-V sparse	V. sparse
Achene Colour	All dark red	Mostly dark	75% dark	51-74% red	50:50 yellow:red	51-74% yellow	75% yellow	Mostly yellow	All yellow
Achene Position	V. pitted (sunken)	Quite pitted	Slightly pitted (sunken)	Very slightly pitted	On surface	Very slightly raised	Slightly raised	Quite raised	V. raised
Glossiness	V. dull	Quite dull	Fairly dull	Slightly dull	Medium	Slightly glossy	Quite glossy	Glossy	V. glossy
Skin Colour	Pale orange	Orange	Orange-red	Paler red	Mid red	Mid to brick red	Brick-red	Dark red	Wine red
Calyx Position	Tightly clasped	Slightly clasped	Flat calyx	Slightly inflexed	Fully inflexed				
Neck Position	Very sunken	Slightly sunken	Flat	Slightly raised	Very raised				
Skin Firmness	V. fragile	V.fragile-fragile	Fragile	Fragile - medium	Medium	Medium - strong	Strong	Strong to v. strong	V. strong
Flesh Firmness	V. soft	V. soft to soft	Soft	Soft -Medium	Medium	Medium - Firm	Firm	Firm -v. firm	V. firm
Flesh Colour	White	Yellow/orange	Pale red	Pale-mid red	Mid red	Mid-dark red	Dark	Dark-v. dark	V. dark
Sweetness Perception	None	Slightly	Low	Low-moderate	Moderate	Moderately sweet	Sweet	Sweet-v. sweet	V. sweet
Acidity Perception	None	Slightly	Low	Low-moderate	Moderate	Moderate-high	High	High-v. high	V. high
Aromatics	None	Small trace	V. slightly	Slightly	Some aromatics	Quite aromatic	Aromatic	Strongly aromatic	V. strongly aromatic
Flavour Perception	V. poor	Poor	Quite poor	Below average	Average	Average-Good	Good	Good-Excellent	Excellent
Texture Type	Slimy	Stringy	Woolly	Mealy	Acceptable	Quite meaty	Meaty	V. meaty	Too crunchy
Texture Rating	V. poor*	V. poor-poor	Poor	Poor-Average	Average	Average-Good	Good	Good-Excellent	Excellent

*V. poor =too woolly/mealy/stringy/crunchy

376
 377
 378 **Supplementary Table 1** Visual, textural and organoleptic trait category descriptors of
 379 strawberries. Texture type and shape were assessed as discrete ordinal categorical traits and
 380 provide context for Texture Rating and Height: Width measures respectively. Texture Type and
 381 Shape were not assessed for genetic components.

382

Trait	Block	Importance	H ²	h ²	Significance of Block	Significance of Date	Significance of Assessor	Significance of Genotype	GxE	QTL no PC Bonferroni p=0.001	R ² adjusted	GBLUP Prediction Accuracy	Prediction Ability
Truss Number	NA	7	0.90	0.26	NA	***	NA	***	***	0	NA	0.33	0.17
Skin Colour	D	7	0.58	0.26	***	***	NS	***	***	0	NA	0.32	0.16
Neck Position	DV	3	0.74	0.45	***	***	**	***	***	12	0.23	0.49	0.33
Achene Density	D	6	0.50	0.23	***	***	NS	***	***	3	0.16	0.38	0.18
Achene Colour	D	3	0.63	0.36	*	***	NS	***	***	2	0.16	0.36	0.22
Acidity Perception	D	7	0.49	0.27	NS	***	NS	***	***	2	0.19	0.29	0.15
Flesh Firmness	D	8	0.68	0.29	***	***	NS	***	***	24	0.33	0.54	0.29
Brix	D	7	0.54	0.19	***	***	NA	***	***	0	NA	0.35	0.15
Firmness - Instrument	D	1	0.32	0.09	***	***	***	***	***	0	NA	0.34	0.10
Calyx Position	DV	6	0.41	0.13	***	***	***	***	***	4	0.16	0.34	0.12
Achene Position	DV	6	0.69	0.44	***	***	***	***	***	8	0.26	0.47	0.31
Flavour Perception	DO	9	0.36	0.20	NS	***	***	***	***	0	NA	0.26	0.12
Aromatics	DO	4	0.03	0.02	***	***	***	***	***	0	NA	-0.02	0.00
Flesh Colour	DO	4	0.68	0.43	NS	***	***	***	***	20	0.34	0.46	0.30
Sweetness Perception	DO	7	0.41	0.16	NS	***	***	***	***	0	NA	0.21	0.08
Texture Rating	DO	7	0.37	0.13	***	***	***	***	***	0	NA	0.34	0.12
Skin Firmness	DV	8	0.40	0.13	***	***	***	***	***	15	0.31	0.46	0.17
Glossiness	DV	7	0.32	0.03	**	***	***	***	***	0	NA	0.13	0.02
pH	D	2	0.38	0.13	NS	***	NA	***	***	1	0.12	0.40	0.15
Shape (Height: Width)	D	5	0.30	0.05	***	***	NA	***	***	3	0.19	0.40	0.09
Class1 Mass	B	9	0.34	0.13	***	NA	NA	***	NA	0	NA	0.18	0.06
Class1 Number	B	9	0.23	0.15	.	NA	NA	***	NA	1	0.10	0.30	0.12
Waste number	B	7	0.22	0.15	*	NA	NA	***	NA	6	0.19	0.28	0.11
Waste Mass	B	8	0.08	0.08	*	NA	NA	***	NA	0	NA	0.22	0.06
Class2 Number	B	7	0.35	0.20	***	NA	NA	***	NA	9	0.2	0.33	0.15
Class2 Mass	B	8	0.36	0.17	***	NA	NA	***	NA	8	0.26	0.30	0.12
Total Number	B	7	0.36	0.17	NS	NA	NA	***	NA	14	0.31	0.34	0.14
Total Mass	B	8	0.36	0.12	***	NA	NA	***	NA	1	0.09	0.21	0.07
Marketable Number	B	9	0.35	0.18	NS	NA	NA	***	NA	11	0.27	0.33	0.14
Marketable Mass	B	9	0.39	0.12	***	NA	NA	***	NA	0	NA	0.18	0.06
Percentage Marketable	B	9	0.21	0.11	***	NA	NA	*	NA	0	NA	0.16	0.05

0 **** 0.001 *** 0.01 ** 0.05 ' 0.1 ' NS ' 1

383

384

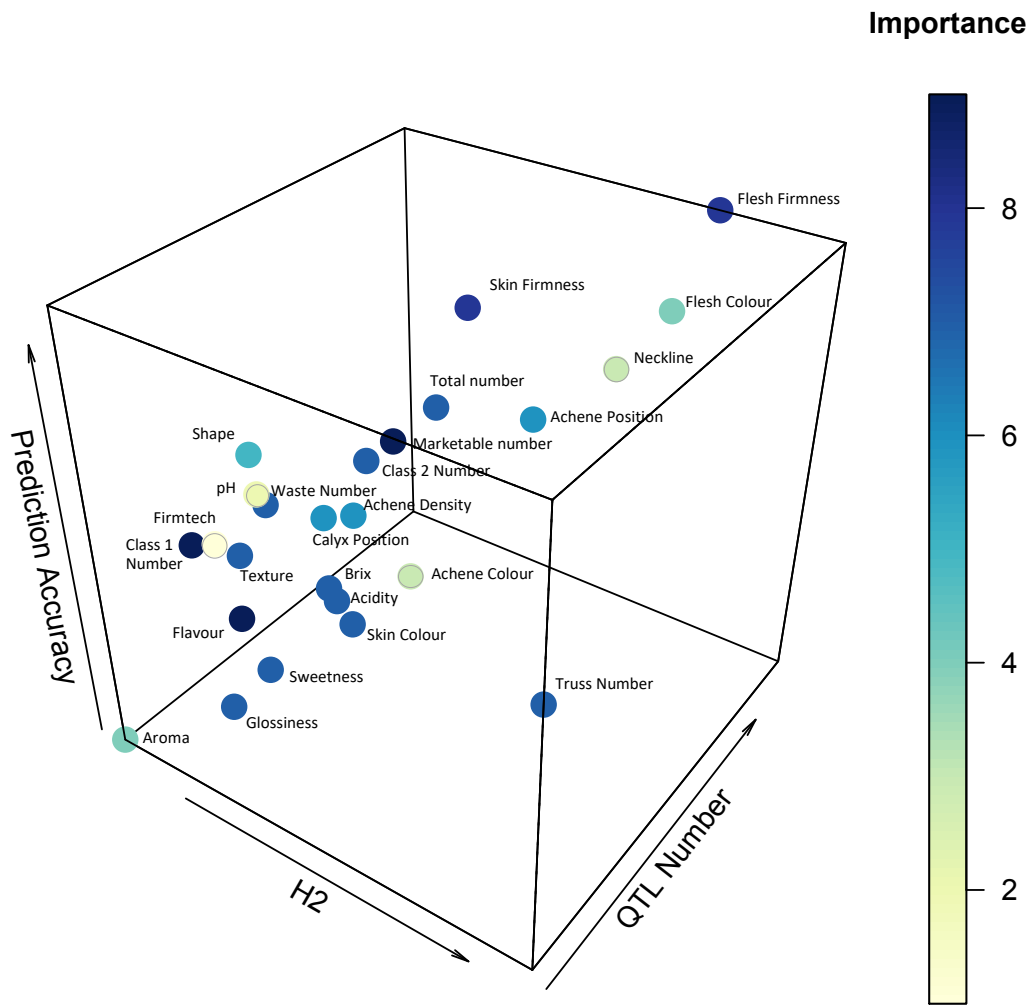
385 **Supplementary Table 2** Upper and lower bounds of broad sense heritability (H^2) and narrow
386 sense heritability (h^2) for strawberry fruit quality and yield traits across the multi-parental
387 population. Model denotes the BLUES model fitted per trait where the term DV represents date of
388 picking and visual recorder specified as random effects, DO represents date of picking and visual
389 recorder specified as random effects. Variation in date superseded variation in block. B represents
390 block specified as a random effect D represents date of picking specified as a random effect. All
391 prediction models were weighted by replicate number. The impact of block, picking date and
392 genome by environment interactions (GxE) on traits; significance values are ANOVA tests
393 comparing mixed models. p values are denoted by stars: *** < 0.001, ** < 0.01, * < 0.05, . < 0.01
394 NS - not significant. Importance denotes the importance in breeding on a scale from 1 (not
395 important) to 9 (highly important). The number of quantitative trait loci (QTL) identified through
396 GWAS after Bonferroni correction. The coefficient of determination (R^2) indicates the proportion
397 of variation explained by the combined QTL.

Chromosome	Marker name	Position (Mb)	Log ₁₀ P Value	Trait	Chromosome	Marker name	Position (Mb)	Log ₁₀ P Value	Trait
3B	AX-166512618	11.90	7.31	Ache Colour	4D	AX-89888668	3.26	14.03	Flesh Firmness
5A	AX-166524404	8.06	7.29	Ache Colour	5A	AX-123614613	3.48	14.26	Flesh Firmness
5B	AX-89788513	3.89	7.97	Ache Density	5B	AX-166514881	27.02	7.55	Flesh Firmness
5C	AX-89794999	10.67	7.97	Ache Density	5C	AX-166524140	27.30	8.56	Flesh Firmness
6A	AX-166516580	5.52	8.21	Ache Density	6A	AX-123365573	31.63	14.32	Flesh Firmness
1B	AX-123357020	1.39	8.74	Ache Position	6B	AX-123525554	6.73	8.35	Flesh Firmness
1C	AX-123366451	9.15	9.07	Ache Position	6C	AX-166525689	8.67	10.22	Flesh Firmness
2C	AX-166520935	23.85	7.26	Ache Position	7A	AX-166516936	18.49	10.16	Flesh Firmness
3A	AX-166504889	1.83	12.17	Ache Position	7B	AX-166526179	17.00	8.31	Flesh Firmness
3B	AX-166522083	27.09	7.15	Ache Position	7D	AX-123363650	7.54	10.86	Flesh Firmness
5C	AX-166514471	18.75	10.11	Ache Position	1C	AX-166516255	8.62	8.82	Neck Position
5D	AX-123364793	6.53	8.40	Ache Position	2A	AX-166503963	0.32	9.19	Neck Position
7B	AX-123364497	0.15	8.02	Ache Position	2B	AX-166520873	21.57	7.38	Neck Position
1A	AX-89875747	5.09	9.00	Acidity Perception	3D	AX-166504587	11.31	11.81	Neck Position
5A	AX-123525124	3.01	12.01	Acidity Perception	4A	AX-166523331	3.36	7.91	Neck Position
1B	AX-123365021	22.98	8.57	Calyx Position	4D	AX-89887238	21.56	10.05	Neck Position
5C	AX-166514700	23.52	7.78	Calyx Position	5A	AX-89858173	4.64	10.84	Neck Position
7A	AX-166526371	23.21	7.39	Calyx Position	5B	AX-123358636	0.27	7.81	Neck Position
7B	AX-166526371	23.21	7.39	Calyx Position	5C	AX-89872609	11.81	9.39	Neck Position
5C	AX-123361870	2.25	8.35	Class1 Number	6A	AX-123367361	20.22	9.95	Neck Position
2B	AX-89783653	7.13	7.45	Class2 Mass	6B	AX-166516553	0.99	8.75	Neck Position
3A	AX-166505049	6.02	7.76	Class2 Mass	7B	AX-166517297	0.15	9.20	Neck Position
3C	AX-166504244	23.79	7.28	Class2 Mass	1B	AX-89873309	15.52	7.89	Number Market
4D	AX-166523280	31.64	7.24	Class2 Mass	2B	AX-89783653	7.13	8.32	Number Market
5A	AX-166514409	13.71	7.88	Class2 Mass	2C	AX-123357423	4.31	8.60	Number Market
6C	AX-166515770	23.98	8.98	Class2 Mass	3A	AX-166505049	6.02	7.67	Number Market
6D	AX-166520085	2.17	8.02	Class2 Mass	3C	AX-166504183	16.74	7.19	Number Market
7D	AX-166518385	19.45	9.15	Class2 Mass	3D	AX-89826839	34.16	8.06	Number Market
1B	AX-89873309	15.52	9.22	Class2 Number	4B	AX-123367138	1.31	7.78	Number Market
2B	AX-89783653	7.13	9.15	Class2 Number	5A	AX-123361756	13.71	7.32	Number Market
2C	AX-123357423	4.31	7.92	Class2 Number	5C	AX-123361870	2.25	9.20	Number Market
3C	AX-166504244	23.79	9.30	Class2 Number	6C	AX-166515770	23.98	8.12	Number Market
4D	AX-166523280	31.64	7.63	Class2 Number	7D	AX-166518385	19.45	9.56	Number Market
5A	AX-123367281	7.69	8.76	Class2 Number	5A	AX-123525124	3.01	9.76	Ph
5D	AX-166506177	8.92	7.65	Class2 Number	1A	AX-166510826	3.44	10.65	Skin Firmness
6C	AX-166515770	23.98	10.34	Class2 Number	1B	AX-166502675	2.33	9.00	Skin Firmness
7D	AX-166527137	19.17	9.05	Class2 Number	1C	AX-166510504	9.91	8.31	Skin Firmness
1A	AX-166525522	8.68	10.40	Flesh Colour	1D	AX-89816903	2.29	7.16	Skin Firmness
1B	AX-123359925	17.97	8.26	Flesh Colour	2D	AX-166511667	26.34	10.44	Skin Firmness
1C	AX-166509612	14.44	10.74	Flesh Colour	3B	AX-166510079	8.99	9.12	Skin Firmness
1D	AX-166520330	4.30	8.41	Flesh Colour	3C	AX-166504505	12.34	7.66	Skin Firmness
2B	AX-166520993	27.02	7.57	Flesh Colour	4D	AX-89790990	3.72	10.26	Skin Firmness
2D	AX-123366894	25.78	7.34	Flesh Colour	5A	AX-123358616	3.13	12.61	Skin Firmness
3B	AX-89911919	26.02	7.89	Flesh Colour	5C	AX-89890707	21.22	9.54	Skin Firmness
3C	AX-123524621	12.30	7.23	Flesh Colour	6A	AX-123365573	31.63	9.34	Skin Firmness
3D	AX-123357787	26.91	8.87	Flesh Colour	6C	AX-166525682	8.88	9.73	Skin Firmness
4A	AX-166505532	15.74	10.41	Flesh Colour	7A	AX-166516933	18.42	8.91	Skin Firmness
4C	AX-123367100	26.71	7.51	Flesh Colour	7B	AX-123364491	11.41	7.47	Skin Firmness
5A	AX-123364118	7.26	7.37	Flesh Colour	7D	AX-123363650	7.54	7.61	Skin Firmness
5B	AX-123358397	15.47	11.14	Flesh Colour	3A	AX-166505049	6.02	7.56	Total Mass
5C	AX-166506190	12.95	7.48	Flesh Colour	1B	AX-89873309	15.52	7.72	Total Number
5D	AX-89784272	22.51	8.16	Flesh Colour	2C	AX-123357423	4.31	7.71	Total Number
6A	AX-123525365	24.98	10.51	Flesh Colour	3A	AX-166505049	6.02	8.75	Total Number
6C	AX-123366303	21.38	7.18	Flesh Colour	3C	AX-166504244	23.79	8.02	Total Number
6D	AX-123525365	24.98	10.51	Flesh Colour	3D	AX-89826839	34.16	7.53	Total Number
7A	AX-166527038	6.06	8.07	Flesh Colour	4B	AX-123367138	1.31	7.63	Total Number
7B	AX-166526605	0.08	9.10	Flesh Colour	4D	AX-166522841	32.62	7.79	Total Number
1A	AX-123360104	0.95	9.95	Flesh Firmness	5A	AX-123361756	13.71	7.38	Total Number
1B	AX-166502675	2.33	11.50	Flesh Firmness	5C	AX-123361870	2.25	8.95	Total Number
1C	AX-166502611	9.64	10.18	Flesh Firmness	5D	AX-166523870	19.35	7.55	Total Number
1D	AX-89816903	2.29	7.53	Flesh Firmness	6C	AX-166515770	23.98	8.68	Total Number
2A	AX-166511806	11.39	8.42	Flesh Firmness	6D	AX-166520085	2.17	7.53	Total Number
2B	AX-166520676	17.02	11.54	Flesh Firmness	7A	AX-166508748	23.50	7.42	Total Number
2C	AX-166521343	7.21	9.30	Flesh Firmness	7D	AX-166518385	19.45	9.72	Total Number
2D	AX-166511667	26.34	13.97	Flesh Firmness	3A	AX-166505049	6.02	7.17	Waste number
3A	AX-123363704	28.94	8.31	Flesh Firmness	3C	AX-166512335	23.74	7.34	Waste number
3B	AX-166510079	8.99	9.68	Flesh Firmness	4B	AX-166522765	15.23	8.51	Waste number
3C	AX-166522678	0.56	7.60	Flesh Firmness	4D	AX-166522841	32.62	8.02	Waste number
3D	AX-166522585	7.33	9.60	Flesh Firmness	6C	AX-166515613	20.25	8.05	Waste number
4B	AX-123358277	3.50	9.53	Flesh Firmness	7A	AX-166508748	23.50	7.64	Waste number
4C	AX-123358284	8.82	8.18	Flesh Firmness					

400 **Supplementary Table 3** QTL associated with strawberry yield and fruit quality traits identified
401 through a GWAS. Bold marker names were associated with multiple traits.

402 **Genetic architecture of traits**

403 Through plotting the importance of a trait as defined through breeding priorities against
404 heritability, predictive accuracy and number of QTL on a 3D scatter plot it was possible to visualise
405 the relative ability versus desire to improve yield and fruit quality traits within the study population
406 (Figure 7). The figure provides an indication of whether the observed variation is highly heritable
407 and whether it may be appropriate to adopt a genomic prediction or MAS breeding approach.
408 Explicitly, traits possessing high QTL numbers and high prediction accuracy values, such as flesh
409 firmness, are appropriate for selection using a genomic prediction breeding approach. By contrast,
410 traits possessing low QTL numbers (one or two) and high heritability may be suitable for MAS,
411 particularly where QTL effect sizes are high.



412

413 **Figure 7** Heritability (H^2), QTL number, and prediction accuracy for strawberry yield and fruit
414 quality traits as assessed across the multi-parental population. Dark blue represents the most
415 important traits to select upon, yellow the least important traits.

416

417

418 **Discussion**

419 **Trade-off Between Class One Yield and Soluble Sugar Content**

420 We confirm a well-established challenge for strawberry breeders: a trade-off was observed
421 between total soluble sugars and class one plant yield metrics in June-bearing plants grown under
422 a protected production system. Physiological or genetically linked trade-offs fundamentally limit
423 the possibility that some combinations of phenotypes can occur (47). Ultimately, the traits are
424 diametrically opposed, with the benefit gained by increasing the class one yield of strawberries,
425 associated with a cost that leads to reduced sugar content in the resulting berries. Conceptually,
426 should the mechanism be defined, gene editing offers a solution to overcome genetically linked
427 traits, unfortunately physiological trade-offs represent a potential “roadblock” in the pursuit of an
428 unattainable goal (47). Dividing a finite amount of sugar between a defined number of berries may
429 be considered a physiological trade-off. However, gene editing or extensive breeding can still
430 provide a solution; through the introduction of compounds that increase the perception of
431 sweetness and flavour without the need for sugars (16). Volatile organic compounds have a lower
432 carbon cost and can improve strawberry flavour perception (16) introduction of these compounds
433 into germplasm may become a critical component of mitigating the observed trade-off.

434
435 Further investigation is required to confirm the mechanism underpinning the relationship between
436 yield and sugar content. Nonetheless, other studies of strawberry have hinted at the existence of
437 this phenomenon, with a similar trade-off found in one out of three years across a biparental
438 population (10) and a 27% increase in yield associated with an 8% reduction in soluble sugars
439 (48). Our results indicate that breeders and strawberry plants alike may have to “decide” whether
440 to invest in a greater number of berries or produce a smaller number of higher sugar content berries,
441 with the elected strategy influencing both commercial success for the breeder and reproductive
442 success for the plant.

443

444

445 **Genetics informed breeding**

446 Here we study the power to breed for traits versus the relative importance in breeding for them.
447 Improving yield is a key goal of plant breeding. Our findings suggest that the *number of marketable*
448 *fruit* per plant may be the best trait to select upon when breeding for high cropping strawberry
449 varieties, particularly when using genomic prediction approaches. Enhancing the accuracy of
450 selection is a critical component for enhancing genetic gain (49). The only way improvement that
451 can be made via breeding is through selecting upon the variation that is caused by genetic
452 components. Therefore, selection of variation that is largely influenced by environmental
453 conditions (such as mass) will lead to lower genetic gain. It must be acknowledged that mass traits
454 were more influenced by environmental components and had lower narrow sense heritability
455 scores. As such, using mass traits for yield selection is associated with a lower accuracy. We
456 therefore suggest that selecting based upon the number of marketable strawberries could improve
457 the accuracy of selection and thus lead to greater genetic gain. However, in order to prevent
458 selection for smaller and yet marketable berries it is recommended that breeders increase the
459 threshold for acceptable berries.

460 **Environmental Influence on Fruit Quality**

461 Homeo-QTL, whereby QTL were located at the same physical position across different sub
462 genomes, have been identified in previous studies for fruit shape, size, glucose content, pH, malate
463 content and firmness traits (50). The researchers found that different QTL homologs were
464 expressed under different environmental conditions. Therefore, it was hypothesised that, as fruit
465 quality is an important trait associated with reproductive success, and that multiple gene homologs
466 remain functional. Environmental variation has a large impact on strawberry fruit production,
467 indeed, some cultivars of strawberries grown under high temperatures, have been shown to
468 produce lower yields (51) and poorer flavour (52). Our experimental setup, whereby blocks were
469 temporally separated across the season, prohibits homeo-QTL detection but allows us to mitigate
470 the significant impact of environmental variation on traits (Suppl. Table 1) and thus strengthens
471 the ability to detect stable alleles operational across multiple environments.

472

473 **Increasing Class One Yield**

474 We highlight a commercially relevant QTL associated with an 11% increase in class one fruit
475 number. Here we have used a diverse multi-parental population generated from temperate
476 European germplasm, therefore linkage between the trait and the associated QTL can be seen to
477 be conserved across germplasm. Past work using very sparse linkage maps have been able to
478 identify weak signals of QTL controlling fruit number on a number of chromosomes including
479 chromosome 5 (10). This may be reflected in our findings, but crucially, our analysis used a large
480 number of SNPs and has provided a fine scale resolution of the region of interest. Dissection of
481 the components which underlie the class one category will reveal the biologically relevant
482 attributes believed to result in higher class one yield: fruit size, truss architecture or truss number.

483

484 **Flavour**

485 The use of a multi-parental population has the advantage over biparental QTL mapping studies as
486 it allows the assessment of genetic components across diverse germplasm. A similar analysis has
487 been conducted across a multi-parental population in strawberry where multiple QTL were
488 identified for titratable acidity, pH and total soluble sugars, (61) multiple QTL for pH were found
489 in a biparental study, one of which was on chromosome 5B (50). However, the large effect acidity
490 perception and pH QTL was observed on linkage group 5A, and so may represent a novel source
491 of flavour that has not been reported in the literature previously. Others have characterised the
492 complex relationship between soluble sugar content and sweetness perception and how perception
493 can be influenced by volatiles (16). However, less has been reported on the relationship between
494 acidity and acidity perception and our finding suggests the relationship could be more
495 straightforward.

496 **Fruit Firmness**

497 Firmness is an essential component of fruit quality which is linked to increased shelf life, lower
498 mechanical injury and reduced susceptibility to storage rots (53,54). Overall, breeders aim for an
499 intermediate level of firmness, striking a balance between durability and a desirable eating texture.
500 The identified fruit firmness QTL accounted for a large proportion of the variation observed in the

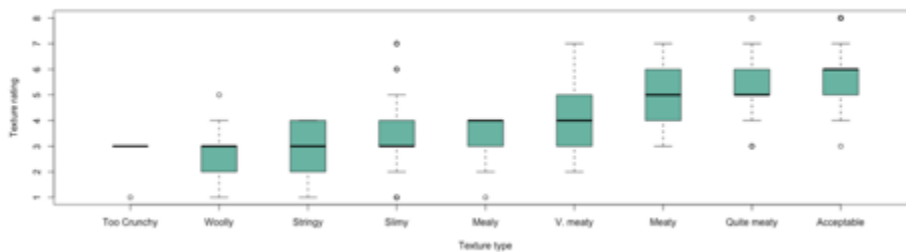
501 multi-parental population. Therefore, firmness is likely to show improvement through the adoption
502 of genomic prediction approaches.

503

504 A non-destructive, firmness measuring instrument was used to produce an objective measure of
505 fruit firmness. However, these measures were not associated with high heritability, predictive
506 ability nor QTL number. Such inconsistent results between methods of measuring strawberry
507 firmness have been well documented (55), and our results highlight the difficulty associated with
508 objective measurement of this trait. We confirm that tactile human perception can be used as a
509 robust measure to assist the genetic guided improvement of skin and flesh firmness. Destructive
510 penetrometer instruments may be more effective in capturing human perceived firmness
511 particularly where injury to the fruit is not prohibited due to downstream assessment requirements.

512

513 Firmness is not only important for longevity, but also related to strawberry texture in a nonlinear
514 fashion; here texture type was recorded alongside the texture rating, and we see that texture types
515 from across the firmness spectrum score low texture ratings i.e., “woolly”, “slimy”, “stringy” and
516 “too crunchy” (Sup. Figure 3). Limited genetic studies have been conducted on strawberry texture,
517 and this may be due to the complexities associated with quantifying the trait. Nonetheless, texture
518 has been reported to play a significant role in the overall fruit quality score of strawberries (56),
519 therefore desirable texture of strawberries must continue to be selected for in spite of the associated
520 challenges.



521

522

523 **Supplementary Figure 3** Subjective overall texture rating for each strawberry texture type

524

525

526 **Fruit Shape**

527 The height / width (H/W) ratio can be used to discriminate between some strawberry shape types,
528 particularly long conic fruit. However, the H/W ratio did not segregate desirable and undesirable
529 fruit shapes into discrete groups and so cannot be used as a straightforward metric to select for
530 fruit shape. This is because the breeders' definition of desirable strawberry shape does not align
531 with the H/W measure. More comprehensive methods of fruit shape quantification have been
532 conducted through the use of machine learning approaches (57) alongside 3D imaging studies
533 describing fruit uniformity (58). Strawberry shape has been studied extensively in the diploid
534 strawberry *F. vesca* and the genes responsible for controlling the height and width of the berries
535 have been identified (59,60) Plant hormones have been shown to define fruit shape, with auxin
536 boosting the width of receptacle expansion, GA increasing height and ABA inhibiting overall
537 expansion (59,60). Further work may determine whether similar genetic components control the
538 complexities of fruit shape in octoploid strawberry.

539 **Conclusions**

540 Through studying the genetic architecture of strawberry traits, we conclude that selecting upon the
541 number of marketable fruit produced per plant may lead to the production of high yielding
542 strawberry varieties. We show that subjective human scores of firmness and acidity were superior
543 to surrogate measures of non-destructive instruments and pH meters and recommend the
544 implementation of genomic prediction and MAS to capture the observed variation, respectively.
545 Finally, we highlight the dilemma faced by many strawberry breeders: greater class one yield or
546 sugar content?

547 **Acknowledgements**

548 The authors acknowledge project partners Soloberry, Sainsburys, Botanicair and Agrovista for
549 their involvement and support of the project. The authors acknowledge Dr Robert Vickerstaff for
550 generating the octoploid consensus map as part of other projects and Dr Beatrice Denoyes, INRA
551 and Dr Amparo Monfort, CRAG for granting the use of their informative markers in the production
552 of the strawberry consensus linkage map and Dr Daniel Sargent for helpful comments and editorial

553 suggestions. We also acknowledge and thank the many field staff and visiting workers that assisted
554 with phenotyping. The authors acknowledge funding from the Biotechnology and Biological
555 Sciences Research Council (BBSRC) BB/M01200X/2, BB/P005039/1 and Innovate UK project
556 101914.

557 **Conflicting Interests**

558 On behalf of all authors, the corresponding author states that there is no conflict of interests
559 regarding the publication of this work.

560 **Contributions**

561 AJ, HMC, BL, ES, RJH - Conceived and designed experiments

562 HMC – Conducted quantitative genetics analysis

563 AJ, HMC, KH, AW, AK, BL - Performed experiments

564 AK - Performed genotyping & wrote initial GBLUP script

565 BL - Performed image analysis

566 HMC wrote the manuscript with contributions from all authors.

567 **List of Abbreviations**

568 i35k - Istraw35 Affymetrix chip

569 GEBV - Genomic Estimated Breeding Value

570 GWAS - Genome Wide Association Study

571 QTL- Quantitative Trait Loci

572 QR - Quick Response

573 SNP - Single Nucleotide Polymorphism

574

575 **References**

576

- 577 1. Beal WJ. Some unique examples of dispersion of seeds and fruits. *The American*
578 *Naturalist*. 1898 Nov 1;32(383):859-66.
- 579 2. McAtee WL. Distribution of seeds by birds. *American Midland Naturalist*. 1947
580 Jul;38(1):214.
- 581 3. Nakamura S. Germination of strawberry seeds. *Engei Gakkai zasshi*. 1972;41(4):367-75.
- 582 4. Vazačová K, Münzbergová Z. Simulation of seed digestion by birds: how does it reflect
583 the real passage through a pigeon's gut? *Folia Geobot*. 2013 Jun;48(2):257-69.
- 584 5. Schaefer HM, Levey DJ, Schaefer V, Avery ML. The role of chromatic and achromatic
585 signals for fruit detection by birds. *Behavioral Ecology*. 2006 Sep 1;17(5):784-9.
- 586 6. Du X, Song M, Rouseff R. Identification of New Strawberry Sulfur Volatiles and Changes
587 during Maturation. *J Agric Food Chem*. 2011 Feb 23;59(4):1293-300.
- 588 7. Ménager I, Jost M, Aubert C. Changes in Physicochemical Characteristics and Volatile
589 Constituents of Strawberry (Cv. Cigaline) during Maturation. *J Agric Food Chem*. 2004
590 Mar;52(5):1248-54.
- 591 8. Rodríguez A, Alquézar B, Peña L. Fruit aromas in mature fleshy fruits as signals of
592 readiness for predation and seed dispersal. *New Phytol*. 2013 Jan;197(1):36-48.
- 593 9. Hummer KE, Hancock J. Strawberry genomics: botanical history, cultivation, traditional
594 breeding, and new technologies. In: Folta KM, Gardiner SE, editors. *Genetics and*
595 *genomics of rosaceae*. New York, NY: Springer New York; 2009. p. 413-35.
- 596 10. Zorrilla-Fontanesi Y, Cabeza A, Domínguez P, Medina JJ, Valpuesta V, Denoyes-Rothan
597 B, et al. Quantitative trait loci and underlying candidate genes controlling agronomical and
598 fruit quality traits in octoploid strawberry (*Fragaria × ananassa*). *Theor Appl Genet*. 2011
599 Sep;123(5):755-78.
- 600 11. Klee HJ, Tieman DM. The genetics of fruit flavour preferences. *Nat Rev Genet*.
601 2018;19(6):347-56.
- 602 12. Cardello AV. Food quality: Relativity, context and consumer expectations. *Food Qual*
603 *Prefer*. 1995 Jan;6(3):163-70.
- 604 13. Brennan R, Graham J. Improving fruit quality in *Rubus* and *Ribes* through breeding. *Func*
605 *Plant Sci Biotech*. 2009;3(1):22-9.
- 606 14. Mitcham B. Quality assurance for strawberries: a case study. *Perishables Handling*
607 *Newsletter*. 1996;85:6-9.
- 608 15. Ikegaya A, Toyozumi T, Ohba S, Nakajima T, Kawata T, Ito S, et al. Effects of
609 distribution of sugars and organic acids on the taste of strawberries. *Food Sci Nutr*. 2019
610 Jul;7(7):2419-26.
- 611 16. Schwieterman ML, Colquhoun TA, Jaworski EA, Bartoshuk LM, Gilbert JL, Tieman DM,
612 et al. Strawberry flavor: diverse chemical compositions, a seasonal influence, and effects
613 on sensory perception. *PLoS ONE*. 2014 Feb 11;9(2):e88446.
- 614 17. Chambers AH, Pillet J, Plotto A, Bai J, Whitaker VM, Folta KM. Identification of a
615 strawberry flavor gene candidate using an integrated genetic-genomic-analytical chemistry

- 616 approach. *BMC Genomics*. 2014 Apr 17;15:217.
- 617 18. Sanz C, Richardson DG, Pérez AG. 2,5-Dimethyl-4-hydroxy-3(2*H*)-furanone and
618 Derivatives in Strawberries During Ripening. In: Rouseff RL, Leahy MM, editors. *Fruit*
619 *flavors: biogenesis, characterization, and authentication*. Washington, DC: American
620 Chemical Society; 1995. p. 268–75.
- 621 19. Lado J, Inés Moltini A, Vicente E, Rodríguez G, Arcia P, Rodríguez M, et al. Integration
622 of sensory analysis into plant breeding: review. *Agrociencia Uruguay*. 2019 Jun 1;23(01).
- 623 20. Lawless HT, Heymann H. *Sensory evaluation of food: principles and practices*. Springer
624 Science & Business Media; 2013 Dec 11.
- 625 21. Migicovsky Z. Tasting improvement in fruit flavor using genomics. *New Phytol*. 2020 Apr
626 28;
- 627 22. Diehl DC, Sloan NL, Bruhn CM, Simonne AH, Brecht JK, Mitcham EJ. Exploring
628 Produce Industry Attitudes: Relationships between Postharvest Handling, Fruit Flavor, and
629 Consumer Purchasing. *Horttechnology*. 2013 Oct;23(5):642–50.
- 630 23. Ruth TK, Rumble JN. Branding the berries: consumers' strawberry purchasing intent and
631 their attitude toward Florida strawberries. *Journal of Applied Communications*. 2016 Jun
632 1;100(2):28-43.
- 633 24. Colquhoun TA, Levin LA, Moskowitz HR. Framing the perfect strawberry: An exercise in
634 consumer-assisted selection of fruit crops. *J Berry Res*. 2012;2(1):45–61.
- 635 25. Wang X, Xu Y, Hu Z, Xu C. Genomic selection methods for crop improvement: Current
636 status and prospects. *Crop J*. 2018 Apr;6(4):330–40.
- 637 26. Karlström A, Cabo-Medina M, Harrison R. Advances and challenges in apple breeding' in
638 Lang, G. A. (ed.), Volume 2: . In: Lang GA, editor. *Achieving sustainable cultivation of*
639 *temperate zone tree fruits and berries*. Burleigh Dodds Science Publishing; 2019.
- 640 27. Hayes BJ, Bowman PJ, Chamberlain AJ, Goddard ME. Invited review: Genomic selection
641 in dairy cattle: progress and challenges. *J Dairy Sci*. 2009 Feb;92(2):433–43.
- 642 28. Swan AA, Johnston DJ, Brown DJ, Tier B, Graser H-U. Integration of genomic
643 information into beef cattle and sheep genetic evaluations in Australia. *Anim Prod Sci*.
644 2012;52(3):126–32.
- 645 29. Cleveland MA, Hickey JM. Practical implementation of cost-effective genomic selection
646 in commercial pig breeding using imputation. *J Anim Sci*. 2013 Aug;91(8):3583–92.
- 647 30. Wolc A, Stricker C, Arango J. Breeding value prediction for production traits in layer
648 chickens using pedigree or genomic relationships in a reduced animal model. *Genet Sel*
649 *Evol*. 2011;43(5).
- 650 31. Gezan SA, Osorio LF, Verma S, Whitaker VM. An experimental validation of genomic
651 selection in octoploid strawberry. *Hortic Res*. 2017 Jan 11;4:16070.
- 652 32. Savini G, Neri D, Zucconi F, Sugiyama N. Strawberry growth and flowering. *Int J of Fruit*
653 *Sci*. 2005;5:29–50.
- 654 33. Rife TW, Poland JA. *Field Book: An Open-Source Application for Field Data Collection*
655 *on Android*. *Crop Sci*. 2014;54(4):1624–7.

- 656 34. Li B, Cockerton HM, Johnson AW, Karlström A, Stavridou E, Deakin G, et al. Defining
657 strawberry shape uniformity using 3D imaging and genetic mapping. *Hortic Res.* 2020
658 Dec;7(1):115.
- 659 35. Verma S, Bassil NV, E Van De Weg, et al. Development and evaluation of the Axiom®
660 IStraw35 384HT array for the allo-octoploid cultivated strawberry *Fragaria* × *ananassa*.
661 *Acta Hortic.* 2017;1156:75–82.
- 662 36. Cockerton HM, Vickerstaff RJ, Karlström A, Wilson F, Sobczyk M, He JQ, et al.
663 Identification of powdery mildew resistance QTL in strawberry (*Fragaria* × *ananassa*).
664 *Theor Appl Genet.* 2018 Sep;131(9):1995–2007.
- 665 37. van Dijk T, Pagliarani G, Pikunova A, Noordijk Y, Yilmaz-Temel H, Meulenbroek B, et
666 al. Genomic rearrangements and signatures of breeding in the allo-octoploid strawberry as
667 revealed through an allele dose based SSR linkage map. *BMC Plant Biol.* 2014 Mar
668 1;14:55.
- 669 38. Sargent DJ, Yang Y, Šurbanovski N, Bianco L, Buti M, Velasco R, et al. HaploSNP
670 affinities and linkage map positions illuminate subgenome composition in the octoploid,
671 cultivated strawberry (*Fragaria* × *ananassa*). *Plant Sci.* 2016 Jan;242:140–50.
- 672 39. Kruijer W, Flood P, Kooke R. heritability: Marker-based estimation of heritability using
673 individual plant or plot data. URL [http://CRAN.R-project.org/package= heritability](http://CRAN.R-project.org/package=heritability), r
674 package version. 2016;1.
- 675 40. Granato I, Fritsche-Neto R. snpReady: Preparing genotypic datasets in order to run
676 genomic; analysis. R package version 0.9. 6. R; 2018.
- 677 41. Revelle W. psych: Procedures for psychological, psychometric, and personality research. R
678 package version. 2018 Mar;1(10).
- 679 42. Wei T, Simko V. R package “corrplot”: Visualization of a Correlation Matrix (Version
680 0.84). R; 2017.
- 681 43. Endelman JB. Ridge Regression and Other Kernels for Genomic Selection with R Package
682 rrBLUP. *Plant Genome.* 2011;4(3):250–5.
- 683 44. Yu J, Pressoir G, Briggs WH, Vroh Bi I, Yamasaki M, Doebley JF, et al. A unified mixed-
684 model method for association mapping that accounts for multiple levels of relatedness. *Nat*
685 *Genet.* 2006 Feb;38(2):203–8.
- 686 45. Erbe M, Pimentel ECG, Sharifi AR, Simianer H. Assessment of cross-validation strategies
687 for genomic prediction in cattle. 2010;
- 688 46. Li Y, Pi M, Gao Q, Liu Z, Kang C. Updated annotation of the wild strawberry *Fragaria*
689 *vesca* V4 genome. *Hortic Res.* 2019 May 1;6:61.
- 690 47. Weih M. Trade-offs in plants and the prospects for breeding using modern biotechnology.
691 *New Phytologist.* 2003 Apr;158(1):7–9.
- 692 48. Whitaker VM, Osorio LF, Hasing T. Estimation of Genetic Parameters for 12 Fruit and
693 Vegetative Traits in the University of Florida Strawberry Breeding Population. *J Am Soc*
694 *Hortic Sci.* 2012 Sep;137(5):316–24.
- 695 49. Cobb JN, Declerck G, Greenberg A, Clark R, McCouch S. Next-generation phenotyping:
696 requirements and strategies for enhancing our understanding of genotype-phenotype

- 697 relationships and its relevance to crop improvement. *Theor Appl Genet*. 2013
698 Apr;126(4):867–87.
- 699 50. Lerceteau-Köhler E, Moing A, Guérin G, Renaud C, Petit A, Rothan C, et al. Genetic
700 dissection of fruit quality traits in the octoploid cultivated strawberry highlights the role of
701 homoeo-QTL in their control. *Theor Appl Genet*. 2012 Apr;124(6):1059–77.
- 702 51. Sun P, Mantri N, Lou H, Hu Y, Sun D, Zhu Y, et al. Effects of elevated CO₂ and
703 temperature on yield and fruit quality of strawberry (*Fragaria × ananassa* Duch.) at two
704 levels of nitrogen application. *PLoS ONE*. 2012 Jul 24;7(7):e41000.
- 705 52. Wang SY, Camp MJ. Temperatures after bloom affect plant growth and fruit quality of
706 strawberry. *Sci Hortic*. 2000 Aug;85(3):183–99.
- 707 53. Mass JL. Screening for resistance to fruit rot in strawberries and red raspberries: a review.
708 *Hortsci*. 1978;
- 709 54. Hietaranta T, Linna M-M. Penetrometric measurement of strawberry fruit firmness: device
710 testing. *Horttechnology*. 1999 Jan;9(1):103–5.
- 711 55. Døving A, Måge F, Vestrheim S. Methods for testing strawberry fruit firmness. *Small*
712 *Fruits Review*. 2005 Mar;4(2):11–34.
- 713 56. Cockerton H, Blanco Unzueta M, Johnson A, Vadillo Dieguez A, Fernández Fernández F.
714 Pathway analysis to determine factors contributing to overall quality scores in four berry
715 crops. *Journal of Horticultural Research*. 2020 Dec 28 (2): 35–42.
- 716 57. Feldmann MJ, Hardigan MA, Famula RA, Lopez CM, Tabb A, Cole GS, Knapp SJ. Multi-
717 dimensional machine learning approaches for fruit shape phenotyping in strawberry.
718 *GigaScience*. 2020 May;9(5):giaa030.
- 719 58. Li B, Cockerton HM, Johnson AW, Karlström A, Stavridou E, Deakin G, Harrison RJ.
720 Defining strawberry shape uniformity using 3D imaging and genetic mapping. *Horticulture*
721 *research*. 2020 Aug 1;7(1):1-3.
- 722 59. Wang S, Li W, Liu Y, Li H, Ma Y, Zhang Z. Comparative transcriptome analysis of
723 shortened fruit mutant in woodland strawberry (*Fragaria vesca*) using RNA-Seq. *J Integr*
724 *Agric*. 2017 Apr;16(4):828–44.
- 725 60. Liao X, Li M, Liu B, Yan M, Yu X. Interlinked regulatory loops of ABA catabolism and
726 biosynthesis coordinate fruit growth and ripening in woodland strawberry. *Proc Natl Acad*
727 *Sci*. 2018;115(49):E11542–50.
- 728 61. Verma S, Zurn JD, Salinas N, Mathey MM, Denoyes B, Hancock JF, et al. Clarifying sub-
729 genomic positions of QTLs for flowering habit and fruit quality in U.S. strawberry
730 (*Fragaria × ananassa*) breeding populations using pedigree-based QTL analysis. *Hortic Res*.
731 2017 Nov 8;4:17062.